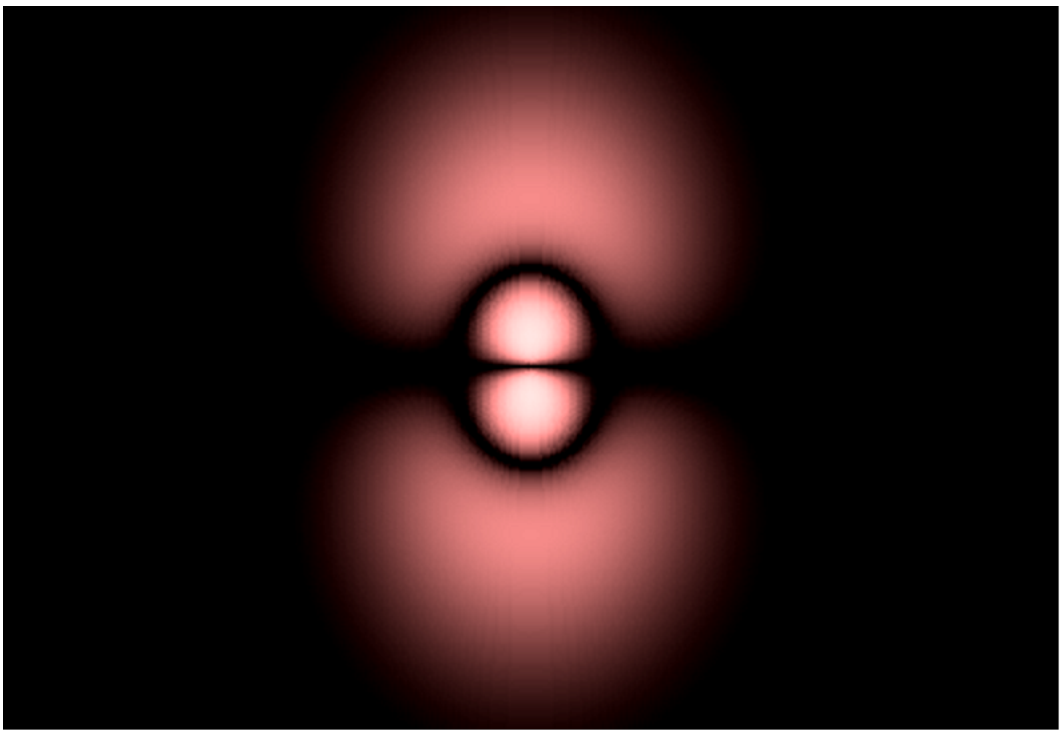
Quantum Chemistry

Dr. Marcel Drabbels



EPFL

Autumn 2025

Table of Contents

1 Introduction and Historical Perspective

1.1 General comments about the class

Dr. Marcel Drabbels Room: CH G0-613 Telephone: 33022

Assistants: Liam Marsh
Peter Gispert
Pratham Deshmukh

Documentation: Lecture notes and additional information are available on Moodle (http://moodle.epfl.ch/)

Reference Books

Primary Reference:

- D. A. McQuarrie, *Quantum Chemistry*

Secondary References:

- P. W. Atkins, *Molecular Quantum Mechanics*
- Cohen-Tannoudji, Diu, and Laloë, *Quantum Mechanics* (originally published in French)
- B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*

Exercises: **Solving problems is an essential part of the course**. I encourage you not only to come to the exercise sessions, but also to work on the problems before you come. Two hours each week is not enough to do and understand all the exercises.

This course will give you a basic introduction to the principles of quantum chemistry. In the time we have here we can only cover the basics, but once you understand these basics you will be able to go on and learn about more advanced topics on your own. It is therefore not important exactly how much material we cover in class. What I am most concerned about is that you understand the material that we do cover. However, this puts also a certain burden on you, *i.e*. you should ask questions if a subject or point made is not clear to you. I have found that students are often afraid to ask questions when they don't understand something. They feel they are going to pose a stupid question. To me there are no stupid questions. When you leave this room you should feel that you understand everything that I covered. It is your responsibility to speak up when you don't understand.

1.2 Importance and Usefulness of Quantum Chemistry

One can think about chemistry and physics on two levels. In your physics course up to now, you have learned about the physics of macroscopic objects. This is described by the fields of classical mechanics, electricity and magnetism, and geometrical optics. In a similar way, in your engineering courses you will learn about the bulk or macroscopic chemical properties of matter, *i.e.* mass and energy transport properties, heat capacity, viscosity, density, etc. Since we interact with the world around us at a macroscopic level, it is clearly important to treat chemistry and physics at this macroscopic level.

However, molecules are microscopic objects, and there are physical phenomena of molecules that only become apparent when one looks at microscopic dimensions. Some of these phenomena are very different than what one would expect from our experience in the macroscopic world.

From an engineering point of view, the microscopic realm is also becoming more and more important as all types of engineering attempt to measure and control processes on a microscopic level. Chemists and physicists have been doing “nanotechnology” for a long time insofar as molecules are microscopic objects that we attempt to design, fabricate, and manipulate.

Quantum chemistry, or more generally quantum mechanics, describes a field that treats the microscopic properties of matter. Classical mechanics, which is what you have been studying in your physics courses up to now, can describe physical and chemical systems very well on a macroscopic level but fails badly on a microscopic level. Quantum mechanics provides a framework that describes matter at a microscopic level but in the limit of large dimensions generalizes to classical mechanics. It is important for us, both as scientists and engineers, to understand the microscopic basis of the macroscopic theories.

1.3 Historical Perspective

Towards the end of the 19th century, many physicists felt that all of the basic principles of physics had been discovered, and there was little that was fundamentally new to be found.

- Newtonian Mechanics had become a mature branch of science (200 years old), particularly by the work of Hamilton and Lagrange.
- The field of Thermodynamics was essentially in the form that it is today.
- The work of Maxwell brought together many of the unanswered questions of electricity and magnetism, and unified it with optics.

These fields constitute what is considered *Classical Physics*.

In physics as well as in any field of science, the basic approach is to postulate a theory and then test the theory by comparison with experimental observations. If the theory fails to describe adequately the observed phenomena, it is either rejected or modified. When the theory gets to the point where it describes the entirety of experimental observations, it generally becomes accepted as being correct. Most of *classical physics* had reached this point by the end of the 19th century.

One of the basic assumptions of *classical physics* was that physical quantities such as energy, momentum, etc. could take on a continuous range of values. However, the beginning of the 20th century witnessed several important experiments that did not fit the classical picture. When enough such observations mount, the current theory is overturned and a new one replaces it. Hence, in the first quarter of this century, a revolution occurred in the world of physics, and the development of *Quantum Mechanics* was at the center of this revolution.

I would like to discuss briefly just a few of these early experiments that served to overturn much of classical physics.

These experiments involve:
- Blackbody Radiation
- The Photoelectric Effect
- The Line Spectra of Atoms

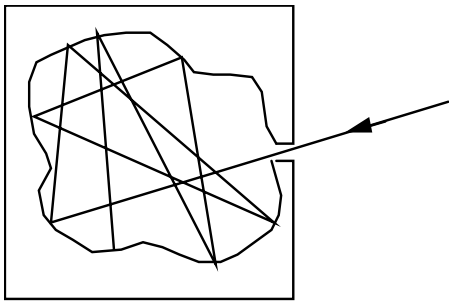
1.3.1 Blackbody Radiation (Planck - 1900)

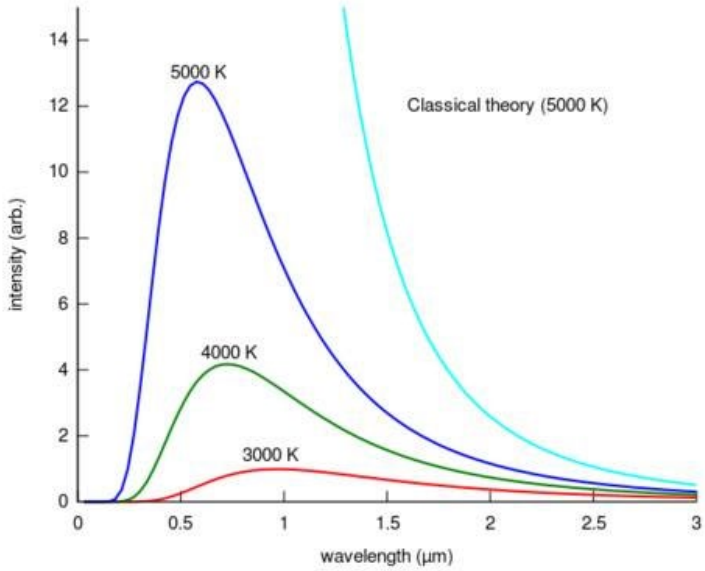
An ideal blackbody is an object that is perfectly black, *i.e*. it absorbs all wavelengths of light that fall onto it. A good model for a perfect black body is a cavity with a small hole to the outside.

Light that falls on the hole will undergo reflections in which it is partially absorbed and partially reflected. If the hole is small with respect to the area of the cavity wall, essentially all light that is incident on the hole will be totally absorbed. Therefore, the hole behaves just like the surface of a totally black body. At low temperatures, the hole looks black. It can be shown that it will also radiate like a perfect black body.

Now, if the cavity is heated, the hole will become self-luminous. This is because as the solid is heated the atoms vibrate and cause oscillations in the electrons which are responsible for the emission. The cavity walls will thus emit thermal radiation and some will escape the hole. Since the hole acts like a black surface, the emitted radiation will be characteristic of a perfect black body. The spectrum of light coming out of the hole can according

to Kirchhoff (1859) be described by a function $\rho(\nu,T)$, *i.e.* energy density in the frequency interval ν and $\nu+d\nu$, that depends only on the frequency of the emitted light and the temperature of the object.
Experimentally, that radiation looks something like this, where the energy density is given as function of the wavelength of the light :





You know from experience that when you heat things up hotter, the color changes from dull red, to bright yellow, even to blue, *i.e*. the peak frequency/wavelength shifts.

The problem comes in calculating this energy spectrum using the laws of classical physics. Early in 1900, this was done by Lord Rayleigh and Sir James Jeans as well as others, and their basic result was:

$$\rho(\nu,T)d\nu = \frac{8\pi kT\nu^2}{c^3}d\nu$$

where $\rho(\nu,T)$ is the energy density between ν and $\nu+d\nu$

This is called the Rayleigh-Jeans formula for blackbody radiation. If we compare this value to that observed experimentally, one finds that it does ok at low frequencies, but at high frequencies it fails miserably. This has been termed the ultraviolet catastrophe, since it fails in the ultraviolet region of the spectrum.

Later, in the same year, Max Planck was able to deduce a theoretical result that was in complete agreement with the experimental observations. He reasoned that some quantization phenomena must be occurring to give a maximum in the intensity distribution, *i.e.* the energy emitted by the electron oscillations could only take on discrete values:

$$E = n\Delta\varepsilon$$

He then assumed that the energy increment $\Delta\varepsilon$ is simply proportional to the frequency of the vibrating electron, hence:

$$\Delta\varepsilon = h\nu$$

where *h* is an adjustable parameter which was varied to fit the experimental observations.

Thus, the emitted energy *E* could only take on integral values of $h\nu$

$E = nh\nu$ n = 1, 2, 3, . . .

Lord Rayleigh

Sir James Jeans

Max Planck

From statistical thermodynamics, the above condition gives

$$\rho(\nu,T)d\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/kT}-1}d\nu$$

This result matched the experimental observations for blackbody emission perfectly and yielded a value for h of $6.626 \cdot 10^{-34}$ J·sec. The constant h is nowadays called Planck's constant. This formula for blackbody radiation is frequently used in astronomy to estimate the temperature of stars.

Note what happens in certain limits. If we were to make the value of h arbitrarily small or make T very large, we could expand the exponential in a power series.

$$e^{h\nu/kT} \approx 1 + h\nu/kT + \ldots \qquad \text{for } h\nu/kT = 1$$

If we neglect the higher order terms (which will be small) and put this back into the expression we get:

$$\rho(\nu,T)d\nu \approx \frac{8\pi h\nu^3}{c^3}\frac{1}{1+h\nu/kT-1}d\nu$$

$$\approx \frac{8\pi kT\nu^2}{c^3}d\nu$$

which is identical to the classical result, *i.e*. the Rayleigh-Jeans formula.

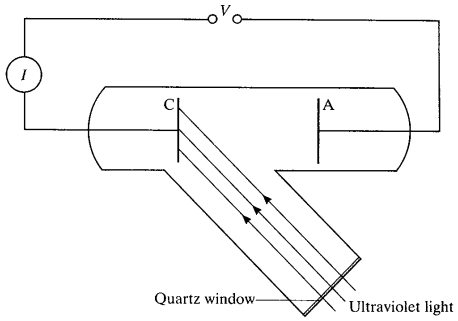
This is a general principle that we will see over and over again. In the limit of $h \to 0$ quantum mechanics goes over to classical mechanics. In addition, in the limit of high temperature quantum mechanics goes over to classical mechanics. Note also that as the frequency gets small, $h\nu/kT = 1$. This is why the original Rayleigh-Jeans formula worked at low frequency.

So Planck's contribution here was that the energy of oscillations of the electrons in a black body was quantized, not the radiation itself. He was awarded the Nobel Prize for this work in 1918.

1.3.2 The Photoelectric Effect (Einstein - 1905)

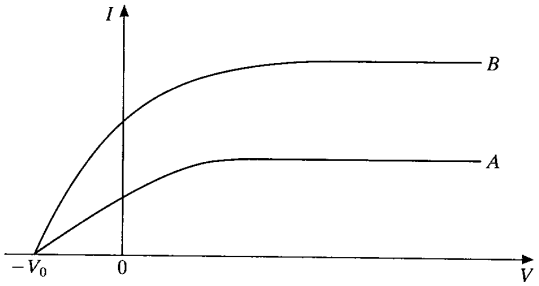
In about 1887, Heinrich Hertz discovered that ultraviolet light causes electrons to be ejected from the surface of a metal. This is called the *photoelectric effect.*

There were several observations involving the photoelectric effect that were in stark contrast to the predictions of classical theories. In 1905 Einstein put forth an explanation of these observations that was central in the development of a *quantum theory*.
First, let me tell you about how the experiments are done. A typical experimental arrangement would have been as follows:



A piece of metal is placed in an evacuated glass tube where light shines on it. Electrons are emitted from the surface of the metal and have some kinetic energy. An electrode is placed opposite the metal surface and a negative potential is applied to retard the electron flow to that electrode.

When the current is measured as a function of the applied voltage between the electrodes one obtains curves as shown in the figure below for low (A) for high (B) light intensity. In both cases, one finds the same voltage $-V_0$ to stop the electrons.



How can we interpret these results? The kinetic energy of the electrons is given by

$$E_{kin} = \tfrac{1}{2} m_e \mathrm{v}^2$$

If we measure the voltage required to stop the electrons, V_0, the energy of those electrons will just be

$$E_{kin} = \tfrac{1}{2} m_e \mathrm{v}^2 = eV_0$$ where e is the electron charge (Note: Volts = Joules/Coulomb)

Thus by measuring the stopping voltage, V_0, one has essentially measured the kinetic energy of the electrons. If one plots the stopping voltage as a function of the frequency of light one obtains something like the following (note that this curve was not known when Einstein proposed his explanation, it rather was predicted by him).

There are several significant observations to be made here:

1. There is a threshold frequency, ν_t, below which no electrons are emitted that is metal dependent.
2. The slope is found to be identical for all metals.
3. The kinetic energy of the electrons is independent of the intensity of light, depending only on the frequency.
4. Increasing the intensity increases the *number* of electrons emitted, *not their energy*.
5. Electrons are released immediately

These observations were contrary to the classical physical description of light. The wave nature of light had been very well established by interference phenomena. The energy of a classical wave of with amplitude E_0, is proportional to the intensity, $I \propto E_0^2$. Increasing the intensity should therefore increase the energy of the incident light and hence increase the kinetic energy of the electrons. In addition, a classical wave of a particular frequency can have any energy, *i.e.* by simply increasing the wave amplitude or intensity.

If the classical description of light were correct, the photoelectric effect should work for any frequency, if intense enough. However this was not found. Red light, no matter how intense, would not cause electrons to be emitted from the metal, whereas blue light that was above a threshold, φ, would eject electrons no matter how weak.

In his work on blackbody radiation, Planck had restricted his concept of the quantization of energy to the emission and absorption process, and he presumed that once emitted, the light behaved like a wave. Einstein proposed the energy itself consisted of concentrated bundles, or photons, where the energy of a single photon is given by

$$E = h\nu \qquad \text{where } \nu \text{ is the frequency of the light and } h \text{ is Planck's constant.}$$

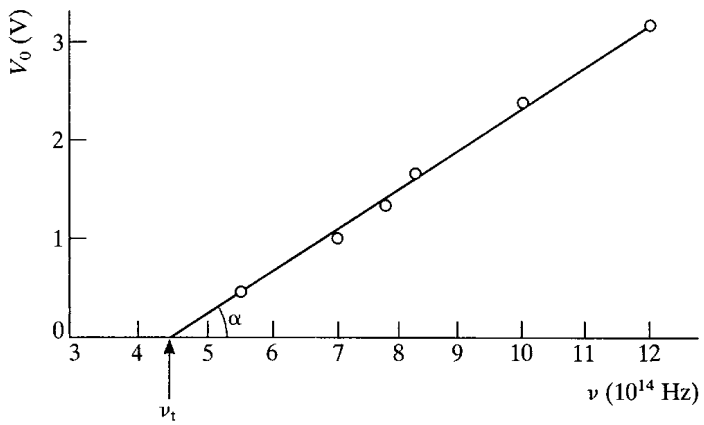
Applying the photon concept to the photoelectric effect, Einstein proposed that the entire energy of a photon is transferred to a single electron in a metal. When emitted, the electron's kinetic energy will be:

$$E_{kin} = h\nu - \varphi \qquad \text{where } \varphi \text{ is the energy required to remove the electron.}$$

Since the stopping voltage is proportional to the kinetic energy, one can write

$$E_{kin} = eV_0 = h\nu - \varphi$$

By plotting V_o versus ν one gets a slope of h/e. By comparing this to the slope of the experimental data, Einstein obtained a value for h of $6.63 \cdot 10^{-34}$ J·sec, in agreement with Planck's results. It was quite amazing at the time that a completely different experimental approach produced a value of this constant so nearly the same.



Heinrich Hertz

Albert Einstein

This model also explains the fact that there is a threshold frequency ν_t, since the process will not occur unless ν satisfies the equation:

$$h\nu_t \geq \varphi$$

Since a single photon ejects a single electron, if the energy of the photon does not exceed the work function of the metal, no electrons will be emitted. However, as ν increases the energy and not the number of electrons will increase. In addition, according to this model, more intense light means more photons and hence more electrons. Therefore, Einstein's model was in complete agreement with the observations.

Einstein's major contribution here was that not only did matter exhibit quantized energy states, but also light was quantized. He was awarded the Nobel Prize in 1921 for this contribution.

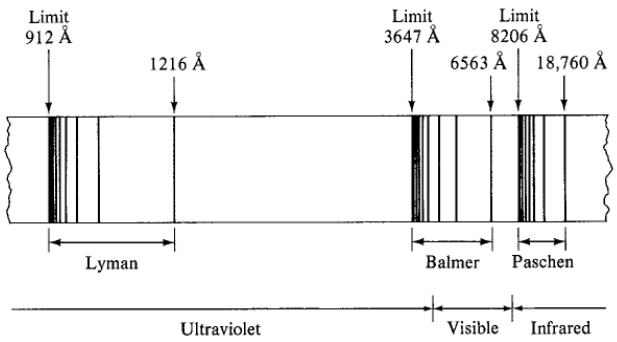
It is interesting to note that Newton had proposed a "corpuscular" theory of light 200 years earlier but this notion had since been dismissed due to the success of wave theories for describing phenomena such as diffraction and interference.

It is now well accepted that light has both wave and particle behavior, and there is a principle formulated by Niels Bohr that we will discuss later which states that you cannot simultaneously observe the wave and particle properties in one experiment.

1.3.3 The Line Spectra of Atoms

It was well known for some time that when gas phase atoms are subjected to high temperatures or an electrical discharge they emit light that is not continuous, but consists of discrete frequencies or lines that are characteristic of a particular element. This is in stark disagreement with classical physics.

Hydrogen, which is the simplest atom, shows the simplest emission spectrum, looking something like this:



An amateur Swiss scientist, Johann Balmer, showed in 1885, that the spacing between some of the lines in the hydrogen spectrum could be expressed in a simple algebraic form. If each line in the series, now called the Balmer series, is assigned an integer *n*, the frequency of lines is given quantitatively by the formula:

$$\nu \propto \left(\frac{1}{2^2} - \frac{1}{n^2} \right) \qquad n = 3,4,5,\ldots$$

A Swedish spectroscopist, Johannes Rydberg, was later able to account for all the lines in the hydrogen spectrum by using the empirical formula:

$$\tilde{\nu} = \frac{1}{\lambda} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \qquad n_1 = 1,2,3,\ldots \quad \text{and} \quad n_2 > n_1$$

where R_H is an empirically determined constant known as the Rydberg constant.

$$R_H = 109\,680 \text{ cm}^{-1}$$

The fact that integers are involved here is very different from what you would expect from a classical picture and strongly suggest that the energy of the hydrogen atom is quantized.

There was no explanation for this observation until the work of Niels Bohr who was working for Rutherford in 1911 when Rutherford proposed the nuclear model for the atom: that is that atoms exist as a heavy, positively charged nucleus with electrons surrounding it. Bohr abandoned classical physics by making the following assumptions about the hydrogen atom.

Bohr's (non-classical) assumptions:

1. The electron in an atom has only certain definite stationary states of motion; each of these states has a definite, fixed energy. In other words, *the energy is quantized*.

2. In any of these states the electron moves in a circular orbit around the nucleus.

3. When an atom is in one of these states it does not radiate; but when changing from a high-energy state to a state of lower energy the atom emits a photon whose energy $h\nu$, is equal to the difference in energy of the two states, *i.e.*,. $\Delta E = h\nu$

4. The states of allowed electronic motion are those in which the angular momentum of the electron is an integral multiple of $h/2\pi$. In other words, *the angular momentum is quantized.*

$$l = n\frac{h}{2\pi} = n\hbar \qquad \text{where} \qquad \hbar = \frac{h}{2\pi}$$

These ideas were very non-classical. Quantized energy levels were still not widely accepted at that point. Particularly troubling was the fact that the electrons moved in circular orbits yet did not radiate. In electromagnetic theory, a charge that is accelerated will continuously emit energy and hence gradually spiral into the nucleus. Although this picture of circular orbits *was later shown to be wrong*, it gave the correct result for the hydrogen atom.

Finally, the quantization of orbital angular momentum was done in an *ad hoc* manner. There was no physical justification for it.

Having made the non-classical assumptions of circular orbits of fixed energy and angular momentum, Bohr went on to apply the laws of classical mechanics.

Let us look at his derivation. As the electron rotates about the nucleus, the attractive Coulomb force between the two charged particles directed in towards the center (*i.e.*, toward the nucleus, assuming it is infinitely heavy compared to the electron) supplies the force needed to keep the electron in a stable, fixed circular orbit. (Think of a ball rotating on a string. The inward force is supplied by the tension on the string.)

Classical Equations:

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{m\mathrm{v}^2}{r}$$

Solving for *r*:

$$r = \frac{e^2}{4\pi\varepsilon_0 m\mathrm{v}^2}$$

The angular momentum of a particle is given by:

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

Assuming a circular orbit for the particle, one can write:

$$l = |\mathbf{r} \times \mathbf{p}| = |\mathbf{r}||\mathbf{p}|\sin\theta = |\mathbf{r}||\mathbf{p}|\sin 90^\circ = rp = m\mathrm{v}r$$

The quantization of angular momentum thus gives:

$$l = m\mathrm{v}r = n\hbar \qquad n = 1,2,3\ldots$$

We can solve this equation for v and substitute it in to our equation for *r*.

$$\mathrm{v} = \frac{n\hbar}{mr}$$

Thus
$$r = \frac{e^2}{4\pi\varepsilon_0 m\left(\frac{n\hbar}{mr}\right)^2}$$

Rearranging:

$$r = \frac{4\pi\varepsilon_0\hbar^2}{me^2}n^2 \qquad n = 1,2,3,\ldots$$

$$r = a_0 n^2 \qquad a_0 = 5.29\cdot 10^{-11}\,\mathrm{m} = 0.529\text{Å}$$

where a_0 is nowadays known as the Bohr radius. You can see that the circular orbits have radii that are characteristic of the quantum number *n*.

To find the energy of the electron in the hydrogen atom, we need to write down the sum of the kinetic energy of the electron and the potential energy of the electron-proton interaction

$$E = \tfrac{1}{2}m\mathrm{v}^2 - \frac{e^2}{4\pi\varepsilon_0 r}$$

We can substitute for mv^2 in the kinetic energy term using our first equation:

$$E = \frac{1}{2}\frac{e^2}{4\pi\varepsilon_0 r} - \frac{e^2}{4\pi\varepsilon_0 r}$$

$$= -\frac{e^2}{8\pi\varepsilon_0 r}$$

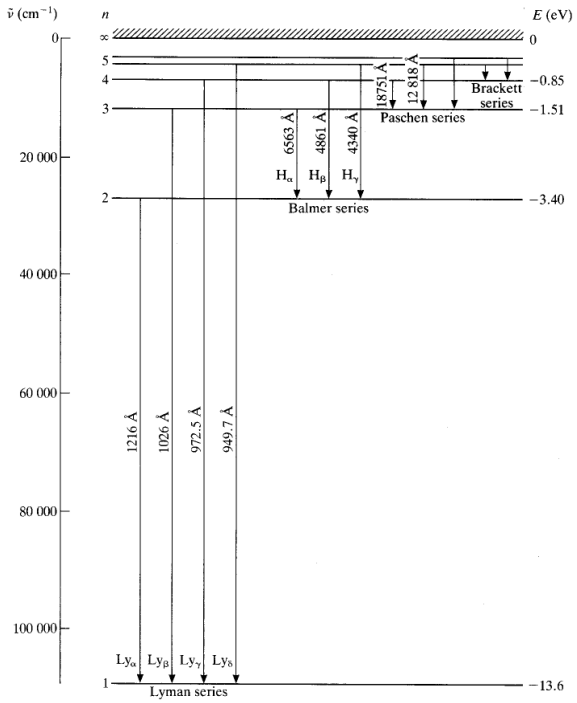
Finally, substituting our formula for r we get

$$E_n = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{n^2} \qquad n = 1,2,3,\ldots$$

Now these are the energies of the stationary states of the electron. Light is emitted when the atoms go from a higher state down to a lower state.

Thus $$\Delta E = \frac{me^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) = h\nu$$

Setting $\Delta E = h\nu$ gives us what is called the ***Bohr frequency condition***, *i.e.* as the electron falls from one level to another, the photon given off has energy $h\nu$. You can see that this formula agrees with the empirically determined Rydberg formula. We can now see where the hydrogen atom spectrum arises.



The Lyman series occurs when electrons relax to n=1 from higher levels. The Balmer series occurs when excited electrons relax back into the n=2 state, and so on.

One now thus finds a theoretical value for the Rydberg constant given by:

$$R_H = \frac{me^4}{8\varepsilon_0^2 h^2}$$

Johann Balmer

Johannes Rydberg

Niels Bohr

This permitted calculation of the empirical Rydberg constant from other fundamental constants to within 0.5% accuracy. Including the reduced mass of the electron improves the agreement even further. This model works for any one-electron atom, *e.g.*, He^{+}, Li^{2+}.

Although this model of the atom is oversimplified and was replaced later, this concept of quantization of electronic energy levels was important to the development of quantum theory.

To summarize the major conclusions drawn from these and other experiments in the early 20th century:

- Planck's description of blackbody radiation and Bohr's analysis of the line spectra of atoms suggest that molecules and atoms emit energies in discrete or quantized amounts. Hence, their energies are quantized.

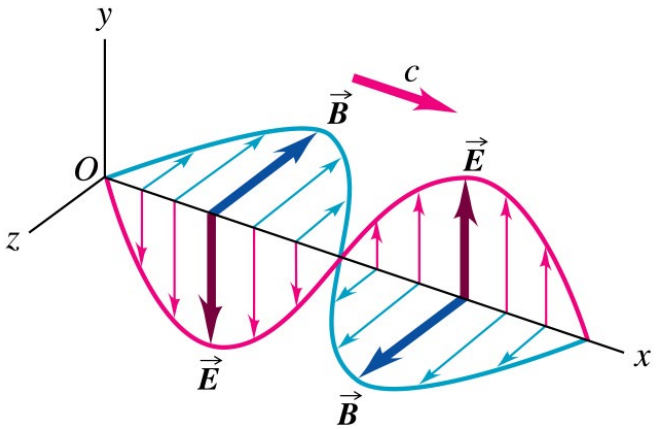
- The photoelectric effect suggests that such quantization is not only a property of matter, but also an intrinsic property of light itself.

The idea that light consisted of "particles" had actually been suggested by Newton, but the overwhelming amount of data displaying the wave-like properties of light (particularly the work of Maxwell) had caused this notion to be rejected. However, new experiments revealed that light exhibits both wavelike and particle-like properties.

The wave-particle duality of light (and matter) is of central importance to the development of quantum theory. It is therefore worth our considering it further.

1.3.4 The Wavelike Nature of Light

The wave nature of light is given by Maxwell's equations and is described by mutually perpendicular oscillating electric and magnetic fields.



For plane polarized light:

$$E(x,t) = E_{y0}\cos(kx - \omega t) = E_{y0}\cos\left(2\pi\left[\frac{x}{\lambda} - \nu t\right]\right)$$

$$B(x,t) = B_{z0}\cos(kx - \omega t) = B_{z0}\cos\left(2\pi\left[\frac{x}{\lambda} - \nu t\right]\right)$$

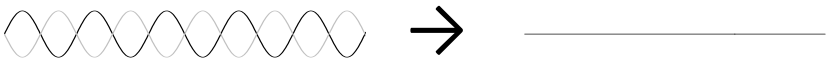
where $k = \frac{2\pi}{\lambda}$ and $\omega = 2\pi\nu$

The wave is thus characterized by a wavelength, λ and a frequency, ν such that

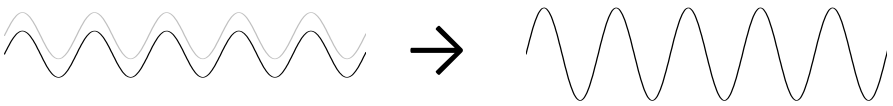
$$\lambda\nu = c \quad \text{or} \quad \nu = \frac{c}{\lambda}$$

One property of waves, which will be of great importance to us in this course, is that of interference or superposition. Interference effects were demonstrated in 1801 by Thomas Young.

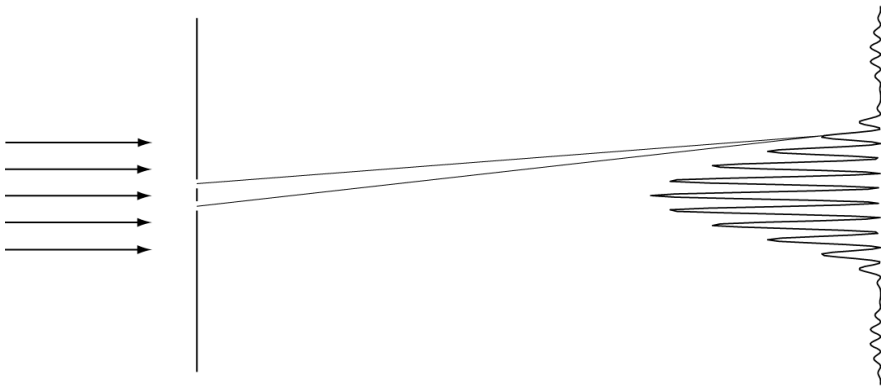
The result of superimposing two waves can be determined by simply adding the wave amplitudes. If two waves are 180° out of phase, they interfere destructively.



Conversely, if they are in phase they constructively interfere.



You may also recall that for a double slit experiment one gets a diffraction pattern that is modulated by interference fringes.



These patterns occur because waves emanating from each slit differing in path length by integral numbers of wavelengths interfere.

James Clerk Maxwell

Thomas Young

1.3.5 The Particle-like Nature of Light

The particle nature of light is characterized by:

- Discrete energies $E = h\nu$ (we have discussed this already)
- Its momentum

We know that the rest mass of a photon is zero. However, Einstein's work on Relativity tells us that a photon has a relativistic mass given by:

$$E = mc^2$$

If we combine this with Planck's equation:

$$E = h\nu = mc^2$$

Rearranging we get:

$$mc = \frac{h\nu}{c} = \frac{h}{\lambda} \qquad \text{since} \quad \frac{\nu}{c} = \frac{1}{\lambda}$$

The momentum of the photon is than given by:

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

That photons actually do have momentum was confirmed by Compton in 1924, when he observed scattering of photons by electrons. He described the process theoretically assuming a photon momentum of h/λ, and this prediction was confirmed by experiment. He was awarded the 1927 Nobel Prize in physics for this work. Note that Einstein never discussed the momentum of a photon, even though he proposed the photon concept and developed the theory of Relativity.

Light exhibits both wave- and particle-like properties.

At this point it is useful to discuss the units used by atomic and molecular scientists to express the characteristics of light.

We already saw that the energy of a photon is given by:

$$E = h\nu$$

If one gives the frequency of the light one can directly convert it to energy using this expression. However, the frequency of light at which atoms absorb light is in the order of 10^{15} Hz, not a convenient order of magnitude to use.

We saw before that that the frequency of the light is related to its wavelength according to:

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

Thus we can express the energy of a photon as:

$$E = hc\frac{1}{\lambda} = hc\tilde{\nu}$$

where $\tilde{\nu} = \frac{1}{\lambda}$

is known as the wavenumber (not to be confused with the wavevector) and has the units of cm^{-1}. If the wavenumber is given the energy can be simply calculated using the expression above. The use of wavenumbers has the advantage that the relevant part of the electromagnetic spectrum is covered by 1-100'000 cm^{-1}.

Once the wavelength of the light is known the corresponding wavenumber can be readily calculated. Let us calculate the wavenumber for light with a wavelength of 500 nm (blue-green light).

$$\lambda = 500 \text{ nm} = 500 \cdot 10^{-9} \text{ m} = 5 \cdot 10^{-7} \text{ m} = 5 \cdot 10^{-5} \text{ cm}$$

Consequently, one finds

$$\tilde{\nu} = \frac{1}{5 \cdot 10^{-5}} = 20'000 \text{ cm}^{-1}$$

1.3.6 The wave-like Nature of Matter

Experiments in the beginning of last century not only indicated that light exhibits wave-particle duality, *but that matter does as well.* We are quite familiar with the particle-like properties of matter, since our intuition is calibrated by classical mechanics. However, we need to discuss the wave properties of matter.

In 1923, DeBroglie suggested that not only does light show properties of particles, but particles of matter exhibit wavelike properties as well. DeBroglie was awarded the Nobel Prize in 1929 for his Ph.D. thesis.

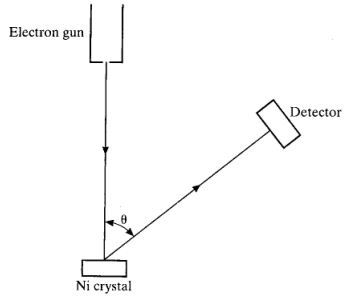
He suggested that an electron of mass *m* and speed v would have a wavelength:

$$\lambda = \frac{h}{m\text{v}} = \frac{h}{p}$$

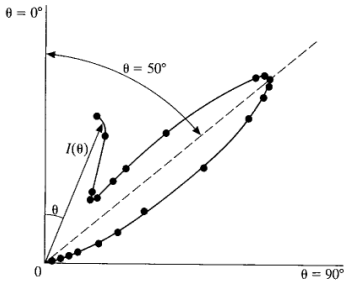
So analogous to radiation, matter also obeys the laws:

$$E = h\nu \qquad \text{and} \qquad p = \frac{h}{\lambda}$$

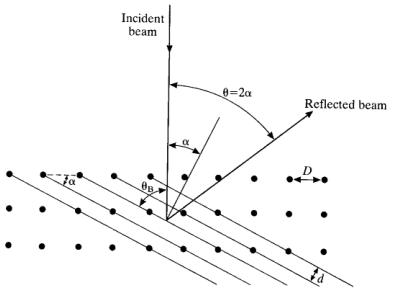
In 1927, Davisson and Germer confirmed this hypothesis by demonstrating that electrons could exhibit Bragg diffraction off a crystal surface by accelerating them to a velocity such that their wavelength is of comparable dimensions to the crystal lattice. A schematic overview of their setup is shown below. Using this setup they observed a maximum in the scattered electron intensity at an angle of 50°. This observation could be explained by constructive interference of the waves by the crystal lattice.



Experimental setup used by Davisson and Germer



Measured angular distribution of the scattered electrons



Scattering of electron waves by a crystal

Davisson was awarded the Nobel Prize in 1937 for this work along with G. P. Thomson. The latter was the son of J. J. Thomson who discovered the electron and characterized it as a particle with definite mass to charge ratio. He was awarded the Nobel prize in 1906 for this. So the father won it for demonstrating the electron as a particle and the son for demonstrating that it is a wave.

To get a feel for the wavelengths of typical objects consider the following:

Particle	Mass (kg)	Velocity (m/sec)	λ (Å)
100 volt electron	$9.1\cdot10^{-31}$	$5.9\cdot10^{6}$	1.2
Dust particle	~ 10^{-15}	$1\cdot10^{-3}$	$6.6\cdot10^{-6}$
Tennis ball	0.057	57 (≈200 km/h)	$1.18\cdot10^{-24}$

You can see why it is impossible to detect quantum effects on a macroscopic scale. The smallness of *h* makes these effects apparent only for small mass objects.

It is important to note that: **Both matter and radiation exhibit wave-particle duality.**

It is important to note that in any given measurement, only one model applies. Niels Bohr summarized this situation by stating: "If a measurement proves the wave character of radiation or matter, then it is impossible to prove the particle character in the same experiment, and conversely."

This is known as the *Bohr Complementarity Principle*. The two views can be linked by a probabilistic interpretation. Einstein made the following argument:

In the particle picture, light intensity is given by:

$$I = N h\nu ,$$

where *N* is the number of photons per unit time crossing a unit area $\perp$ to the direction of propagation. From wave theory, the intensity is proportional to the average value of the square of the electric field over one cycle, $\overline{E}^2$. More specifically,

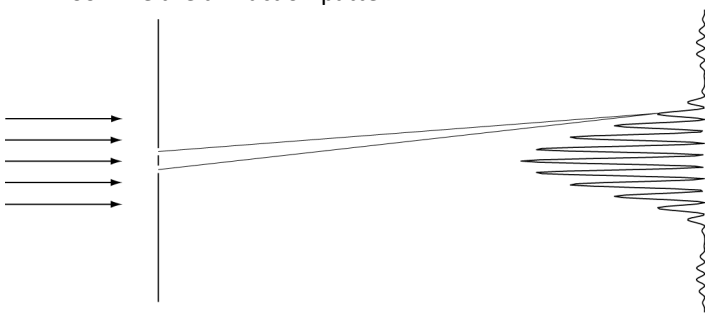
$$I = \varepsilon_0 c\overline{E}^2$$

Equating these we see:

$$N \propto \overline{E}^2$$

Thus, Einstein interpreted $\overline{E}^2$ as the probability measure of photon density.

Let us look back at our diffraction pattern. One detects the intensity at each point along the screen. If we turn down our light intensity such that one photon goes through at a time and use a CCD camera where we could detect individual photons, any one photon will make one spot on the screen (Measuring particle properties here). However, if we repeat this experiment a large number of times, one could build up a histogram of photon impacts. That histogram will look like the diffraction pattern.



In analogy to Einstein's view of the wave-particle duality of radiation, Max Born proposed a similar view of the wave-particle duality of matter. We talked about the wavelength associated with matter, but we can also assign it an amplitude.

The function representing the DeBroglie wave is called a wave function, signified by Ψ. The square of the amplitude of this matter wave is related to a probability.

I would like to make one last point about matter waves. The DeBroglie formula for matter waves gives a nice interpretation of Bohr's quantization condition for electron orbits. Since the electron in the hydrogen atom has a particular wavelength and travels in circular orbits (according to Bohr), then it is reasonable to assume that the wave must "match" in this circular orbit, otherwise there will be interference and cancellation of intensity. One could show that a non-match would lead to total destructive interference of the wave.

For the orbits to match, we will have the following condition:

$$2\pi r = n\lambda \qquad n = 1, 2, 3, \ldots$$

If we substitute DeBroglie's relation:

$$\lambda = \frac{h}{p} = \frac{h}{m\mathrm{v}}$$

one gets

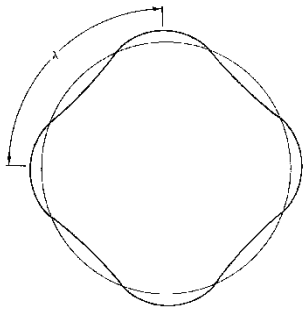
$$2\pi r = \frac{nh}{m\mathrm{v}}$$

or

$$l = m\mathrm{v}r = \frac{nh}{2\pi} = n\hbar$$

This is exactly the assumption Bohr made.

Below is a schematic of the standing wave produced by an electron in a Bohr orbit with $n = 4$. In this case the circumference of the circle is 4λ.



[Arthur Compton]

[Louis-Victor DeBroglie]

[Davisson] and [Germer]

[G.P Thomson]

1.3.7 The Heisenberg Uncertainty Principle

A very important consequence of this wave-particle duality of matter is a principle called the Heisenberg Uncertainty Principle.

It states that an experiment cannot *simultaneously* determine the exact component of momentum p_x say, of a particle and its exact corresponding coordinate position in x. The precision of the measurement is inherently limited by the measurement itself, such that $\Delta p_x \Delta x \approx h$ This principle has nothing to do with the precision of the instrumentation (but with the magnitude of h).

Let's say we want to make the most accurate measurement possible of the position of an electron. For the electron to be "seen", a photon must interact or collide in some way with the electron, or else we have no way of knowing it is there. If we wish to locate the electron within a distance Δx, we need to use light with a wavelength at least that small. However, the photon has a momentum $p = h/\lambda$, and during the collision some of the momentum will be transferred to the electron. The act of locating the electron, changes its momentum in an uncertain way. If we wish to locate the electron more accurately, we need shorter wavelength light, but these photons will have higher momentum, and will thus cause more uncertainty in the momentum of the particle.

To look at this more quantitatively, say we wanted to locate an electron to within 1 Å.

$$\Delta x = 1 \cdot 10^{-10}\,\text{m}$$

$$\Delta p_x = \frac{h}{\Delta x} = \frac{6.626 \cdot 10^{-34}\,\text{Js}}{1 \cdot 10^{-10}\,\text{m}}$$

$$= 6.6 \cdot 10^{-24}\,\text{kgms}^{-1}$$

Since $p = m\text{v}$,

we have $\Delta p = m\Delta\text{v}$ or $\Delta\text{v} = \frac{\Delta p}{m}$

Having $m = 9.11 \cdot 10^{-31}$ kg

yields for the uncertainty in speed

$$\Delta\text{v} = \frac{6.6 \cdot 10^{-24}\,\text{kgms}^{-1}}{9.11 \cdot 10^{-31}\,\text{kg}}$$

$$= 7.2 \cdot 10^{6}\,\text{ms}^{-1} \qquad \textit{This is } 2.4\%\ \textit{the speed of light}!!$$

Note that:

1. if Δx increases, Δv decreases
2. if m increases, Δv decreases

You can see that in the limit of large masses and macroscopic sizes or distances, this phenomenon becomes unimportant.

1.3.8 Summary

We discussed 3 different discoveries made in the early 20th century (there were others, but I picked three of the most important ones). Below these discoveries and the most important new concepts introduced are listed:

(1) Theory of blackbody radiation (Planck, 1900; Nobel Prize 1918)

Key concepts introduced:

- Energy emitted by a blackbody takes on integral values of $h\nu$ hν; $E = Nh\nu$
- Introduced Planck's constant, h

(2) Photoelectric effect (Einstein, 1905; Nobel Prize 1921)

Key concept introduced:

- Light has quantized energy, $E = h\nu$

(3) Line spectra of atoms

Key concepts introduced:

- Quantized energy of atoms
- Energy and angular momentum of electron quantized

Rydberg and Balmer found empirical formula:

$$\tilde{\nu} = R_H\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \qquad n_1 = 1,2,3,\ldots \text{ and } n_2 > n_1$$

Bohr (Nobel Prize 1922):

- Stationary states of electron, angular momentum quantized $I = \frac{nh}{2\pi}$
- $\Delta E = h\nu$ upon changing state

(4) Wave-particle duality of light and matter

Wave-particle duality of light	*Wave-particle duality of matter*
$E = h\nu$ (Einstein)	$\Delta E = h\nu$ (Planck, Rydberg, Bohr)
$p = \frac{h}{\lambda}$ (Compton)	$\lambda = \frac{h}{p}$ (DeBroglie)

(5) Heisenberg Uncertainty Principle

$\Delta p_x \, \Delta x \geq h$ result of wavelike nature of matter. We will see that the principle is much broader than this simple statement.

Quantum chemistry deals with the wavelike nature of matter!!

2 The Postulates of Quantum Mechanics

2.1 Introduction

In light of the series of experiments performed in the early 20th century that we have discussed, it became clear that existing theories of classical physics did not adequately describe these newly discovered phenomena, particularly the quantization of the energies of atoms and molecules, the quantization of light itself, and the wave-particle duality of both light and matter.

Think about the process by which science moves forward. One proposes a theory that describes the existing experimental phenomena. One then tests the theory by performing additional experiments. If the theory describes all know phenomena, one begins to accept it as being true. There may be a small number of observations that a theory doesn't predict, and in some cases it may cause one to reexamine the experiment to make sure it was done properly. However, as the number of experimental observations that the theory cannot account for increases, the theory must be replaced. Sometimes this happens by making slight modifications of the existing theory, and sometimes the old theory is completely overturned and replaced by one with a totally different perspective.

This is what occurred in 1925 when Erwin Schrödinger and Werner Heisenberg, working independently, formulated a general quantum theory. The quantum theory provided a completely different way to view physical systems. It was a revolutionary way of thinking and requires abandoning much of the intuition that one develops from the macroscopic world. In certain limits, the quantum theory reduces to the classical theory, so that not only does it describe the newly observed phenomena, but the existing body of scientific results as well.

The theories independently formulated by Schrödinger and Heisenberg are mathematically different. Schrödinger used a differential equation approach whereas Heisenberg used a matrix formalism. While at first the two theories appeared different in substance, a year later Schrödinger demonstrated that the two formulations are mathematically equivalent.

In this course, I will follow the differential equation approach formulated by Schrödinger, although later in the course I may introduce aspects of the Heisenberg approach. I will not present quantum mechanics in an historical manner, however--that is by following the way that Schrödinger first developed it. We will take advantage of hindsight, particularly in the interpretations of certain aspects of the theory. I will present to you a new way of thinking, and ask you to try to abandon much of the intuition you have developed during your first year of physics.

I will introduce the quantum theory by presenting to you a series of 6 ***postulates***. All my further development of quantum mechanics will be based on those six postulates. It is important to realize what a postulate is (and what it is not). A ***postulate*** is a statement of thesis, not particularly a statement of fact. It is not a law. It can always come under scrutiny and should do so. The test of whether a postulate is correct is if it describes experimental results -- that is, that it works!!

I will present and briefly discuss each of these 6 postulates and give examples of applications to simple quantum mechanical systems where necessary.

We will treat these postulates like laws in the sense we will base our intuition and predictions about the properties of molecules on these postulates. One must always recognize, however that tomorrow someone might observe some property or phenomena that is not in accord with these postulates.

This may seem a bit disturbing to approach science in this way in that we like to have laws that one can derive mathematically. While it might seem a bit uncomfortable, it is no different from classical physics, for example. At the very depth of our laws and rules in science there are always postulates that are accepted simply because they work, *i.e.*, the theory based upon them predicts the results of experiment. Examples are Coulombs law or the gravitational force.

Werner Heisenberg

Erwin Schrödinger

Two primary concepts are involved in the postulates of quantum mechanics:

- The state of a system
- Physical observables

A physical observable is simply some variable that can be measured. Examples are position, momentum, angular momentum, and energy.

2.2 Postulate 1

The state of a quantum mechanical system is completely specified by a function $\Psi(\mathbf{r},t)$ *that depends on the coordinates of the particle and on the time.* $\Psi(\mathbf{r},t)$ *is called the wave function or state function and has the property that* $\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t)dxdydz$ *is the probability that the particle lies in the volume element* $dxdydz$ *at position* $\mathbf{r}$ *at time* t.

- This postulate says that $\Psi(\mathbf{r},t)$ exists and in principle can be determined.
- If $\Psi(\mathbf{r},t)$ is known, it can predict how the system changes with time (*i.e.* how the probability distribution changes).

This situation is in contrast to that of classical mechanics. In a classical mechanical system, given Newton's equations of motion and a set of initial coordinates and momenta, one can determine the trajectories of the particles for all time. This is not possible in Quantum Mechanics due to the Heisenberg Uncertainty Principle. All we can do is see how the probability distribution changes with time.

Because of the interpretation of $\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t)dxdydz$ as a probability, $\Psi(\mathbf{r},t)$ must meet certain conditions:

- $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t)dxdydz=1$$

  The probability integrated over all space equals 1.

  To save time and paper I will abbreviate integrals like this as $\int_{-\infty}^{\infty}\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t)d\tau=1$

- $\Psi(\mathbf{r},t)$ must be finite over the possible range of values of the coordinates
- $\Psi(\mathbf{r},t)$ must be single valued, continuous (*i.e.* a well behaved function)
- $\nabla\Psi(\mathbf{r},t)$ must be single valued, continuous

Remember we discussed Einstein's explanation of the relationship between the wave and particle behavior of light. In the wave picture, the intensity of an electromagnetic wave is proportional to the amplitude of the wave squared.

$$I = \varepsilon_0 c\bar{E}^2 \qquad \text{where} \quad E = E_0 \cos(\omega t) \quad \text{for the electric field}$$

In the particle picture however,

$I = N h \nu$ where N is the number of photons per unit area per unit time.

Thus, $N \propto \bar{E}^2$

That is, the probability of detecting a photon at a certain point is proportional to the square of the average wave amplitude.

In analogy to this, the square of a wavefunction $\Psi(\mathbf{r},t)$

$$\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t) \qquad \text{or} \qquad \left|\Psi(\mathbf{r},t)\right|^2$$

can be interpreted as the probability density of finding a particle at a particular point in space. This is known as the Kopenhagen interpretation and was proposed by Niels Bohr. One should realize that also other interpretations exist, but none of them is as widely accepted as the Kopenhagen interpretation. Although the interpretations might be different, the physics described by the theory remains the same.

2.3 Digression on operators

Before we go on to postulate 2, I need to make a brief digression and discuss operators.

Definition of an Operator:
A rule which transforms one function, f(x) for example, into some other function, g(x).
Basically it is any mathematical operation.

Examples of operators:

$$\int \quad , \quad \frac{d}{dx} \quad , \quad \sqrt{} \quad , \quad x^2 \quad , \quad \log \quad , \quad 3$$

These are (in order), integrating, taking the derivative with respect to x, taking the square root, multiplying by x^2, taking the log, multiplying by the number 3.

I will sometimes designate an operator with a hat or carrot, $\hat{A}$, for example, but I will often omit it. In such cases, the fact that it is an operator should be evident from the context.

It is important to note that in Quantum Mechanics, we will deal only with *linear operators*.

Definition of a Linear Operator:
An operator is said to be linear if

$$\hat{A}\left[c_1 f_1(x) + c_2 f_2(x)\right] = c_1 \hat{A} f_1(x) + c_2 \hat{A} f_2(x)$$

where c_1 and c_2 are constants.

Clearly, $\frac{d}{dx}$ (differentiation) or $\int$ (integration) are linear since

$$\frac{d}{dx}\left[c_1 f_1(x) + c_2 f_2(x)\right] = c_1 \frac{df_1(x)}{dx} + c_2 \frac{df_2(x)}{dx}$$

$$\int \left[c_1 f_1(x) + c_2 f_2(x)\right] dx = c_1 \int f_1(x) dx + c_2 \int f_2(x) dx$$

Whereas the $\sqrt{\ }$ is *not* a linear operator

$$\sqrt{c_1 f_1(x) + c_2 f_2(x)} \neq c_1 \sqrt{f_1(x)} + c_2 \sqrt{f_2(x) dx}$$

When an operator operates on a function and returns the function multiplied by a constant, the function is said to be an **eigenfunction** of that operator. The constant is called an **eigenvalue** of that operator.

That is: $\hat{A} f(x) = a f(x)$

where f is an eigenfunction of $\hat{A}$ and a is the corresponding eigenvalue.

For example:

Let $\hat{A} = \frac{d^2}{dx^2}$ and $f(x) = e^{\alpha x}$

$$\hat{A} f(x) = \frac{d^2}{dx^2} e^{\alpha x} = \alpha^2 e^{\alpha x}$$

Here we would say that $e^{\alpha x}$ is an **eigenfunction** of $\frac{d^2}{dx^2}$ with an **eigenvalue** of α^2.

The equation is called an **eigenvalue problem** or **eigenvalue equation**. For a given operator, solving an eigenvalue equation involves finding both the eigenfunctions and the eigenvalues for a given operator. As we will see shortly, solving the eigenvalue equation plays a central role in quantum mechanics.

We also need to define what we mean by the square of an operator. This simply means the operator applied twice.

If we define an operator (we will see soon that this is the operator describing momentum in the *x*-direction):

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

then the square of the operator would be given by:

$$\hat{p}_x^{\,2} = \hat{p}_x \hat{p}_x = \left(-i\hbar \frac{d}{dx}\right)\left(-i\hbar \frac{d}{dx}\right) = -\hbar^2 \frac{d^2}{dx^2}$$

It is obvious that one cannot simply take the square root to get back the original.

2.4 Postulate 2

To every physical observable in classical mechanics, there corresponds a linear, Hermitian operator in quantum mechanics.

Hermitian operators are linear operators and have the property that their eigenvalues are real. This is necessary if they are to represent physical observables.

For a Hermitian operator $\hat{A}$ the following relation holds for all permissible wavefunctions $\Psi(\mathbf{r},t)$:

$$\int_{-\infty}^{\infty} \Psi^*(\mathbf{r},t)(\hat{A}\Psi(\mathbf{r},t))d\tau = \int_{-\infty}^{\infty} (\hat{A}\Psi(\mathbf{r},t))^*\Psi(\mathbf{r},t)d\tau$$

Prescription for finding an operator for a particular observable:
Write the classical expression for the observable expressed in terms of the Cartesian coordinates and their corresponding momenta.

$x \rightarrow \hat{x} = x$	or in general	$q \rightarrow \hat{q} = q$
$p_x \rightarrow \hat{p}_x = -i\hbar\frac{d}{dx}$	or in general	$p_q \rightarrow \hat{p}_q = -i\hbar\frac{d}{dq}$

where $q = x, y,$ or z.

This prescription **must** involve the **Cartesian coordinates**. After this substitution, one can convert to other coordinate systems.

You can see from the table on the next page that just knowing the expressions for position and momenta operators will enable you to find the others.

For example, one can write the expression for kinetic energy as:

$$K = \tfrac{1}{2}mv^2$$

Since we know that the momentum $\mathbf{p} = m\mathbf{v}$, one can write:

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

The same relationships that hold between the observables will hold between the operators.

One can therefore get the kinetic energy operator in the x-direction by applying the momentum operator twice and dividing by $2m$:

$$\hat{K}_x = \frac{\hat{p}_x\hat{p}_x}{2m} = \frac{1}{2m}\left(-i\hbar\frac{d}{dx}\right)\left(-i\hbar\frac{d}{dx}\right) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

One can do the same with angular momentum. Recall that the angular momentum vector **l** is given by:

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

Observable		**Operator**	
Name	**Symbol**	**Symbol**	**Operation**
Position	x	$\hat{x}$	multiply by x
	$\mathbf{r}$	$\hat{\mathbf{r}}$	multiply by $\mathbf{r}$
Momentum	p_x	$\hat{p}_x$	$-i\hbar\frac{\partial}{\partial x}$
	$\mathbf{p}$	$\hat{\mathbf{p}}$	$-i\hbar\left(\mathbf{i}\frac{\partial}{\partial x}+\mathbf{j}\frac{\partial}{\partial y}+\mathbf{k}\frac{\partial}{\partial z}\right)$
Kinetic Energy	K_x	$\hat{K}_x$	$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$
	K	$\hat{K}$	$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)=-\frac{\hbar^2}{2m}\nabla^2$
Potential Energy	$U(x)$	$U(\hat{x})$	multiply by $U(x)$
	$U(x,y,z)$	$U(\hat{x}, \hat{y}, \hat{z})$	multiply by $U(x,y,z)$
Total Energy	E	$\hat{H}$	$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+U(x,y,z)=-\frac{\hbar^2}{2m}\nabla^2+U(x,$
Angular Momentum	$L_x = y\,p_z - z\,p_y$	$\hat{L}_x$	$-i\hbar\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)$
	$L_y = z\,p_x - x\,p_z$	$\hat{L}_y$	$-i\hbar\left(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z}\right)$
	$L_z = x\,p_y - y\,p_x$	$\hat{L}_z$	$-i\hbar\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)$

The cross product of two vectors, $\mathbf{r}\times\mathbf{p}$ gives you another vector of magnitude $|\mathbf{r}||\mathbf{p}|\sin\theta$ with the direction determined by the right hand rule. To find the individual components of the resultant vector one calculates:

$$\mathbf{r}\times\mathbf{p}=\begin{vmatrix}\hat{\mathrm{i}} & \hat{\mathrm{j}} & \hat{\mathrm{k}}\\ x & y & z\\ p_x & p_y & p_z\end{vmatrix}$$

$$= \left(yp_z-zp_y\right)\hat{\mathrm{i}}-\left(xp_z-zp_x\right)\hat{\mathrm{j}}+\left(xp_y-yp_x\right)\hat{\mathrm{k}}$$

Therefore:

$$L_x = yp_z - zp_y$$

$$L_y = zp_x - xp_z$$

$$L_z = xp_y - yp_x$$

We can take these expressions and substitute the appropriate operators for the position and momentum coordinates.

As you will soon see, a particularly important operator in quantum chemistry is the total energy operator, otherwise known as the Hamiltonian and given the symbol $\hat{H}$.

This operator is found by writing down the classical expression for the sum of the kinetic and potential energy of a system.

$$H = E_{kin} + U$$

In three dimensions this reads in terms of position and momentum coordinates:

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

where $U(x,y,z)$ is the potential energy as a function of the coordinates.
Making the substitution as prescribed above one finds for the Hamiltonian operator:

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(x,y,z) = -\frac{\hbar^2}{2m}\nabla^2 + U(x,y,z)$$

As we will see shortly, solving the eigenvalue problem for this operator will provide us with the energy of the system.

2.5 Postulate 3

In any measurement of the observable associated with the operator $\hat{\mathbf{A}}$*, the only values that will ever be observed are the eigenvalues,* $\boldsymbol{a}$*, which satisfy the eigenvalue equation*

$$\hat{A}\Psi = a\Psi$$

Remember that for any operator we can find a set of **eigenfunctions** and **eigenvalues**.

Recall that an eigenfunction of an operator is defined by the fact that when it is operated on, the result is a constant times the same function. That constant is called the eigenvalue.

This postulate says that a measurement of A can only result in one of its eigenvalues -- **no matter what the wave function is!!.** Thus even if the wavefunction is not an eigenfunction of the operator

Example: Total energy operator $\hat{H}$ (Hamiltonian)

If one makes a measurement of the energy, the only possible values you could obtain are the E_n such that:

$$\hat{H}\Psi_n = E_n\Psi_n$$

where the Ψ_n are the eigenfunctions of $\hat{H}$.

In other words, if you measure the energy, you can only get as a result one of the eigenvalues of the energy operator, the Hamiltonian. Thus even if the wave function is not one of the eigenfunctions of $\hat{H}$, upon measurement, you will only get one of the eigenvalues of $\hat{H}$.

There is an eigenvalue equation for each operator, and any possible measurement of a quantity corresponding to $\hat{A}$ will result in a_n, one of the eigenvalues of $\hat{A}$.

The eigenvalue equation for the total energy operator (*i.e.*, the Hamiltonian) plays a very important role in quantum chemistry. As we will see shortly, the equation

$$\hat{H}\Psi_n = E_n\Psi_n$$

is called the **Time Independent Schrödinger Equation**

This equation provides us with the possible values of the energy of a system and which are of course extremely important in chemistry.

In order to discuss specific examples, I want to make a digression and look at the eigenvalues of the Hamiltonian for a specific quantum mechanical system called the particle in a one-dimensional well or a particle-in-a-box. Although this example might seem farfetched at first it turns out that it can be used to describe the delocalized π electrons in linear conjugated hydrocarbons.

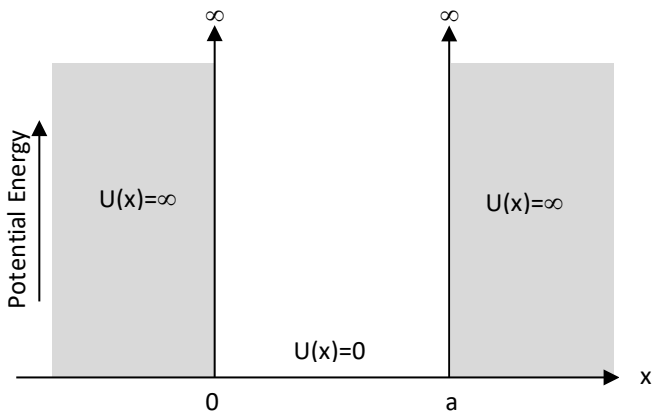
2.5.1 Particle in a 1-dimensional potential well

Consider a particle in a one-dimensional "box" with infinitely high walls (*i.e.,* an infinitely high potential for escaping the box).

Outside the box, the potential is infinite, which means it would take an infinite amount of energy for the particle to exist in that region. Therefore, the wave amplitude or the probability of finding the particle in that region is zero.

So for $x \le 0 \quad \psi(x) = 0$ and for $x \ge a \quad \psi(x) = 0$

These define the boundary conditions for the wave function "inside" the box.



Let us say we want to determine the possible values for the total energy of the system and the eigenfunctions that are associated with those eigenvalues. According to postulate 2, we must write down the operator for the total energy, which we called $\hat{H}$, the Hamiltonian.

Recall that the procedure is to first write down the classical expressions and then substitute operators according to the prescription of postulate 2. We have already done this in the table that I gave you.

The Hamiltonian operator, $\hat{H}$, in one dimension is:

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

The first term is the kinetic energy term and the second term the potential energy.

Since the potential inside the "box" equals zero, the eigenvalue equation for the energy is:

$$\hat{H}\psi(x) = E\psi(x)$$

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

Rearranging this gives:

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0 \qquad 0 \le x \le a$$

and the boundary conditions are $\psi(0) = \psi(a) = 0$. This is because we must require the wave function to be continuous.

This is a second order linear differential equation with constant coefficients. We can solve this by simply looking at it and guessing.

Try $\psi(x) = e^{\alpha x}$ then $\frac{d^2\psi(x)}{dx^2} = \alpha^2 e^{\alpha x}$

Substitute $\psi(x)$ and its second derivative back into the differential equation to get the auxiliary equation.

$$\alpha^2 e^{\alpha x} + \frac{2mE}{\hbar^2}e^{\alpha x} = 0$$

This has to hold for all values of *x*, hence

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

$$\alpha = \pm i\frac{\sqrt{2mE}}{\hbar}$$

The most general solution of the differential equation is a linear combination of the + and - solutions.

$$\psi(x) = c_1 e^{i\frac{\sqrt{2mE}}{\hbar}x} + c_2 e^{-i\frac{\sqrt{2mE}}{\hbar}x}$$

or using Euler's formula: $e^{\pm i\theta} = \cos\theta \pm i\sin\theta$

$$\psi(x) = A\sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) + B\cos\left(\frac{\sqrt{2mE}}{\hbar}x\right)$$

The relation between the coefficients c_1 and c_2 and *A* and *B* is simply

$$A = i(c_1 - c_2) \qquad B = c_1 + c_2$$

Let's now apply the boundary conditions:

At $x = 0$: $\psi(0) = A\sin(0) + B\cos(0) = 0$

$\Rightarrow B = 0$

At $x = a$ we then find:

$$\psi(a) = A\sin\left(\frac{\sqrt{2mE}}{\hbar}a\right) = 0$$

If we say $A = 0$, this gives us a trivial solution (*i.e.* the solution will be zero everywhere, $\psi(x) = 0$).

Therefore $\sin\left(\frac{\sqrt{2mE}}{\hbar}a\right) = 0$ must hold.

This is only possible if:

$$\frac{\sqrt{2mE}}{\hbar}a = n\pi \qquad n = 1,2,3,\ldots$$

We can then solve this equation for the energy, E, and get

$$E = \frac{n^2h^2}{8ma^2} \qquad n = 1,2,3,\ldots$$

These are the eigenvalues of the energy operator (*i.e.*, the Hamiltonian). They represent the allowed values for the total energy of the system.

The eigenfunctions are then given by:

$$\psi(x) = A\sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) = A\ \sin\left(\frac{n\pi}{a}x\right)$$

We can determine the constant A by requiring the eigenfunction to be normalized. Since $\psi^*\psi$ represents a probability, it must equal one if integrated over all space. I will omit the details, but this gives

$$A = \sqrt{\frac{2}{a}}$$

where a is the length of the box.

There are a few important points to note:

- The energy is quantized, *i.e.* it can only have certain values which are determined by the integer n. Note that the quantization arose when we imposed the boundary conditions. This is a general principle of quantum mechanics.

- The energy level spacing increases with decreasing dimensions of the box. The more localized a particle is, the greater the quantization. Another way to look at this is as the dimensions of the box approach the wavelength of the particle, you get quantum effects. As the size of the box or the mass of the particle increases, the energy levels get closer and closer. Quantum effects begin to become less apparent.

Let us look in some detail at the eigenfunctions of the particle-in-a-box Hamiltonian. Let us assume for the moment that the wave function for a particle-in-a-box happens to be one of the eigenfunctions of the Hamiltonian. This need not be the case. The only requirements for an acceptable wave function are given by postulate 1. Nothing says that the wavefunctions for this system must be eigenfunctions of the Hamiltonian. However, one of the eigenfunctions of the Hamiltonian certainly could be a wavefunction for a particle-in-a-box.

Recall the information carried by the wave function for a system. The wave function times its complex conjugate gives the probability of finding the system at a particular point in space.

The eigenfunctions of the Hamiltonian for a particle-in-a-box are given by

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) \qquad n = 1,2,3,\ldots$$

The probability distributions are given by:

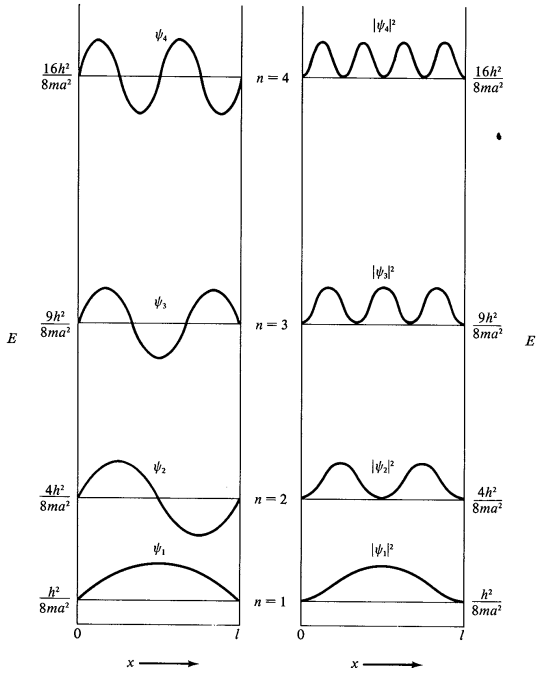
$$\left|\psi_n(x)\right|^2 = \psi_n^*(x)\psi_n(x) = \frac{2}{a}\sin^2\left(\frac{n\pi}{a}x\right)$$

Both of these are plotted in the figure on the next page. As we already saw, there is zero probability of finding the particle outside the box or potential well, because the potential is infinite.

Quantum Mechanically, there are places inside the box where the wave function ψ and $\left|\psi\right|^2$ are zero (*i.e.*, places where the particle has no probability of being found). These are called ***nodes***. The higher the quantum number *n* (and the energy), the more nodes. This will be true for all bound systems. It is always the case in quantum mechanics that the lowest energy wavefunction will have no nodes, the next, one node, etc.... Qualitatively, wavefunctions look similar for different bound potentials.

How can a particle get from one side of the node to the other without crossing the middle of the box? <u>One can't think of it that way</u>! $\left|\psi\right|^2$ simply gives us a probability. One cannot think of trajectories of macroscopic particles in quantum mechanics.

Classically one would expect a particle to bounce back and forth in the box and show equal probability of being found anywhere.



However, note that as *n* gets large, ψ → 

The number of nodes increases, but the probability spreads out evenly. Thus, at large quantum numbers, the quantum mechanical result approaches the classical result. This is an example of the Bohr Correspondence Principle.

Remember that postulate 3 says the only result of measuring A is one of the eigenvalues of $\hat{A}$ ***no matter what the wave function is.***

Postulate 1 does not say that the wave function for a system, Ψ, must be an eigenfunction of the Hamiltonian. For example, in the case of a particle in a 1- dimensional potential well, postulate 1 does not say that the wave function must be one of the eigenfunctions of the Hamiltonian that we found above. Nor does it say that Ψ has to be an eigenfunction of any operator. It simply requires that it be well behaved.

No matter what the wave function is, the only result of measuring A is one of the eigenvalues of $\hat{A}$. If, however, Ψ *happens to be* an eigenfunction of the operator you measure, then, as we will see when discussing postulate 4, you know exactly the result you will get when you make the measurement, the eigenvalue that corresponds to that eigenfunction.

If Ψ *is not* an eigenfunction of the operator you measure, you could get any of the eigenvalues corresponding to operator A.

To help you understand this concept further, I need to introduce several important properties of eigenfunctions.

2.5.2 Some important properties of eigenfunctions

DEFINITION - Complete set of eigenfunctions

A set of eigenfunctions, φ_n, is considered complete if, for any well behaved function, f, we can write:

$$f = \sum_n c_n \varphi_n$$

DEFINITION - Orthonormality of eigenfunctions

A set of eigenfunctions is considered orthonormal if

$$\int_{-\infty}^{\infty} \varphi_i^* \varphi_j \, d\tau = \delta_{ij}$$

here δ_{ij} is the Kronecker delta with the properties:

$$\delta_{ij} = 0 \qquad \text{if } i \neq j$$
$$\delta_{ij} = 1 \qquad \text{if } i = j$$

That is, if $i \neq j$ they are orthogonal, and the case in which $i = j$ expresses the normalization condition.

Auxiliary postulate:
The eigenfunctions of a Hermitian operator form a complete, orthonormal set.

One can prove the fact that the eigenfunctions of a Hermitian operator are orthonormal. However, the completeness of the set cannot be proved mathematically. It must be taken as an additional postulate.

2.5.3 Return to postulate 3

Although a wave function does not have to be a single eigenfunction of any operator, it can **always** be written as a linear combination of eigenfunctions of **any operator that corresponds to a physical observable** (*i.e*. a Hermitian operator), because the eigenfunctions of a Hermitian operator form a complete orthonormal set .
Let us choose to write a wave function for the system we want to measure as a linear combination of the eigenfunctions that correspond to the operator (*i.e*., the observable) we want to measure. (Note that we are free to choose any set of functions as long as it is complete.)

$$\Psi(x,t) = \sum_n c_n \varphi_n(x,t)$$

There could be any number of functions in this expansion.

If you then make a measurement on a system described by such a wave function, the result will be an eigenvalue of one of the eigenfunctions in the linear combination. *Which one you do not know (unless there is only one function in the linear expansion).*

Postulate 4 will tell us the average result you will get if you make repeated measurements on identical systems. Before I discuss postulate 4, I need to remind you of a few concepts related to statistics.

2.6 Some statistical concepts

2.6.1 Expectation values

In general the probability P_j of a certain event can be given by:

$$P_j = \lim_{N\to\infty} \frac{N_j}{N}$$

where N = the total number of tries or measurements

and $N_j =$ the number of times event j was observed.

Note that $0 \le P_j \le 1$ since the maximum value N_j can have is N.

Also, because $\sum_{j=1}^{n} N_j = N$

we have the condition that

$$\sum_{j=1}^{n} P_j = 1$$

where n is the number of possible outcomes for the measurement.
Now consider the case in which a particular event has a value associated with it, as in the case of an experimental measurement. We define the **average value** or **expectation value** to be:

$$\overline{x} = \langle x \rangle = \sum_{j=1}^{n} x_j P_j = \sum_{j=1}^{n} x_j P(x_j)$$

where $P(x_j)$ represents the probability of realizing the number x_j.

Example:
What is the average value observed in rolling a die? *i.e*. What is $\langle x \rangle$ where *x* is the number on the die?

$$\langle x \rangle = \sum_{j=1}^{6} x_j P_j = 1 \cdot P_1 + 2 \cdot P_2 + 3 \cdot P_3 + 4 \cdot P_4 + 5 \cdot P_5 + 6 \cdot P_6$$

If it is an honest die then:

$$P_1 = P_2 = P_3 = P_4 = P_5 = P_6 = \frac{1}{6}$$

Then one would find

$$\langle x \rangle = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = 3.5$$

Note that the average value does not have to equal a value of the die.

2.6.2 Variance

The **variance** is an indication of the distribution or spread in a series of measurements; *i.e*. the distribution about the mean or average value. The variance gives some indication of the precision of a measured result. If a measurement always gives the same number, the variance is zero.

$$\sigma_x^2 = \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle$$

It is a measure of the difference of a measurement *x* from the mean $\langle x \rangle$. Using our definition of average value we can also write this as

$$\sigma_x^2 = \sum_{j=1}^{n} \left(x_j - \langle x \rangle \right)^2 P_j$$

You can see that σ_x^2 is a sum of positive terms and so $\sigma_x^2 \geq 0$

Let's expand the expression for σ_x^2:

$$\sigma_x^2 = \sum_{j=1}^{n} \left(x_j^2 - 2\langle x \rangle x_j + \langle x \rangle^2 \right) P_j$$

$$= \sum_{j=1}^{n} x_j^2 P_j - 2\sum_{j=1}^{n} \langle x \rangle x_j P_j + \sum_{j=1}^{n} \langle x \rangle^2 P_j$$

$$= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2$$

$$= \langle x^2 \rangle - \langle x \rangle^2$$

So, $\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2 \geq 0$

And thus $\sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$

The variance is always greater than or equal to zero and tells us how precise a measurement is or how precisely it can be made. If the variance equals zero we will always measure the same result.

2.7 Postulate 4

If a system is in a state described by a normalized wave function Ψ, then the average value or expectation value of the observable corresponding to the operator $\hat{A}$ is given by:

$$\langle a \rangle = \int_{-\infty}^{\infty} \Psi^* \hat{A} \Psi \, d\tau$$

I need to define more clearly what we mean by an average or expectation value in quantum mechanics.

It ***does not*** mean making repeated measurements on the same system. This is because in making a measurement in quantum mechanics you disturb the system and influence the subsequent result. We will discuss this in more detail later.

It ***does*** mean making the same measurement on a series of identical systems.

Let's have a look at some examples.

Example 1:
Calculate the expectation value for the position of a particle in a one-dimensional box of length *a* that is described by an eigenfunction of the Hamiltonion of the particle-in-a-box problem, furthermore calculate σ_x^2.

We saw before that the wavefunctions are given by:

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a} x\right) \qquad n = 1, 2, 3, \ldots \qquad \text{for } 0<x<a$$

the probability of finding the particle between *x* and *x* + *dx*.

$$\psi^*(x)\psi(x)dx$$

The average position is given by:

$$\langle x \rangle = \int_0^a \psi^*(x)\, x\, \psi(x) dx$$

In the present example we have:

$$\langle x \rangle = \int_0^a \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a} x\right) x \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a} x\right) dx$$

$$= \frac{2}{a} \int_0^a x \sin^2\left(\frac{n\pi}{a} x\right) dx$$

Evaluation of this integral gives us:

$$\langle x \rangle = \frac{2}{a} \frac{a^2}{4} = \frac{a}{2}$$

This result makes sense. The box is perfectly symmetrical, and there is nothing to favor one side of the box over the other, thus the average position is in the center.

We also wanted to calculate the variance σ_x^2. To do this we need to calculate $\langle x^2 \rangle$.

$$\langle x^2 \rangle = \frac{2}{a}\int_0^a x^2 \sin\left(\frac{n\pi}{a}x\right)^2 dx$$

$$= \left(\frac{a}{2\pi n}\right)^2 \left(\frac{4\pi^2 n^2}{3} - 2\right)$$

So,

$$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2 = \left(\frac{a}{2\pi n}\right)^2 \left(\frac{4\pi^2 n^2}{3} - 2\right) - \frac{a^2}{4}$$

$$= \left(\frac{a}{2\pi n}\right)^2 \left(\frac{\pi^2 n^2}{3} - 2\right)$$

Thus there is some distribution of possible values, represented by the fact that $\sigma_x^2 \neq 0$.

Example 2:

Suppose a particle in a 1-dimensional box is in a state described by the following wave function

$$\psi(x) = A(a^2 x^2 - x^4)$$

where $$A = \sqrt{\frac{315}{8a^9}}$$

is the normalization constant and *a* represents the length of the box. This is a perfectly well behaved wave function that satisfies the boundary conditions of the particle in the box properly, but clearly is not an eigenfunction of the Hamiltonian. Let us use postulate 4 to calculate the average energy.

For a particle in a 1-dimensional box:

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

So,

$$\langle E \rangle = A^2 \int_0^a (a^2 x^2 - x^4)\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right)(a^2 x^2 - x^4)dx$$

$$= -\frac{A^2\hbar^2}{2m}\int_0^a (a^2 x^2 - x^4)(2a^2 - 12x^2)dx$$

$$= -\frac{A^2\hbar^2}{2m}\int_0^a (2a^4 x^2 - 14a^2 x^4 + 12x^6)dx$$

$$= -\frac{A^2\hbar^2}{2m}\left[\frac{2}{3}a^4 x^3 - \frac{14}{5}a^2 x^5 + \frac{12}{7}x^7\right]_0^a$$

$$= -\frac{A^2\hbar^2}{2m}\left(-\frac{44}{105}a^7\right) = \frac{22A^2\hbar^2 a^7}{105m}$$

Substituting for A:

$$A = \sqrt{\frac{315}{8a^9}}$$

We finally find:

$$\langle E \rangle = \frac{33}{2\pi^2}\frac{h^2}{8ma^2}$$

This is clearly **not** one of the eigenvalues of the Hamiltonian, but remember, this is an average, not a single measurement. Note, this is similar to what we have seen when we calculated the average value for throwing a die.

Recall that $E_n = n^2\frac{h^2}{8ma^2}$

The **average** or expectation value doesn't need to be an eigenvalue of the Hamiltonian. It is the average of a series of measurements on identical systems.

Postulate 3 says that **upon any one measurement**, the value we obtain can only be one of the eigenvalues of the operator corresponding to the quantity we measure. However, because each time we make a measurement we get a different answer, the average value will not be a single eigenvalue.

Although I won't do it here, if we were to calculate the variance in our measurements

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2$$

we would find that it is larger than zero. One gets a spread in the values obtained upon measuring the energy. Clearly the wavefunction is not an eigenfunction of the Hamiltonian.

Example 3:

Now let's consider a case in which the wave function happens to be an eigenfunction of the particle-in-a-box Hamiltonian, $\hat{H}$.

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$

Let us find the average value for repeated measurements of the energy. We could do this one of two ways. The hard way is to evaluate the integral as we did before:

$$\langle E \rangle = \int_0^a \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right)\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)dx$$

The easy way would be to realize that since the wave function ψ is one of the eigenfunctions of the Hamiltonian, we can write:

$$\langle E \rangle = \int_0^a \psi_n^*(x)\hat{H}\psi_n(x)dx = \int_0^a \psi_n^*(x)E_n\psi_n(x)dx = E_n\int_0^a \psi_n^*(x)\psi_n(x)dx$$

$$= E_n$$

Thus

$$\langle E \rangle = E_n = \frac{n^2h^2}{8ma^2}$$

If we now calculate the variance of the energy measurement:

$$\left\langle E^2 \right\rangle = \int_0^a \psi_n^*(x)\hat{H}^2\psi_n(x)dx = E_n^2$$

$$\sigma_E^2 = \left\langle E^2 \right\rangle - \left\langle E \right\rangle^2 = E_n^2 - E_n^2 = 0$$

we find that the variance equals zero. There is no uncertainty in the measurement!!

If the wave function is an eigenfunction of the operator you measure, then upon making a measurement, you will always get a single value: the eigenvalue corresponding to that eigenfunction.

Example 4:

Let us try one more example in which the wave function is written as a linear combination of eigenfunctions of the particle-in-a-box Hamiltonian. However, let us assume there are only two terms in the expansion.
Let us assume the form:

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

where $\psi_n(x)$ are the eigenfunctions of the particle-in-a-box Hamiltonian and c_1 and c_2 are just the coefficients in the expansion.

As $\psi(x)$ has to be normalized, one can derive a relation between the coefficients c_1 and c_2:

$$\int \psi^*(x)\psi(x)dx = \int \left(c_1^*\psi_1^*(x) + c_2^*\psi_2^*(x)\right)\left(c_1\psi_1(x) + c_2\psi_2(x)\right)dx = 1$$

$$= \int \left(c_1^*c_1\psi_1^*(x)\psi_1(x) + c_2^*c_1\psi_2^*(x)\psi_1(x) + c_1^*c_2\psi_1^*(x)\psi_2(x) + c_2^*c_2\psi_2^*(x)\psi_2(x)\right)dx$$

$$= |c_1|^2 \int \psi_1^*(x)\psi_1(x)dx + c_2^*c_1 \int \psi_2^*(x)\psi_1(x)dx + c_1^*c_2 \int \psi_1^*(x)\psi_2(x)dx + |c_2|^2 \int \psi_2^*(x)\psi_2(x)dx$$

$$= |c_1|^2 + |c_2|^2 = 1$$

Here we made explicit use of the fact that the eigenfunctions of the Hamiltonian form a complete orthonormal set.

Let's choose $c_1 = \sqrt{\frac{2}{3}}$ and $c_2 = \sqrt{\frac{1}{3}}$, so that we have

$$|c_1|^2 = \frac{2}{3} \quad \text{and} \quad |c_2|^2 = \frac{1}{3} \qquad \text{giving} \qquad |c_1|^2 + |c_2|^2 = 1$$

The wave function is then

$$\psi(x) = \sqrt{\frac{2}{3}}\sqrt{\frac{2}{a}}\sin\left(\frac{\pi}{a}x\right) + \sqrt{\frac{1}{3}}\sqrt{\frac{2}{a}}\sin\left(\frac{2\pi}{a}x\right)$$

where I have purposely not combined the constants in front of each function.

Let us now calculate the expectation value. Note that to simplify the writing and to make things more suggestive, I will use $\psi_1(x)$ and $\psi_2(x)$ instead of writing the sine functions explicitly.

$$\langle E\rangle = \int_0^a \psi^*(x)\hat{H}\psi(x)dx$$

$$= \int_0^a \left(\sqrt{\frac{2}{3}}\psi_1^*(x) + \sqrt{\frac{1}{3}}\psi_2^*(x)\right)\hat{H}\left(\sqrt{\frac{2}{3}}\psi_1(x) + \sqrt{\frac{1}{3}}\psi_2(x)\right)dx$$

$$= \frac{2}{3}\int_0^a \psi_1^*(x)\hat{H}\psi_1(x)dx + \frac{\sqrt{2}}{3}\int_0^a \psi_1^*(x)\hat{H}\psi_2(x)dx + \frac{\sqrt{2}}{3}\int_0^a \psi_2^*(x)\hat{H}\psi_1(x)dx + \frac{1}{3}\int_0^a \psi_2^*(x)\hat{H}\psi_2(x)dx$$

$$= \frac{2}{3}E_1\int_0^a \psi_1^*(x)\psi_1(x)dx + \frac{\sqrt{2}}{3}E_2\int_0^a \psi_1^*(x)\psi_2(x)dx + \frac{\sqrt{2}}{3}E_1\int_0^a \psi_2^*(x)\psi_1(x)dx + \frac{1}{3}E_2\int_0^a \psi_2^*(x)\psi_2(x)dx$$

$$= \frac{2}{3}E_1 + \frac{1}{3}E_2$$

To calculate the variance, we need $\langle E^2\rangle$. By repeating the same procedure, you can convince yourself that

$$\langle E^2\rangle = \frac{2}{3}E_1^2 + \frac{1}{3}E_2^2$$

Thus the variance in the measurement of the energy is:

$$\sigma_E^2 = \langle E^2\rangle - \langle E\rangle^2 = \frac{2}{3}E_1^2 + \frac{1}{3}E_2^2 - \left(\frac{2}{3}E_1 + \frac{1}{3}E_2\right)^2$$

$$= \frac{2}{3}E_1^2 + \frac{1}{3}E_2^2 - \frac{4}{9}E_1^2 - \frac{1}{9}E_2^2 - \frac{4}{9}E_1E_2$$

$$= \frac{2}{9}E_1^2 + \frac{2}{9}E_2^2 - \frac{4}{9}E_1E_2 = \frac{2}{9}(E_1 - E_2)^2 \neq 0$$

Therefore, the result of your measurement will not be the same every time. You only get a single value for your measurement each time if ψ is an eigenfunction, not a linear combination of eigenfunctions.

There is one other very important point to note, however. Look again at our expression for the average energy $\langle E\rangle$.

$$\langle E\rangle = \frac{2}{3}E_1 + \frac{1}{3}E_2$$

Remember that we had originally chosen the expansion coefficients $c_1 = \sqrt{\frac{2}{3}}$ and $c_2 = \sqrt{\frac{1}{3}}$

The probability of obtaining either result is simply $|c_1|^2 = \frac{2}{3}$ and $|c_2|^2 = \frac{1}{3}$

In general, the probability of obtaining a particular eigenvalue is simply the square of the expansion coefficient for the eigenfunction corresponding to that eigenvalue.

That is: $P_n = |c_n|^2$

Proof:
We start with the expression for an average value, according to postulate 4:

$$\langle a \rangle = \int_{-\infty}^{\infty} \psi^*(x) \hat{A} \psi(x) dx$$

(Note that I have not included the time dependence of the wave function here. I will explain shortly why I can do that).

Let us substitute for ψ a linear combination of eigenfunctions φ_n of the operator $\hat{A}$, i.e.:

$$\hat{A}\varphi_n(x) = a_n \varphi_n(x)$$

We then find

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

$$= \int_{-\infty}^{\infty} \left(\sum_n c_n^* \varphi_n^*(x) \right) \hat{A} \left(\sum_m c_m \varphi_m(x) \right) dx$$

$$= \sum_n \sum_m c_n^* c_m \int_{-\infty}^{\infty} \varphi_n^*(x) \hat{A} \varphi_m(x) dx$$

$$= \sum_n \sum_m c_n^* c_m a_m \int_{-\infty}^{\infty} \varphi_n^*(x) \varphi_m(x) dx$$

$$= \sum_n \sum_m c_n^* c_m a_m \delta_{nm}$$

$$= \sum_n c_n^* c_n a_n = \sum_n |c_n|^2 a_n$$

Remember how we define the average value:

$$\langle x \rangle = \sum_j x_j P_j$$

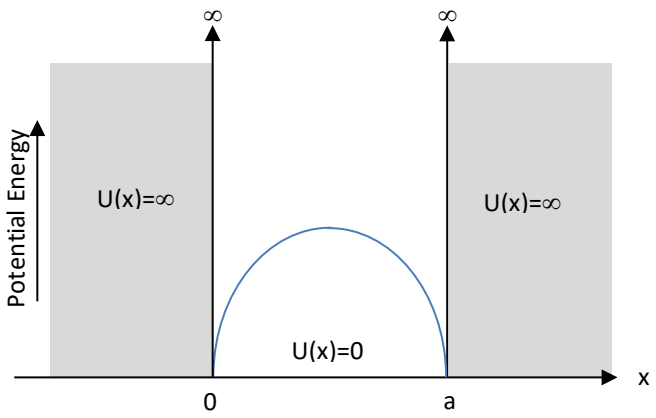
where x_j are the possible values and P_j the probability of obtaining that value.

By analogy, one can see that in our case we see that $P_n = |c_n|^2$

Note that when $c_n = 0$ the probability for obtaining result a_n is zero. One will therefore never measure result a_n. In my previous example, one will only measure E_1 or E_2 and no other value.

In general, the expansion of a wave function may have many terms, and these terms will not be equally weighted; hence, there are many possible results to a measurement. However, those which are weighted more will have a higher probability and will influence the average more heavily.

For the moment, I will demonstrate this physically. Later we will do it in a more mathematical way. Assume that for the particle-in-a-box problem we had a wave function that looked like this.



You can see that this is not an eigenfunction of the particle-in-a-box Hamiltonian (*i.e.,* it is not a sine function) but it has a similar shape. This means that in an expansion of this function in terms of the eigenfunctions of the particle-in-a-box Hamiltonian, the term with the largest expansion coefficient would be ψ_1 (*i.e.,* the term c_1 would be the largest). If you start to think about it you can tell a few other qualitative things as well. Since this wave function is symmetrical, only those functions in the expansion that have the same symmetry will have a non-zero expansion coefficient.

2.7.1 Calculation of coefficients in an eigenfunction expansion

If we are given a functional form for a wavefunction $f(x)$, we can find the coefficients in the following way:

Assume we have a complete orthonormal set of eigenfunctions ψ_n of some Hermitian operator so that we can write

$$f(x) = \sum_{n=1}^{\infty} c_n \psi_n(x)$$

If we multiply both sides of this expression by ψ_m^* and integrate, we get

$$\int_{-\infty}^{\infty} \psi_m^*(x) f(x) dx = \sum_{n=1}^{\infty} c_n \int_{-\infty}^{\infty} \psi_m^*(x) \psi_n(x) dx$$

$$= \sum_{n=1}^{\infty} c_n \delta_{nm} = c_m$$

So,

$$c_m = \int_{-\infty}^{\infty} \psi_m^*(x) f(x) dx$$

That is, to find an expansion coefficient corresponding to a particular eigenfunction, just multiply the function $f(x)$ by the complex conjugate of the eigenfunction of interest and integrate over all space.

The probability of obtaining the m^{th} eigenvalue upon measuring the quantity whose eigenfunctions form the basis for the expansion is then

$$P_m = \left| \int_{-\infty}^{\infty} \psi_m^*(x) f(x) dx \right|^2$$

Example 1:
This example is trivial, but serves to illustrate a point.

Take the wave function for a particle-in-a-box that is a combination of two eigenfunctions

$$\psi(x)=\frac{1}{\sqrt{2}}\psi_1(x)+\frac{1}{\sqrt{2}}\psi_2(x)$$

where ψ_1 and ψ_2 are the first two eigenfunctions of the particle-in-a-box Hamiltonian. We can look at this and see what the expansion coefficients are, but let's go through the procedure we just introduced.

To find c_1 :

$$c_1=\int_{-\infty}^{\infty}\psi_1^*(x)\psi(x)dx$$

$$=\int_{-\infty}^{\infty}\psi_1^*(x)\left(\frac{1}{\sqrt{2}}\psi_1(x)+\frac{1}{\sqrt{2}}\psi_2(x)\right)dx$$

$$=\frac{1}{\sqrt{2}}\int_{-\infty}^{\infty}\psi_1^*(x)\psi_1(x)dx+\frac{1}{\sqrt{2}}\int_{-\infty}^{\infty}\psi_1^*(x)\psi_2(x)dx$$

$$=\frac{1}{\sqrt{2}}$$

The probability of measuring E_1 is then

$$P_1=\left|\frac{1}{\sqrt{2}}\right|^2=\frac{1}{2}$$

Example 2:

Let $\psi(x)=\sqrt{\frac{30}{a^5}}x(a-x) \qquad 0\le x\le a$

We want to expand this as a linear combination of particle-in-a-box wave functions

$$\psi(x)=\sum_{n=1}^{\infty}c_n\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$

and determine the values of the coefficients c_n. Let us calculate the value of the general term c_n.

From the formula that I gave

$$c_n=\int_{-\infty}^{\infty}\psi_n^*(x)\psi(x)dx$$

$$=\sqrt{\frac{2}{a}}\sqrt{\frac{30}{a^5}}\int_0^a\sin\left(\frac{n\pi}{a}x\right)x(a-x)dx$$

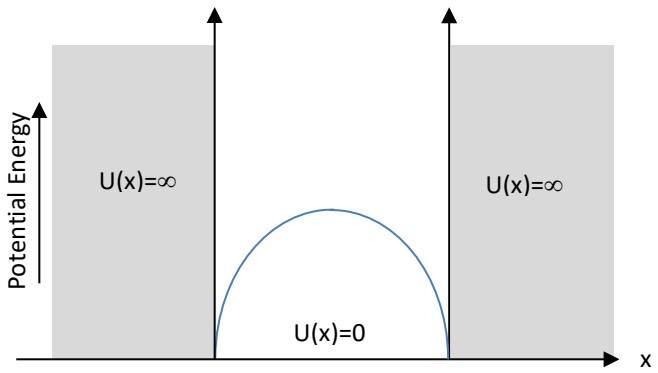
$$=\sqrt{\frac{60}{a^6}}\left[2\left(\frac{a}{n\pi}\right)^3\left(1-\cos n\pi\right)\right]$$

$$= \frac{\sqrt{240}}{\pi^3 n^3}\left[1-(-1)^n\right]$$

Let us look at this result more closely.
Notice that the coefficients corresponding to even values of *n* are 0. That is

$$c_n = \begin{cases} \dfrac{2\sqrt{240}}{\pi^3 n^3} & n = \text{odd} \\ 0 & n = \text{even} \end{cases}$$

Why do the coefficients vanish for even *n*? To understand this we need to look at the function. It looks something like this:



There are two things to notice:

- The symmetry of the function requires the even terms to be zero.
- *n* = 1 should have the largest contribution. You can see that this will be true because $c_n \propto \frac{1}{n^3}$

2.8 Postulate 5

The wave function of a system evolves in time according to the time-dependent Schrödinger equation

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

For most systems $\hat{H}$ does not contain time explicitly. In that case, we can seek solutions of the form

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

That does not mean that all solutions must have this form, however we will see that those that do have a special property: they are the steady-state solutions of the Schrödinger equation, *i.e.*, their probability density does not evolve in time.

If we substitute this into the time-dependent Schrödinger equation:

$$\hat{H}\psi(x)f(t) = i\hbar\frac{\partial}{\partial t}\psi(x)f(t)$$

and realize that since $\hat{H}$ does not depend on time, it will have no effect on $f(t)$ and the derivative with respect to time will have no effect on $\psi(x)$:

$$f(t)\hat{H}\psi(x) = i\hbar\psi(x)\frac{\partial f(t)}{\partial t}$$

If we now divide both sides by $\psi(x)f(t)$ we get

$$\frac{1}{\psi(x)}\hat{H}\psi(x) = \frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t}$$

Note that the left side of this equation is only a function of *x* and the right side is only a function of *t*.

For the two sides to be equal as the variables *x* and *t* are varied, they must each be equal to a constant, which I will call *E*.

$$\frac{1}{\psi(x)}\hat{H}\psi(x) = E = \frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t}$$

First, the left hand side:

$$\frac{1}{\psi(x)}\hat{H}\psi(x) = E \qquad \Rightarrow \qquad \hat{H}\psi(x) = E\psi(x)$$

You can see that this is simply the eigenvalue equation for the Hamiltonian (total energy).

$$\hat{H}\psi(x) = E\psi(x)$$

is called the **Time Independent Schrödinger Equation**.

We will see in a moment that the solution of this equation not only provides the allowed values of the total energy of the system, but the wave functions $\psi(x)$ have special properties.

Now the right hand side:

$$\frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t} = E$$

Rearranging this gives:

$$\frac{\partial f(t)}{\partial t} = -\frac{iE}{\hbar}f(t)$$

The solution of this differential equation is given by:

$$f(t) = Ae^{-\frac{iE}{\hbar}t}$$

Putting the time independent and time dependent parts together, we have

$$\Psi(x,t) = \psi(x)e^{-\frac{iE}{\hbar}t}$$

where $\psi(x)$ is the solution to the Time Independent Schrödinger Equation (note that we have incorporated the constant *A* into $\psi(x)$).

Since there is a whole set of solutions to the Time Independent Schrödinger Equation, denoted by the index *n* (just as we did with the particle-in-a-box wave functions), we write the above equation

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

Recall that from postulate 1, $\left|\Psi(x,t)\right|^2$ represents a probability distribution for finding the particle in an interval *dx* at a particular value of *x*.

Using the wave function above one gets

$$\left|\Psi_n(x,t)\right|^2 = \left(\psi_n^*(x)e^{+\frac{iE_n}{\hbar}t}\right)\left(\psi_n(x)e^{-\frac{iE_n}{\hbar}t}\right)$$

$$= \psi_n^*(x)\psi_n(x)e^{+\frac{iE_n}{\hbar}t}e^{-\frac{iE_n}{\hbar}t}$$

$$= \psi_n^*(x)\psi_n(x)$$

That is for wave functions that satisfy the form

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

where $\psi_n(x)$ are solutions to the Time Independent Schrödinger Equation, the probability density has no time dependence. These are called **stationary states** or **eigenstates**. (Note that the wave function has time dependence, but not the probability density.)

One can find the stationary states of the system by solving the Time Independent Schrödinger Equation and adding the time dependent part.

There is nothing to say that a wave function has to be a stationary state, however if not, the probability will evolve in time.

If we solve the Time independent Schrödinger equation, $\psi^*(x)\psi(x)$ and hence $\Psi^*(x,t)\Psi(x,t)$ are independent of time. In this case, any calculations of probability or of average values need only use $\psi(x)$. These so-called *stationary state wave functions* and their corresponding energies are of central importance in chemistry and in spectroscopy!

Because the stationary states of the system are eigenfunctions of a Hermitian operator (*i.e.*, the Hamiltonian), they form a complete basis set. This means that we can write any wave function as a linear combination of stationary state wave functions.

I showed previously, however, that if we take a linear combination of solutions to the Schrödinger equation with different eigenvalues, they no longer have fixed energy. Such a linear combination is a perfectly valid wave function but does not solve the Schrödinger equation and hence will not be a stationary state.

So the general wave function which we can write as a linear combination of eigenfunctions of the Hamiltonian (or stationary state solutions)

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

is not itself a stationary state solution.

Example:

Let us again consider a particle-in-a-box where the wave function happens to be a linear combination of the first two eigenfunctions of the Hamiltonian, but now let us include the time dependent part explicitly:

$$\Psi(x,t)=\frac{1}{\sqrt{2}}\sqrt{\frac{2}{a}}\sin\left(\frac{\pi}{a}x\right)e^{-\frac{iE_1}{\hbar}t}+\frac{1}{\sqrt{2}}\sqrt{\frac{2}{a}}\sin\left(\frac{2\pi}{a}x\right)e^{-\frac{iE_2}{\hbar}t}$$

or more simply

$$\Psi(x,t)=\frac{1}{\sqrt{2}}\psi_1(x)e^{-\frac{iE_1}{\hbar}t}+\frac{1}{\sqrt{2}}\psi_2(x)e^{-\frac{iE_2}{\hbar}t}$$

Let us look at the probability density of this wave function:

$$\Psi^*(x,t)\Psi(x,t)=\frac{1}{\sqrt{2}}\left(\psi_1^*(x)e^{+\frac{iE_1}{\hbar}t}+\psi_2^*(x)e^{+\frac{iE_2}{\hbar}t}\right)\times\frac{1}{\sqrt{2}}\left(\psi_1(x)e^{-\frac{iE_1}{\hbar}t}+\psi_2(x)e^{-\frac{iE_2}{\hbar}t}\right)$$

$$=\frac{1}{2}\psi_1^*(x)\psi_1(x)+\frac{1}{2}\psi_2^*(x)\psi_2(x)+\frac{1}{2}\psi_1^*(x)\psi_2(x)e^{+\frac{i(E_1-E_2)}{\hbar}t}+\frac{1}{2}\psi_2^*(x)\psi_1(x)e^{-\frac{i(E_1-E_2)}{\hbar}t}$$

You can see that the time dependence **does not drop out** of the probability distribution.

Since the functions $\psi_1(x)$ and $\psi_2(x)$ for the particle in a box are real we can write

$$\psi_1^*(x)\psi_2(x)=\psi_2^*(x)\psi_1(x)=\psi_1(x)\psi_2(x)$$

and then rewrite the probability as

$$\Psi^*(x,t)\Psi(x,t)=\frac{1}{2}\left|\psi_1(x)\right|^2+\left|\psi_2(x)\right|^2+\frac{1}{2}\psi_1(x)\psi_2(x)\left[e^{+\frac{i(E_1-E_2)}{\hbar}t}+e^{-\frac{i(E_1-E_2)}{\hbar}t}\right]$$

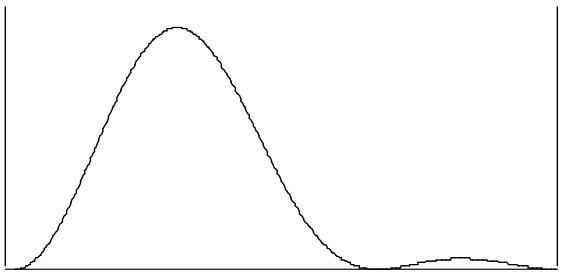
$$=\frac{1}{2}\left|\psi_1(x)\right|^2+\frac{1}{2}\left|\psi_2(x)\right|^2+\psi_1(x)\psi_2(x)\cos\left(\frac{E_1-E_2}{\hbar}t\right)$$

In this case, the probability distribution looks like the average of the distributions from each of the two functions plus an oscillatory cosine term. The frequency of the oscillation depends on the energy separation of the two states; *i.e*. the cosine oscillates with frequency $(E_1 - E_2)/\hbar$.

To get a physical feel for what is happening, let's look at the time dependence of such a probability distribution.

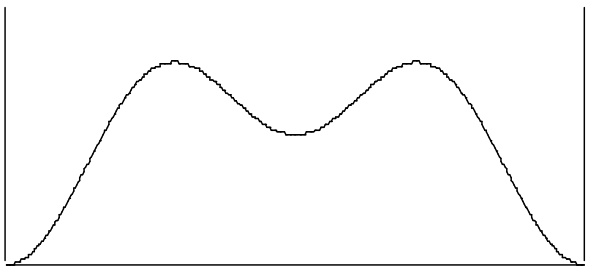
When the cosine term = 1

$$|\psi(x)|^2 = \frac{1}{2}|\psi_1(x)|^2 + \frac{1}{2}|\psi_2(x)|^2 + \psi_1(x)\psi_2(x) = \frac{1}{2}\left(\psi_1(x) + \psi_2(x)\right)^2$$



When the cosine term = 0

$$|\psi(x)|^2 = \frac{1}{2}|\psi_1(x)|^2 + \frac{1}{2}|\psi_2(x)|^2$$



Finally, when the cosine term = -1

$$|\psi(x)|^2 = \frac{1}{2}|\psi_1(x)|^2 + \frac{1}{2}|\psi_2(x)|^2 - \psi_1(x)\psi_2(x) = \frac{1}{2}\left(\psi_1(x) - \psi_2(x)\right)^2$$

So the probability distribution beats back and forth with a frequency $(E_1 - E_2)/\hbar$.

Before I discuss postulate 6, I want to take care of a few loose ends regarding some properties of operators that are important for further understanding measurements in quantum mechanical systems.

2.9 Some properties of operators

When more than one operator operates on a function, you apply the operator closest to the function first (the operator on the right), and then apply the second operator to the function that results from the first operation.

For example:

$$\hat{A}\hat{B}f(x) = \hat{A}\left(\hat{B}f(x)\right) = \hat{A}h(x) \qquad \text{here} \qquad h(x) = \hat{B}f(x)$$

Operators do not necessarily obey the communative law of multiplication.

That is $\hat{A}\hat{B}f(x) \neq \hat{B}\hat{A}f(x)$ in general.

If one can write an = sign here, we say that the operators ***commute***. Otherwise, we say they ***do not commute***.

Let us look at an example.

Let $\hat{A} = \dfrac{d}{dx}$ and $\hat{B} = x$

$$\hat{A}\hat{B}f(x) = \frac{d}{dx}xf(x) = f(x) + x\frac{d}{dx}f(x) = \left(\hat{I} + \hat{B}\hat{A}\right)f(x)$$

where $\hat{I}$ is the identity operator which multiplies a function by one.

So $\hat{A}\hat{B}f(x) \neq \hat{B}\hat{A}f(x)$

One way to characterize this property is to define what is called a ***commutator.***

The commutator of $\hat{A}$ and $\hat{B}$ is given by

$$\left[\hat{A},\hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

and is denoted by two operators in square brackets.

If the operators **commute** then $\left[\hat{A},\hat{B}\right] = 0$. If the operators **do not commute**, $\left[\hat{A},\hat{B}\right] \neq 0$.

The commutator defines a new operator and to evaluate this operator one must operate on a function.

That is $\left[\hat{A},\hat{B}\right]f(x)$

Let me give you a simple example:

Let $\hat{A} = \dfrac{d}{dx}$ and $\hat{B} = x^2$

$$\left[\hat{A},\hat{B}\right]f(x)=\hat{A}\hat{B}f(x)-\hat{B}\hat{A}f(x)$$

$$=\frac{d}{dx}\left(x^2f(x)\right)-x^2\frac{df(x)}{dx}$$

$$=x^2\frac{df(x)}{dx}+2xf(x)-x^2\frac{df(x)}{dx}$$
$$=2xf(x)$$

So $\left[\hat{A},\hat{B}\right]f(x)=2xf(x)$

Thus $\left[\hat{A},\hat{B}\right]=2x$

So we say that the operators $\hat{A}=\frac{d}{dx}$ and $\hat{B}=x^2$ do not commute.

It turns out that the commutator plays an important role in making measurements in quantum mechanical systems, and has a direct relationship to the Heisenberg Uncertainty Principle. To fully understand the connection between commutators and the measurement process, we need to introduce a theorem.

THEOREM: *If (Hermitian) operators* $\hat{A}$ *and* $\hat{B}$ *commute, they possess a common, complete set of eigenfunctions.*

Proof: Consider $\hat{A}$ and $\hat{B}$ such that $\left[\hat{A},\hat{B}\right]=0$,

Start with $\hat{B}\varphi_i=b_i\varphi_i$

If we operate on both sides with $\hat{A}$ we get:

$$\hat{A}\hat{B}\varphi_i=\hat{A}b_i\varphi_i$$

Since $\hat{A}$ and $\hat{B}$ commute we also have:

$$\hat{A}\hat{B}\varphi_i=\hat{B}\hat{A}\varphi_i$$

Consequently, we can write:

$$\hat{B}\left(\hat{A}\varphi_i\right)=\hat{A}\hat{B}\varphi_i=\hat{A}b_i\varphi_i=b_i\left(\hat{A}\varphi_i\right)$$

This says that if φ_i is an eigenfunction of $\hat{B}$ with eigenvalues b_i then the function $\hat{A}\varphi_i$ is also an eigenfunction of $\hat{B}$ with eigenvalue b_i.

If we assume that the eigenvalues of $\hat{B}$ are non-degenerate (no two the same), then b_i determines in a unique way the eigenfunction φ_i. (Remember we showed that eigenfunctions of a Hermitian operator form an orthonormal set. This means that eigenfunctions corresponding to different eigenvalues are orthogonal.) So then if the function $\hat{A}\ \varphi_i$ also has an eigenvalue b_i, it must be linearly dependent on φ_i, that is

$$\hat{A}\varphi_i=a\varphi_i$$

This will be true for every eigenfunction, and hence it forms a complete set. The inverse is also true.

Let us now explore the implications of this. Recall Postulates 3 and 4. Postulate 3 said that upon making a measurement of the quantity represented by $\hat{A}$, the only possible results are the eigenvalues of $\hat{A}$. Postulate 4 gave us a prescription for calculating average values:

$$\langle a \rangle = \int_{-\infty}^{\infty} \psi^*(x)\hat{A}\psi(x)dx$$

Recall what happens when the wave function $\psi(x)$ happens to be one of the eigenfunctions of the operator corresponding to the quantity that you measure.

Let us assume we are measuring the energy of a particle-in-a-box and the wave function is in an eigenfunction of the particle-in-a-box Hamiltonian.

$$\psi(x) = \psi_2(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{2\pi}{a}x\right)$$

What happens when we measure the energy? We could calculate the average value

$$\langle E \rangle = \int_{-\infty}^{\infty} \psi_2^*(x)\hat{H}\psi_2(x)dx = E_2\int_{-\infty}^{\infty} \psi_2^*(x)\psi_2(x)dx = E_2$$

(We could have done this explicitly by substituting in the proper Hamiltonian and operating on the function)

If we calculate the variance in our measurement, we would find that the variance equals zero.

$\rightarrow\rightarrow$ we would get E_2 every time we measured.

If the wave function is an eigenfunction of the operator you are measuring, then you **know** what the result will be

$\rightarrow\rightarrow$ the eigenvalue of that operator corresponding to the eigenfunction.

In this case, you do not need to take an average value. Just use the eigenvalue equation:

$$\hat{H}\psi_2(x) = E_2\psi_2(x)$$

where

$$E_2 = \frac{2^2h^2}{8ma^2}$$

Now let us consider the implications of the theorem we just introduced. Suppose a system is in a wave function that is an eigenfunction of an operator $\hat{A}$. But now, instead of measuring the quantity corresponding to $\hat{A}$ you measure a quantity which is represented by the operator $\hat{B}$ which commutes with $\hat{A}$.

The theorem I just introduced says that if $\hat{A}$ and $\hat{B}$ commute, they have a common set of eigenfunctions. Thus, when you measure $\hat{B}$, you know exactly what you will get: the eigenvalue of $\hat{B}$ that corresponds to that particular eigenfunction. There is no need to calculate an average value here.

Let me give you a specific example of how the commutator of two operators affects the measurement process in quantum mechanical systems.

Let us consider a particle which is fixed to move in a circular orbit of radius *a*. This problem is sometimes called a particle-on-a-ring and represents the rotation of a molecule in two dimensions. I will give you this problem to work out the details in your exercises.

The Hamiltonian operator for a particle constrained to move on a circle of radius *a* is given by:

$$\hat{H} = -\frac{\hbar^2}{2I}\frac{d^2}{d\theta^2}$$

where $I = ma^2$

The eigenfunctions of $\hat{H}$ for this system are:

$$\psi(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}$$

and the energies are:

$$E_n = \frac{n^2\hbar^2}{2I} = \frac{n^2\hbar^2}{2ma^2} \qquad \text{where} \quad n=0, \pm1, \pm2,....$$

(We could verify this simply by operating on the eigenfunction with the Hamiltonian.)

Let us say the system is in an eigenfunction of the particle on a ring Hamiltonian corresponding to n=+3. The wave function is then:

$$\psi(\theta) = \frac{1}{\sqrt{2\pi}} e^{i3\theta}$$

If we made a measurement of the energy we know exactly what we will get, E_3.

Let us just show this.

$$\hat{H} = -\frac{\hbar^2}{2I}\frac{d^2}{d\theta^2}$$

$$\hat{H}\psi(\theta) = -\frac{\hbar^2}{2I}\frac{d^2}{d\theta^2}\frac{1}{\sqrt{2\pi}} e^{i3\theta} = \frac{9\hbar^2}{2I}\frac{1}{\sqrt{2\pi}} e^{i3\theta} = E_3\psi(\theta)$$

But instead of measuring energy, let us measure angular momentum.

By applying the prescription of postulate 2, we could easily determine that the angular momentum operator is

$$\hat{L} = -i\hbar\frac{d}{d\theta}$$

We can also easily show that this operator commutes with the Hamiltonian for the particle on a ring, that is $\left[\hat{H},\hat{L}\right] = 0$

$$\left[\hat{H},\hat{L}\right] f(\theta) = \left(\hat{H}\hat{L} - \hat{L}\hat{H}\right) f(\theta)$$

$$= \frac{i\hbar^3}{2I}\left(\frac{d^2}{d\theta^2}\frac{df(\theta)}{d\theta} - \frac{d}{d\theta}\frac{d^2 f(\theta)}{d\theta^2}\right)$$

$$= \frac{i\hbar^3}{2I}\left(\frac{d^3 f(\theta)}{d\theta^3} - \frac{d^3 f(\theta)}{d\theta^3}\right)$$

$$= 0 f(\theta)$$

So $\left[\hat{H},\hat{L}\right]=0$

Just as an aside, it is a general rule that the commutator of an operator with that operator raised to a power will be zero.

That is $\left[\hat{A},\hat{A}^{n}\right]=0$

So we could have simply looked at the commutator and known it equals zero.

Since $\left[\hat{H},\hat{L}\right]=0$, from the theorem we introduced, if the system is in a state that is an eigenfunction of the angular momentum operator, $\hat{L}$, it is also in an eigenfunction of the Hamiltonian, $\hat{H}$. If we make a measurement of the angular momentum, we know what we will get, the eigenvalue of $\hat{L}$. However, if we make a measurement of the energy, we also know exactly what we will get, the eigenvalue of $\hat{H}$ that corresponds to that eigenfunction. There will be no uncertainty in repeated measurements, *i.e.* the variance is zero, because the system is in an eigenfunction of that operator.

That is $\Delta E\,\Delta L=0$

There is no uncertainty in the relative precision with which these two operators can be measured.

One can prove that there exists a direct relationship between the value of the commutator of two operators and the relative precision with which the corresponding observable quantities can be measured. The prove goes as follows.

Let us consider two observables *A* and *B* with average values,

$$\langle A\rangle=\int\psi^{*}\hat{A}\psi\,d\tau \qquad \text{and} \qquad \langle B\rangle=\int\psi^{*}\hat{B}\psi\,d\tau$$

where ψ is any acceptable wavefunction.

Let us define $$\Delta A=\sqrt{\sigma_A^2}=\sqrt{\left\langle\left(A-\langle A\rangle\right)^2\right\rangle}=\sqrt{\langle A^2\rangle-\langle A\rangle^2}$$

and $$\Delta B=\sqrt{\sigma_B^2}=\sqrt{\left\langle\left(B-\langle B\rangle\right)^2\right\rangle}=\sqrt{\langle B^2\rangle-\langle B\rangle^2}$$

We now introduce new Hermitian operators defined as:

$$\hat{\overline{A}}=\hat{A}-\langle A\rangle \qquad \text{and} \qquad \hat{\overline{B}}=\hat{B}-\langle B\rangle$$

From this follows that

$$\left(\Delta A\right)^2=\left\langle\overline{A}^2\right\rangle \qquad \text{and} \qquad \left(\Delta B\right)^2=\left\langle\overline{B}^2\right\rangle$$

Define a linear (not Hermitian) operator,

$$\hat{C}=\hat{\overline{A}}+i\lambda\hat{\overline{B}}$$

with λ being a real number, and its complex conjugate,

$$\hat{C}^{*}=\hat{\overline{A}}-i\lambda\hat{\overline{B}}$$

We find that

$$\left\langle CC^* \right\rangle = \left\langle \left(\overline{A} + i\lambda\overline{B}\right)\left(\overline{A} - i\lambda\overline{B}\right)\right\rangle = \left\langle \overline{A}^2 + \lambda^2\overline{B}^2 - i\lambda\left(\overline{A}\overline{B} - \overline{B}\overline{A}\right)\right\rangle = \left\langle \overline{A}^2 + \lambda^2\overline{B}^2 - i\lambda\left[\overline{A},\overline{B}\right]\right\rangle$$

Since $\left\langle CC^* \right\rangle = \int \psi^* \hat{C}\hat{C}^* \psi d\tau = \int \left(\hat{C}^*\psi\right)^* \hat{C}^* \psi d\tau \geq 0$ and real

we have that

$$f(\lambda) = \left\langle \overline{A}^2 + \lambda^2\overline{B}^2 - i\lambda\left[\overline{A},\overline{B}\right]\right\rangle = \left\langle \overline{A}^2\right\rangle + \lambda^2\left\langle \overline{B}^2\right\rangle - i\lambda\left\langle \left[\overline{A},\overline{B}\right]\right\rangle = \left(\Delta A\right)^2 + \lambda^2\left(\Delta B\right)^2 - i\lambda\left\langle \left[A,B\right]\right\rangle \geq 0$$

is real and non-negative and consequently $\left\langle \left[A,B\right]\right\rangle$ is purely imaginary.

The function $f(\lambda)$ will have a minimum. We can find this minimum by setting

$$\frac{df(\lambda)}{d\lambda} = 2\lambda\left(\Delta B\right)^2 - i\left\langle \left[A,B\right]\right\rangle = 0$$

giving $\lambda_{\min} = \dfrac{i\left\langle \left[A,B\right]\right\rangle}{2\left(\Delta B\right)^2}$

with which we find that

$$f(\lambda_{\min}) = \left(\Delta A\right)^2 + \left(\frac{i\left\langle \left[A,B\right]\right\rangle}{2\left(\Delta B\right)^2}\right)^2 \left(\Delta B\right)^2 - i\frac{i\left\langle \left[A,B\right]\right\rangle}{2\left(\Delta B\right)^2}\left\langle \left[A,B\right]\right\rangle = \left(\Delta A\right)^2 + \frac{1}{4}\frac{\left\langle \left[A,B\right]\right\rangle^2}{\left(\Delta B\right)^2} \geq 0$$

Consequently,

$$\left(\Delta A\right)^2\left(\Delta B\right)^2 \geq -\frac{1}{4}\left\langle \left[A,B\right]\right\rangle^2 = -\frac{1}{4}\left(\int \psi^* \left[A,B\right]\psi d\tau\right)^2$$

Using the expression before

As an example, consider the commutator $\left[\hat{P}_x,\hat{x}\right]$.

Let us calculate the value of this commutator. Remember the golden rule in evaluating commutators: You must always operate on a function. You can do it other ways, but is very easy to make a mistake.

$$\left[\hat{P}_x,\hat{x}\right] = \left[-i\hbar\frac{d}{dx},x\right]$$

To evaluate:

$$\left[-i\hbar\frac{d}{dx},x\right]f(x) = -i\hbar\left(\frac{d}{dx}xf(x) - x\frac{df(x)}{dx}\right)$$

$$= -i\hbar\left(x\frac{df(x)}{dx} + f(x) - x\frac{df(x)}{dx}\right)$$

$$= -i\hbar f(x)$$

So, $\left[\hat{P}_x,\hat{x}\right] = -i\hbar$

If we now substitute this into our equation for the relative variance we get

$$(\Delta P_x)^2(\Delta x)^2 \geq -\frac{1}{4}\left(\int \psi^*(x)(-i\hbar)\psi(x)dx\right)^2$$

$$\geq -\frac{1}{4}(-i\hbar)^2 = \frac{\hbar^2}{4}$$

By taking the square root of both sides we get

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

which is known as the **Heisenberg Uncertainty Principle**.
So the Heisenberg uncertainty principle relationship does not only hold for position and momentum, but it is a more general rule that pertains to many other pairs of observables. It depends upon the value of their commutator.

If the commutator of two operators is zero (*i.e.,* if they commute), you can measure the observables that correspond to those operators with infinite relative precision.

2.10 Postulate 6

I would like to finish our treatment of the formal postulates of Quantum Mechanics by continuing a little further with our discussion of measurements in quantum mechanical systems. We often talk about quantum mechanical systems like the particle-in-a-box or a particle-on-a-ring being in one particular state or another, but how does the system get into that state initially (or, how does one know what state it is in to start)??

The answer is that the state of the system is determined by the measurement process itself. That is the state of a system is determined by the last measurement on that system.

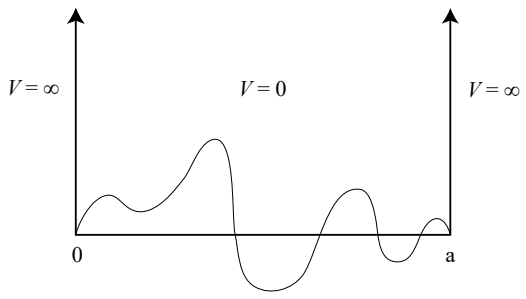
There is one more postulate of quantum mechanics that deals with this.

Postulate 6

Immediately after measuring the quantity corresponding to observable A and getting the result a_n, the wave function representing the state of the system becomes φ_n where $\hat{A}\varphi_n = a_n\varphi_n$. That is, it becomes an eigenfunction of the operator $\hat{A}$.

The system forgets where it was and finds itself in φ_n.

For example, consider the particle in a box. Let us say the original state $\Psi(r,t)$ looks something like this:



Although this is not one of the solutions to the time-independent Schrodinger equation for the particle-in-a-box, it is a perfectly acceptable wave function. There is nothing to say that the wave function Ψ has to be an eigenfunction of the particle-in-a-box Hamiltonian. In general it can be represented by a linear combination of eigenfunctions. We can choose these eigenfunctions however we want, as long as it is a complete, orthonormal set.
Let us choose to expand them in eigenfunctions of the particle-in-a-box Hamiltonian.

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

The exponential comes from the fact that we know the time dependence of the $\Psi_n(x,t)$.

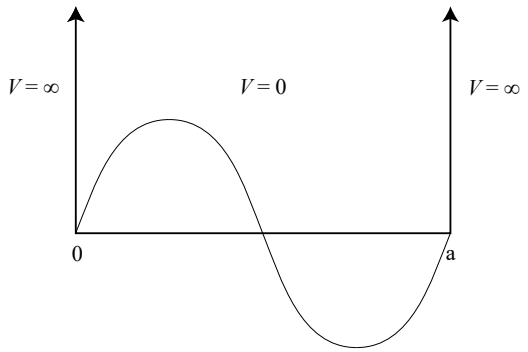
Now, let us say we make a measurement of the energy of the system. The energy we measure will be one of the eigenvalues of the particle-in-a-box Hamiltonian (Postulate 3).

The probability of obtaining a particular result, say E_2 for example, is given by

$$P_2=c_2e^{\frac{-iE_2t}{\hbar}}\cdot c_2^*e^{\frac{+iE_2t}{\hbar}}=|c_2|^2$$

Let's say we make a measurement of the energy and get the result E_2.

Postulate 6 says that the wave function ψ becomes the particle-in-a- box ψ_2. Our wave function now looks like this:



The wavefunction of the system is than given by:

$$\Psi(x,t)=\psi_2(x)e^{-\frac{iE_2}{\hbar}t}$$

Will the system remain in this state in subsequent measurements of the energy?

Inspection of the wavefunction learns us that:

$$P_2=c_2e^{\frac{-iE_2t}{\hbar}}\cdot c_2^*e^{\frac{+iE_2t}{\hbar}}=|c_2|^2=1$$

and

$$P_{n\neq2}=0$$

Thus, if we make a subsequent measurement of the energy on the same system, we know exactly what result we will get – E_2, since the system is in an eigenfunction of the Hamiltonian ψ_2. According to postulate 6, immediately after this measurement the system will be in an eigenfunction of the operator we measured

corresponding to the eigenvalue we obtained. That is: it remains in the same state that it was in before the second measurement.

What if we subsequently measure a quantity whose operator commutes with the Hamiltonian?

If the operator commutes with the Hamiltonian, they have a common complete set of eigenfunctions. We know exactly what result the measurement will give -- the eigenvalue of the operator we measure corresponding to the eigenfunction that the state is in. After the measurement it stays in the same state, since it is put in an eigenfunction of the operator we measure, and that function is also an eigenfunction of the operator we originally measured.

What happens if we first measure the energy and then measure something whose operator does not commute with the Hamiltonian?

For example, we could measure the energy of a particle in a box and then measure its position.

Since $\left[\hat{H}, \hat{x}\right] \neq 0$, then $\hat{x}$ and $\hat{H}$ do not have a common set of eigenfunctions.

When we make a measurement of *x*, postulate 6 says that the system forgets where it was and finds itself in an eigenfunction of the position operator. We haven't talked about what this function looks like, but it is not an eigenfunction of $\hat{H}$. Since this is ***not*** an eigenfunction of $\hat{H}$, when we go back and measure the energy we are not exactly sure what we will get. There will be some probability for getting any one of the energy eigenvalues.

Mathematically this can be simply demonstrated. Although the eigenfunctions of $\hat{x}$ are not eigenfunctions of $\hat{H}$, they can be written as a linear combination of $\hat{H}$ eigenfunctions:

$$\mathrm{X}(x,t) = \sum_n c_n \Psi_n(x,t) = \sum_n c_n \psi_n(x) e^{-\frac{iE_n}{\hbar}t} = \sum_n c_n \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) e^{-\frac{iE_n}{\hbar}t}$$

One can then see that the probability of getting a particular result upon remeasuring the energy is:

$$P_n = \left|c_n\right|^2$$

So we need not get the same result that we got last time we measured the energy. How could the system have changed energy?? **Through the interaction of the measurement!**

This is a fairly hard concept to get a handle on, and it takes a little practice to get it right.

At the risk of making it a little more confusing, let me take this one step further in looking at the role of Postulate 6 on the measurement process.

Let us say we start with a particle in a box and have no idea of what state we are in. We then make a measurement of the energy and find it in E_2, the particle-in-a-box energy corresponding to n=2.

Now let us say that we measure the energy again. What will we get? E_2 again since the system is still in ψ_2 as we saw before.

Now let us measure the position. Since the wave function before the measurement is ψ_2, we know what the probability distribution is for finding the particle at a particular place. Let's say we find it at $a/4$.

Ten seconds later let's measure the position again. Where will we find it? Will it still be at $a/4$? Will it be somewhere else with certainty? Or can it have a distribution of values?

The answer $\rightarrow\rightarrow$ it will have a distribution of values!! The reason is that the wavefunction evolves in time according to:

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

Now notice what happened. When we measured the energy twice in a row we got the same result. However when we measured the position twice in a row we did not necessarily get the same result.

Why is this? Is there something special about measuring the energy?

The answer is **yes!!**

Recall that according to Postulate 5 the wave function $\Psi(x,t)$ evolves in time according to the time dependent Schrödinger equation. Using the method of separation of variables to separate the spatial and time dependent parts we found that the part that just depends on position is given by:

$$\hat{H}\psi(x) = E\psi(x)$$

Note that this is the time independent Schrödinger equation we use to determine the energy of the system.

The time dependent part of the wave function is given by:

$$f(t) = e^{-\frac{iE}{\hbar}t}$$

The total wave function is then

$$\Psi(x,t) = \psi(x) e^{-\frac{iE}{\hbar}t}$$

Remember we called solutions of this form STATIONARY STATE WAVE FUNCTIONS.

This is because the probability density is independent of time, and any average value will be independent of time since any integral involving the product of the complex exponential and its complex conjugate will be independent of time. We can therefore simply use the time independent solution to the Schrödinger equation for calculating average quantities. But these solutions are simply the eigenfunctions of the Hamiltonian.

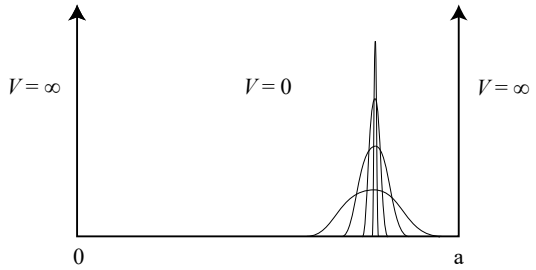
So any time the state of the system becomes one of the eigenfunctions of the Hamiltonian, the wave function has the simple exponential time dependence. Since this state is already a steady state solution to the time dependent Schrödinger equation, we know how it evolves in time--trivially (just a phase factor).

If we measure something which commutes with the Hamiltonian, we have the same situation.

But, if we measure something that does not commute with the Hamiltonian, then the state is changed to a function that is not an eigenfunction of the Hamiltonian and hence will not be a stationary state solution.

Let us say you measure the position. Postulate 6 says that the state of the system becomes an eigenfunction of the position operator. That will not also be an eigenfunction of the Hamiltonian, but can be written as a linear combination of Hamiltonian eigenfunctions. The probability distribution will have a non-trivial time dependence. (REMEMBER that we did an example in class in which we formed a wave function as a superposition of two wave functions and showed that the time dependence didn't drop out).

The time evolution of the eigenfunctions of the position operator is displayed graphically below:



The exact form of the time evolution of the wave function is given by Postulate 5: The time dependent Schrödinger Equation.

But if you subsequently measure the energy or any quantity whose operator commutes with the Hamiltonian, the wave function goes back to the simple complex exponential time dependence (and hence probability distribution stops evolving in time).

Example: Consider the particle on a ring .

1. First we measure the energy and get E_1. Next, we measure the angular momentum. What will we get? A distribution of values or just a single result? Why? What was the wave function before we measured? What was it afterward?

2. Now measure the energy again. What will we get?

3. Now measure the angular momentum again. What will we get this time? (What is the wavefunction after each measurement?)

4. Now we measure the angular position and get some angle θ_0. If we measure it again will we get the same result, why?

5. Now we measure the energy again. What will we get? One result or a distribution?

Now that I have finished my introduction of the postulates of quantum mechanics and you have some idea of the basic principles, we will begin to look at the time independent Schrödinger equation for systems that are relevant to chemistry. We will start with 1-dimensional problems and then go to increasingly more complex systems. Along the way I will introduce a number of new concepts that will allow us to go to the next level of complexity.

Our focus will be on determining the eigenvalues and eigenfunctions for chemically relevant systems.

The principles that we have discussed thus far regarding measurements and time dependence in quantum mechanics are independent of the specific systems that we discuss—they hold for all quantum mechanical systems.

3 The Harmonic Oscillator

What we have been doing over the last few weeks has been somewhat abstract and at times seemed far removed from chemistry, although as you will see throughout the year, it will be quite relevant.

Today I would like to begin a discussion of a problem in quantum mechanics that has clear and important practical implications in chemistry: *the harmonic oscillator.*

Its importance is derived from the fact that to a good approximation, vibrational motions of molecules can be modeled as harmonic oscillators. As we will see more clearly later in the course, transitions between vibrational energy levels of molecules (*i.e.*, the eigenstates of the Hamiltonian for vibrational motion) lead to absorption of light in the infrared region of the spectrum. The manner in which a molecule absorbs infrared light is a fingerprint of that molecule and hence a valuable tool for molecular identification. Moreover, properties such as heat capacities and chemical reaction rates are strongly linked to a molecules vibrational motion. Thus, a good place to begin to understand molecular vibrations is the harmonic oscillator problem.

The energy levels of a diatomic molecule are very closely predicted by solving the quantum mechanics of a harmonic oscillator, and while those of polyatomic molecules may seem quite a bit more complicated, to first order they can be considered as a superposition of harmonic oscillator energy levels. So the principles we will learn here will be applicable not only to diatomics but to polyatomics as well.

Brief outline of what we will do:

1. After defining what we mean by a harmonic oscillator, we will treat the problem using classical mechanics.

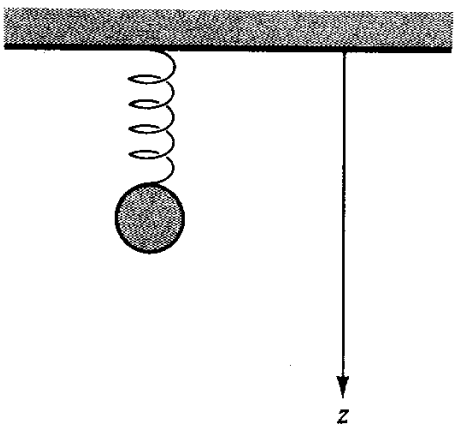
2. We will then introduce the solutions of the time-independent Schrödinger equation for a harmonic oscillator and look at the nature and properties of the eigenvalues and eigenfunctions. Here we will introduce a few new concepts that I did not mention in our previous simple models.

3. We will then go back and work through the mathematics required to arrive at the solution. The form of the Schrödinger equation is more difficult than those we have already solved, and we will have to introduce some new mathematical approaches to solve it.

3.1 Classical Harmonic Oscillator

Let us begin by defining what we mean by a harmonic oscillator.

As depicted below, consider a mass connected to a wall by a spring. The only force on the mass is from the spring.



As you would expect, the force on the mass will be some function of the displacement of the spring from its equilibrium value.

Let the displacement of the spring from its equilibrium value be denoted by x so that $x = z - l_0$. The *harmonic approximation* states that the force on the mass is simply proportional to the displacement from the equilibrium position, that is

$$F = -k(z - l_0) = -kx$$

This is a statement that the spring obeys *Hooke's Law*.

The negative sign indicates that the force acts to restore the mass to its equilibrium position. If we define positive z as the downward direction, then when z is greater than l_0, the force is negative; that is, it acts in the upward direction restoring the mass to its equilibrium position. If $z<l_0$, the force is positive and pushes the mass down.

The proportionality constant k represents the stiffness of the spring. A very high value of k would represent a stiff spring, which would require a large force to compress or extend, and a low value of k represents a loose, floppy spring, which is easy to compress or extend.

To solve the classical problem, we start with Newton's second law, $F = ma$.

$$F = ma = m\frac{d^2z(t)}{dt^2} = -k\left(z(t) - l_0\right)$$

We have simply set the mass times the acceleration equal to the Hooke's Law force.

Using $x(t) = z(t) - l_0$

We find that $\dfrac{d^2z(t)}{dt^2} = \dfrac{d^2x(t)}{dt^2}$

so $m\dfrac{d^2x(t)}{dt^2} = -kx(t)$

$$\frac{d^2x(t)}{dt^2} + \frac{k}{m}x(t) = 0$$

This is a linear second order differential equation with constant coefficients.

Let us guess a solution of the form:

$$x(t) = e^{\alpha t}$$

Substituting this into the differential equation:

$$\alpha^2 e^{\alpha t} + \frac{k}{m}e^{\alpha t} = 0$$

$$\alpha^2 + \frac{k}{m} = 0$$

$$\alpha^2 = -\frac{k}{m} \qquad \Rightarrow \qquad \alpha = \pm i\sqrt{\frac{k}{m}}$$

Let's set

$$\omega = \sqrt{\frac{k}{m}} \qquad \text{so} \qquad x(t) = ce^{\pm i\omega t}$$

The most general solution is

$$x = c_1 e^{+i\omega t} + c_2 e^{-i\omega t}$$

Remember we showed earlier in the course that using Euler's formula we could write this

$$x(t) = A\cos(\omega t) + B\sin(\omega t)$$

where A and B are just combinations of c_1 and c_2.

We can evaluate the constants A and B by considering the initial conditions (*i.e.* the conditions at t=0). Suppose we stretch the spring to a length z_0 so that its initial displacement is $x_0 = z_0 - l_0$ and then let it go.

$$x(t=0) = x_0 = A\cos(0) + B\sin(0) \qquad \Rightarrow \qquad A = x_0$$

Thus

$$x(t) = x_0\cos(\omega t) + B\sin(\omega t)$$

The initial velocity is zero since we start at rest. The velocity is just given by

$$\mathrm{v}(t) = \frac{dx(t)}{dt} = -x_0\omega\sin(\omega t) + B\omega\cos(\omega t)$$

so

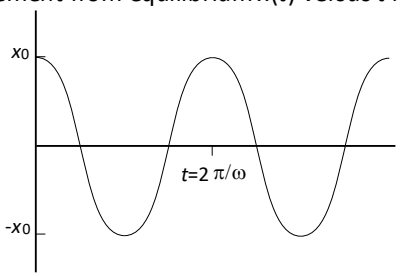
$$\mathrm{v}(t=0) = 0 = -x_0\omega\sin(0) + B\omega\cos(0) \qquad \Rightarrow \qquad B = 0$$

Thus

$$x(t) = x_0\cos(\omega t)$$

where $\omega = \sqrt{\frac{k}{m}}$ from the boundary conditions given above.

If we plot the displacement from equilibrium $x(t)$ versus t it looks like the following:



The displacement of the mass oscillates between x_0 and $-x_0$ with a frequency of ω radians/sec or $\omega/2\pi$ cycles/sec.

Let us look at the potential, kinetic, and total energy of the harmonic oscillator.

Remember the force is given by:

$$F = -kx$$

From classical mechanics, we know that a force can be expressed as a derivative of the potential energy

$$F(x) = -\frac{dU(x)}{dx}$$

where $U(x)$ is the potential.

We can therefore express $U(x)$ as:

$$U(x) = -\int F(x)dx$$

$$= \int kx\,dx$$

$$= \frac{1}{2}kx^2 + c$$

The integration constant here is arbitrary and can be used to fix the absolute zero of energy. This is usually taken to be zero when $x = 0$.

Therefore the potential energy of a harmonic oscillator is

$$U(x) = \frac{k}{2}x^2$$

or, as a function of time:

$$U(t) = \frac{k}{2}x_0^2\cos^2(\omega t)$$

The kinetic energy can be given by

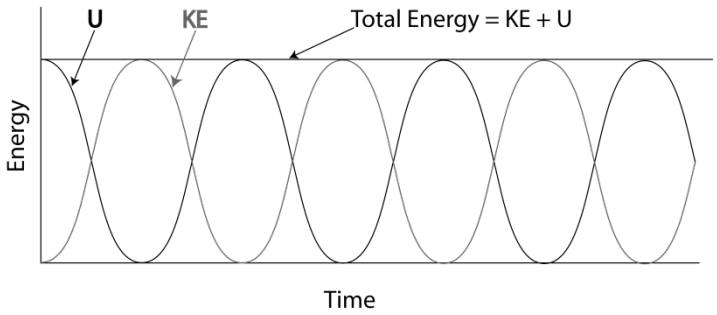
$$K(t) = \frac{1}{2}m\left(\frac{dz(t)}{dt}\right)^2 = \frac{1}{2}m\left(\frac{dx(t)}{dt}\right)^2$$

Using $x(t) = x_0\cos(\omega t)$

gives $\frac{dx}{dt} = -\omega x_0\sin(\omega t)$

So we get $K(t) = \frac{1}{2}m\omega^2 x_0^2\sin^2(\omega t)$

We can plot both the potential energy and kinetic energy:



The fact that the system starts with all potential energy and no kinetic energy comes from our choice of initial conditions (we said that we would stretch the string holding it still and then let go.)

You can see that the potential and kinetic energy are 180° out of phase. The energy transfers back and forth between being all kinetic energy and no potential and then all potential energy and no kinetic. Where the energy starts is determined by the initial conditions.

Note that when the energy is all potential, the spring is at one of its turning points: the mass is turning around and going the other way. When the energy is all kinetic is when the mass passes through its equilibrium configuration.

You can see that because of the phase difference, the total energy remains constant, that is *total energy is conserved*. The system is called a *conservative system*. This will be the case whenever the force can be written as a derivative of the potential. (Cases that *do not* fall into this category are those that have forces like frictional or viscous forces.)

To show quantitatively that the energy of a classical harmonic oscillator remains constant, we need only sum our expressions for U and K.

$$E(t) = U(t) + K(t) = \frac{k}{2}x_0^2\cos^2(\omega t) + \frac{1}{2}m\omega^2 x_0^2\sin^2(\omega t)$$

Remembering that

$$\omega = \sqrt{\frac{k}{m}}$$

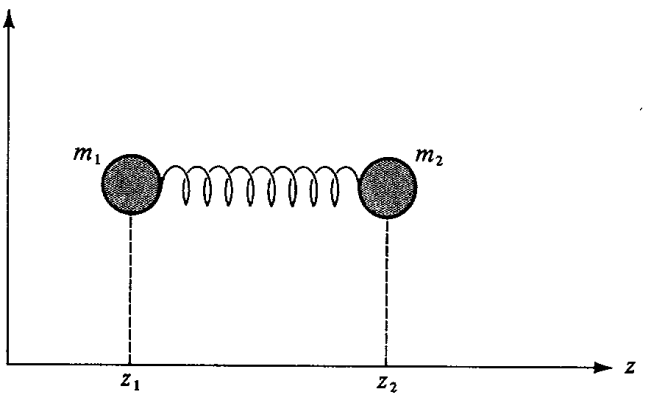
we can write

$$E(t) = \frac{kx_0^2}{2}\left[\cos^2(\omega t) + \sin^2(\omega t)\right]$$

$$= \frac{kx_0^2}{2}$$

So the total energy is a constant and is equal to the potential energy at its turning point.

We now want to make the Harmonic Oscillator problem look more like a diatomic molecule; not just a mass attached to a wall.



Consider two masses connected by a spring. We now get two equations of motion, one for each mass.

$$m_1 \frac{d^2 z_1(t)}{dt^2} = k\left(z_2(t) - z_1(t) - l_0\right)$$

$$m_2 \frac{d^2 z_2(t)}{dt^2} = -k\left(z_2(t) - z_1(t) - l_0\right)$$

Notice the sign convention:

$z_2(t) - z_1(t) > l_0$ the spring is stretched

$z_2(t) - z_1(t) < l_0$ the spring is compressed

Note that the force on each mass is in the direction to restore it to the equilibrium position. Notice also that the force on mass 1 is equal and opposite from the force on mass 2. This must be so (Newton's third law)

This means that

$$m_1 \frac{d^2 z_1(t)}{dt^2} + m_2 \frac{d^2 z_2(t)}{dt^2} = 0$$

Or

$$\frac{d^2}{dt^2}\left(m_1 z_1(t) + m_2 z_2(t)\right) = 0$$

Let us introduce what is called the *center-of mass-coordinate*.

$$Z(t) = \frac{m_1 z_1(t) + m_2 z_2(t)}{M}$$

where $M = m_1 + m_2$

We can then write the equation above as:

$$M \frac{d^2 Z(t)}{dt^2} = 0$$

This means that the acceleration of the particle as a whole equals zero. (Remember we said that the force was due only to the spring). Thus, the whole system moves through space at constant velocity (no acceleration).

The motion of the two mass system *must* depend only on the relative separation of the two masses, *z*, where

$$z(t) = z_2(t) - z_1(t)$$

We can take our original differential equations, divide them by their relative masses, and subtract them

$$\frac{d^2 z_2(t)}{dt^2} - \frac{d^2 z_1(t)}{dt^2} = -\frac{k}{m_2}\left(z_2(t) - z_1(t) - l_0\right) - \frac{k}{m_1}\left(z_2(t) - z_1(t) - l_0\right)$$

$$\frac{d^2}{dt^2}\left(z_2(t) - z_1(t)\right) = -k\left(\frac{1}{m_1} + \frac{1}{m_2}\right)\left(z_2(t) - z_1(t) - l_0\right)$$

Using the definition of a reduced mass,

$$\frac{1}{m_1}+\frac{1}{m_2}=\frac{m_1+m_2}{m_1 m_2}=\frac{1}{\mu}$$

we can write this as

$$\frac{d^2 z(t)}{dt^2}=-\frac{k}{\mu}\left(z(t)-l_0\right)$$

If we now let

$$x(t)=z(t)-l_0$$

we finally arrive at:

$$\mu\frac{d^2 x(t)}{dt^2}+kx(t)=0$$

This is an important result. This result is the same as the mass attached to a wall, but the mass is replaced by the reduced mass μ. You can see that if m_1 or $m_2 \to \infty$ then $\mu \to m_2$ or m_1 and the equation would be the same as before.

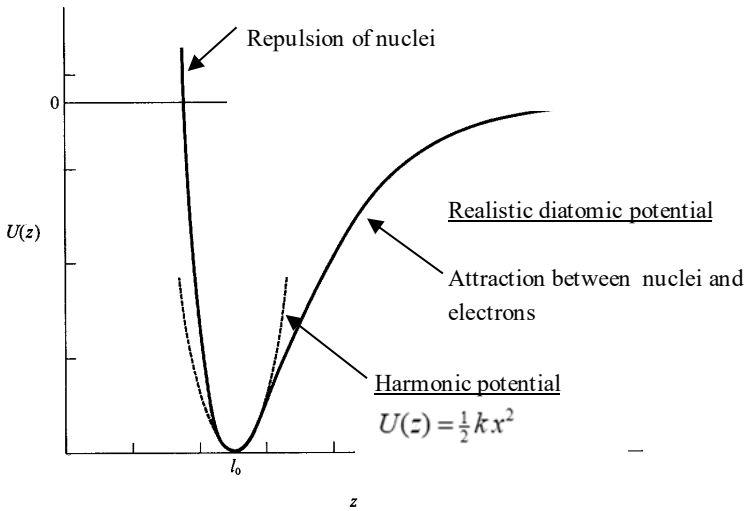
The fact that $m \to \mu$ means that the frequency will now be

$$\omega=\sqrt{\frac{k}{\mu}}$$

In general, if the potential only depends upon relative coordinates, then we can separate off the center of mass motion and reduce the two-body problem to a one body problem with mass equal to μ. We will therefore use μ in the rest of our discussion.

Before we go on to discuss the quantum mechanical harmonic oscillator problem, it is important to consider how good such a harmonic oscillator ball and spring model is for a vibrating diatomic molecule. Below is a typical internuclear potential for a diatomic molecule. At small internuclear separation, the potential rises sharply. This is due to the repulsion of the positively charged nuclei. The well is due to the balance between the nuclear repulsion and the attraction between nuclei and electrons. The flattened out part at large internuclear distance z indicates that as a bond is stretched, the restoring force is no longer linear because the chemical bond begins to break. The asymptotic energy is the bond dissociation energy.

We can superimpose a harmonic potential on this typical internuclear potential curve:



You can see that in some sense it is rather unrealistic to model this potential by a harmonic oscillator in that it takes an infinite amount of energy to break the bond (*i.e*. go to an infinite internuclear separation). However, the shape of the well near the bottom does a very good job of fitting the real potential, and as we will see when we solve the quantum mechanics of this problem, most of the energy levels that one observes experimentally are in the part of the well which is fit well by a harmonic potential.

So the harmonic oscillator approximation is good for small amplitude vibrations where Hooke's Law holds (*i.e*., those near the bottom of the well).

To put this into more mathematical terms, we could write our real potential function as a Taylor series in *z* about the equilibrium position, l_0.

$$U(z) = U(l_0) + \left(\frac{dU(z)}{dz}\right)_{z=l_0}(z-l_0) + \frac{1}{2!}\left(\frac{d^2U(z)}{dz^2}\right)_{z=l_0}(z-l_0)^2 + + \frac{1}{3!}\left(\frac{d^3U(z)}{dz^3}\right)_{z=l_0}(z-l_0)^3 + \ldots$$

The first term determines the absolute energy at the bottom of the well. This is not very important since we usually look at the difference in energy between two levels. We will therefore set it to zero.

The second term is the slope of the curve (first derivative), and by definition, this must vanish near the bottom of the well.

If we let

$$\left(\frac{d^2U(z)}{dz^2}\right)_{z=l_0} = k \qquad \text{and} \qquad \left(\frac{d^3U(z)}{dz^3}\right)_{z=l_0} = \gamma$$

we can write

$$U(z) = \frac{1}{2}k(z-l_0)^2 + \frac{1}{6}\gamma(z-l_0)^3 + \ldots$$

or

$$U(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \ldots$$

where $x = z - l_0$.

If the displacement from equilibrium is small, *x* is small and we can neglect the x^3 term.

We are then left with

$$U(x) = \frac{1}{2}kx^2$$

which is the potential for the Harmonic Oscillator.

This shows that the Harmonic Oscillator should be a good approximation for small amplitude vibrations. One can make corrections to account for "anharmonic" terms later.

It is important to realize what the meaning of *k* in this expression is, *i.e*. it is related to the curvature of the potential well at the minimum.

3.2 Quantum Mechanical Harmonic Oscillator Problem

Remember I said that I would introduce the solutions to the problem first and look at their physical significance. In doing so, I will introduce a few new concepts that we have not seen before. After having looked at the

solutions, we will go back and work through the mathematics required to arrive at the solution. The form of the Schrödinger equation is more difficult, and we will need some new mathematical tools to solve it.

Even though we will not solve the problem right now, let us write down the Schrödinger equation.

$$\hat{H}\psi(x) = E\psi(x)$$

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

Recall that μ is the reduced mass and using it allows us to reduce a two-body problem to that of single body of mass μ.

We can put in

$$U(x) = \frac{1}{2}kx^2$$

and rearrange to get

$$\frac{d^2\psi(x)}{dx^2} + \frac{2\mu}{\hbar^2}\left(E - \frac{1}{2}kx^2\right)\psi(x) = 0$$

This differential equation is more difficult to solve than those we have done previously in that it does not have constant coefficients. There is no "simple" way to solve it. We will present the results first and then go back and solve this differential equation.

It turns out that you only get well-behaved finite solutions if the energy is quantized. We will see shortly that this quantization of energy occurs when applying the **boundary conditions.**

The result for the energy is:

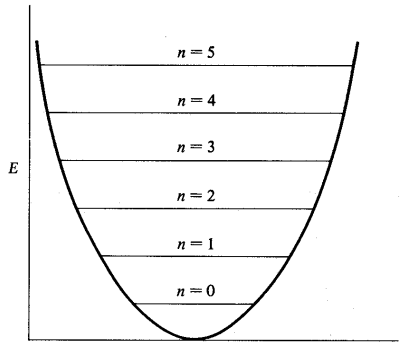
$$E_n = \hbar\sqrt{\frac{k}{\mu}}(n + \tfrac{1}{2})$$

$$= \hbar\omega(n + \tfrac{1}{2}) = h\nu(n + \tfrac{1}{2}) \qquad n = 0,1,2,3,\ldots$$

where $\omega = \sqrt{\frac{k}{\mu}}$ and $\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$

Note that ν (or ω) is the *classical* expression for the frequency of the harmonic oscillator.

I will superimpose these energy levels on the potential energy curve:



There are several important things to note about the harmonic oscillator energy levels:

1) The energy levels are equally spaced, in integral units of the classical frequency.

Think about what this means. The classical frequency is related to the force constant *k*

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$$

Remember that *k* is related to the width of the potential (which is a parabola).

$$U(x) = \frac{1}{2}kx^2$$

Larger *k* means stiffer spring (*i.e*. the energy rises faster with *x* the higher *k* is). So a stiffer spring (narrower parabola) means that the energy levels are spaced more widely. Conversely, a smaller *k* means a floppier spring (wider parabola) and more closely space energy levels.

The harmonic oscillator potential occupies a unique position among simple one-dimensional potentials in that the energy levels are spaced evenly. Any potential that has more curvature than the harmonic oscillator potential will have levels which increase in spacing as you go higher in energy. A good example is the particle in a box whose levels increase as n^2. Any potential which has less curvature than the harmonic oscillator potential will have more closely spaced levels. As we will see, the hydrogen atom potential has negative curvature and hence has energy levels that get closer together as you go up in energy.

2) The second point to notice is that even when the quantum number n=0, there is still energy in the amount of $\frac{1}{2}h\nu$.

This is called *zero point energy*. Its existence is highly *NON-CLASSICAL*. It implies that if you were to cool a system down to absolute zero and every molecule were in its lowest energy state, there would still be some energy in the oscillator.

If there were not zero point energy, the system would violate the Heisenberg Uncertainty Principle. If the energy were identically zero then the kinetic energy would be zero and hence the momentum would be zero, $p_x = 0$. The particle would be at the bottom of the well with no motion, so *x* = 0. This violates the Heisenberg uncertainty principle, since you would know both p_x and *x* precisely.

Another way to look at it is that the total energy of the oscillator can be written:

$$E = \frac{p_x^2}{2m} + \frac{1}{2}kx^2$$

To have zero energy requires both p_x and *x* to be zero (or their expectation values to be zero). This would violate the Heisenberg Uncertainty Principle.

We have looked at the eigenvalues of the Harmonic Oscillator Hamiltonian (*i.e*. the energies of the stationary states). Now let us look at the eigenfunctions. The eigenfunctions corresponding to the eigenvalues E_n are non-degenerate and are given by

$$\psi_n(x) = N_n H_n(\alpha^{\frac{1}{2}}x)e^{-\frac{1}{2}\alpha x^2}$$

where

$$\alpha = \sqrt{\frac{k\mu}{\hbar^2}}$$

The normalization constant N_n is

$$N_n = \frac{1}{\sqrt{2^n n!}}\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}}$$

and the $H_n(\alpha^{\frac{1}{2}}x)$ are polynomials called ***Hermite polynomials***

These polynomials are defined by the following equation or generating function:

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

where $\xi = \alpha^{\frac{1}{2}}x$

You can verify for yourselves that the first few Hermite polynomials are

$$H_0(\xi) = 1$$

$$H_1(\xi) = 2\xi$$

$$H_2(\xi) = 4\xi^2 - 2$$

$$H_3(\xi) = 8\xi^3 - 12\xi$$

$$H_4(\xi) = 16\xi^4 - 48\xi^2 + 12$$

Although we have not solved the Schrödinger equation for the Harmonic Oscillator yet, we can show that they are solutions to the differential equation.

Recall that the Schrödinger equation is:

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

Let us show that ψ_0 satisfies this equation:

$$\psi_0(x) = N_0 H_0(\alpha^{\frac{1}{2}}x)e^{-\frac{1}{2}\alpha x^2} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

We need to differentiate it an plug it back into the differential equation.

$$\frac{d\psi_0(x)}{dx} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}}\left(-\alpha x e^{-\frac{1}{2}\alpha x^2}\right)$$

$$\frac{d^2\psi_0(x)}{dx^2} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}}\left(\alpha^2 x^2 e^{-\frac{1}{2}\alpha x^2} - \alpha e^{-\frac{1}{2}\alpha x^2}\right)$$

Substituting ψ_0 and its second derivative back into the differential equation we get

$$-\frac{\hbar^2}{2\mu}\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}}\left(\alpha^2 x^2 e^{-\frac{1}{2}\alpha x^2} - \alpha e^{-\frac{1}{2}\alpha x^2}\right) + \frac{1}{2}kx^2\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} = E_0\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

Realizing that

$$\alpha = \sqrt{\frac{k\mu}{\hbar^2}}$$

this gives:

$$-\frac{\hbar^2}{2\mu}\frac{k\mu}{\hbar^2}x^2\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} + \frac{\hbar^2}{2\mu}\sqrt{\frac{k\mu}{\hbar^2}}\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} + \frac{1}{2}kx^2\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} = E_0\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

Rearing gives:

$$\frac{\hbar}{2}\sqrt{\frac{k}{\mu}}\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} = E_0\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

Taking the definition of

$$\omega = \sqrt{\frac{k}{\mu}}$$

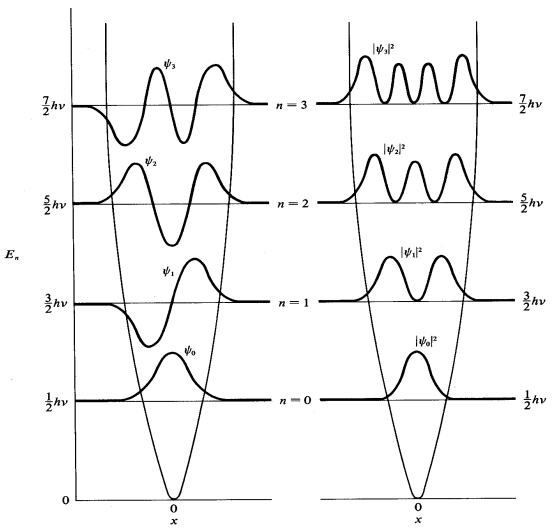
one finds:

$$\frac{1}{2}\hbar\omega\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2} = E_0\left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

Hence we have shown that $\psi_0(x)$ is an eigenfunction of the Hamiltonian and that the corresponding eigenvalue is:

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

The solutions to the Schrödinger equation give the following picture for a vibrating diatomic molecule. Note that *x* is the displacement from equilibrium.



There are several important things to notice about these eigenfunctions and related probability distributions:

1. They are qualitatively similar to the particle in the box.

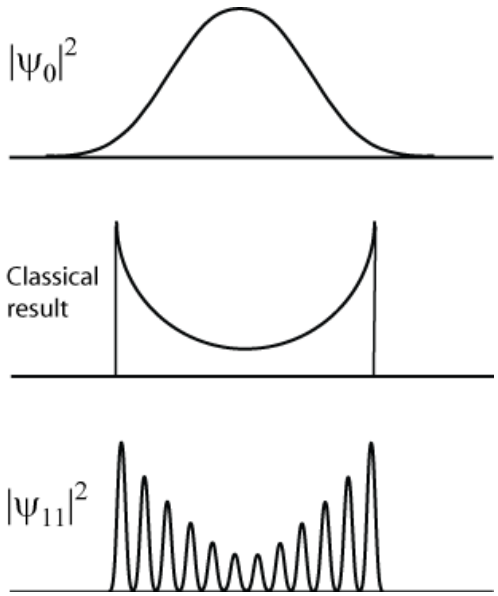
   - Simple oscillatory functions. The number of nodes increases with energy.
   - Note that it will always be true that the lowest energy wavefunction will have no nodes, the next wavefunction will have one, etc.
   - Also note that here the lowest wavefunction is ψ_0 and not ψ_1 like in the particle in a box. This results from the different boundary conditions.
   - In the particle-in-a-box wavefunctions there are *n*-1 nodes. Here there are *n* nodes.

2. The even numbered wave functions are even functions about *x*=0 and the odd numbered functions are odd about *x*=0.

   even function: $f(-x) = f(x)$
   odd function: $f(-x) = -f(x)$

3. There exists a finite probability for the quantum mechanical oscillator to exist outside the classical boundary. Recall that the particle-in-a-box wave functions went identically to zero at the boundary, but this was because the potential went to infinity. This penetration into the classically forbidden region is called tunneling and we will elaborate on it in just a moment.

   The probability distribution for ψ_0 of the harmonic oscillator is in stark contrast to the classical result. Classically the oscillator spends most of its time at the turning points. However, as *n* gets higher, the probability distribution begins to build up at the turning point and approaches the classical result. This is an example of the *Bohr Correspondence Principle*, that in the limit of high quantum number, the quantum mechanical result approaches the classical result.



I would like to digress briefly and discuss **even and odd functions.**

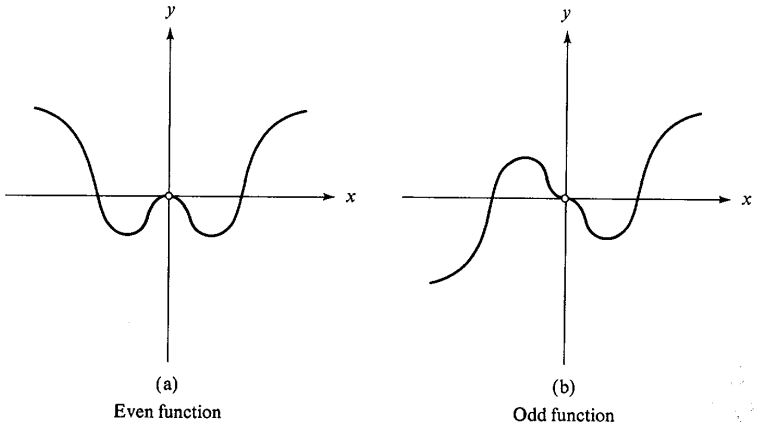
An even function is one in which $f(-x) = f(x)$ and therefore for an even function:

$$\int_{-\infty}^{\infty} f(x)dx = 2\int_{0}^{\infty} f(x)dx$$

An odd function is one in which $f(-x) = -f(x)$ and therefore for an odd function:

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

Since the positive contribution to the area at $x>0$ exactly cancels the negative contribution at $x<0$, see the figure below.



(a) Even function (b) Odd function

Examples of even and odd functions:

$f(x) = \cos(x)$	Even
$f(x) = \sin(x)$	Odd
$f(x) = x$	Odd
$f(x) = e^{-x^2}$	Even
$f(x) = e^{-x}$	Neither

Not all functions are even or odd. Many have no symmetry about $x=0$.

Note:

(even) (even)	$\rightarrow$	even function
(odd) (odd)	$\rightarrow$	even function
(even) (odd)	$\rightarrow$	odd function

Example:

$x \sin(x)$ (odd) (odd)	$\rightarrow$	even function

proof: $f(x) = x\sin(x)$

$$f(-x) = (-x)\sin(-x) = -x(-\sin(x)) = x\sin(x) = f(x)$$

The Hermite polynomials are even and odd functions, and this greatly simplifies doing integrations.

Note that the exponential part of the wavefunction, $e^{-\frac{1}{2}\alpha x^2}$ is even, so the harmonic oscillator wavefunctions reflect the evenness or oddness of the Hermite polynomial. As we noted before, the Hermite polynomials are either even or odd functions.

$H_0(\xi) = 1$ Even

$H_1(\xi) = 2\xi$ Odd

$H_2(\xi) = 4\xi^2 - 2$ Even

$H_3(\xi) = 8\xi^3 - 12\xi$ Odd

In general: n = even $\rightarrow$ H_n is even

n = odd $\rightarrow$ H_n is odd

This property will help us greatly in evaluating integrals. For example, integrals of the type

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi_n^*(x) x \psi_n(x) dx = 0$$

We do not need to explicitly do the integral in this case because if ψ_n is even or odd, the integrand will be odd and the integral equals zero.

The even/odd behavior of the harmonic oscillator wavefunctions can help us to verify that they are orthogonal to one another.

Example: Prove that ψ_0 and ψ_1 are orthogonal to one another.

$$\psi_0(x) = N_0 e^{-\frac{1}{2}\alpha x^2} \qquad \psi_1(x) = N_1 (2\alpha^{\frac{1}{2}} x) e^{-\frac{1}{2}\alpha x^2}$$

$$\int_{-\infty}^{\infty} \psi_0(x)\psi_1(x) dx = N_0 N_1 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\alpha x^2} (2\alpha^{\frac{1}{2}} x) e^{-\frac{1}{2}\alpha x^2} dx$$

$$= 2N_0 N_1 \alpha^{\frac{1}{2}} \int_{-\infty}^{\infty} x e^{-\alpha x^2} dx = 0$$

since the integrand is odd.

Note: this general rule cannot be used to evaluate integrals of the type

$$\int_{-\infty}^{\infty} \psi_{n1}(x)\psi_{n2}(x) dx = N_{n1} N_{n2} \int_{-\infty}^{\infty} H_{n1}(x) H_{n2}(x) e^{-\alpha x^2} dx$$

where H_{n1} and H_{n2} are both even. One must calculate the integral in this case.

3.3 Tunneling Effects

I would now like to make a brief digression into an aspect of the harmonic oscillator wave functions that I mentioned earlier. Recall that when we looked at the Harmonic oscillator wavefunctions we observed that there was a finite amplitude outside the classical potential, that is in a classically forbidden region. This corresponds to the particle having a finite probability of being in a classically forbidden region. This phenomenon is called tunneling.

It arises from the fact that a quantum mechanical particle is a wave and does not respond like a particle to a potential barrier. I would like to make a brief digression to consider tunneling a little further.

Recall the Schrödinger equation for the particle-in-a-box:

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

$$\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + E\psi(x) = 0$$

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0$$

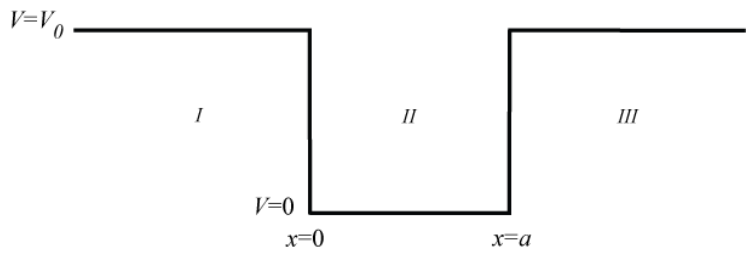
The solution to this is a function which when differentiated twice must be multiplied by a negative constant to fulfill this equation. Recall that the general solutions are complex exponentials, *i.e.*

$$\psi(x) = e^{\pm i\alpha x} \qquad \text{where} \qquad \alpha = \sqrt{\frac{2mE}{\hbar^2}} \qquad \text{(note that } \alpha \text{ is real)}$$

We then converted the complex exponentials to sines and cosines using Euler's formula.

Then we applied the boundary conditions that $\psi(0) = 0$ and $\psi(a) = 0$ since the potential was infinite in the region outside the box. Remember it was in applying the boundary conditions that the energy quantization was required.

Consider now the following potential:



In any one-dimensional potential like this where the potential abruptly changes from one region to another, one simply solves the Schrödinger equation separately in each region since the potential (and hence the Hamiltonian) is different in each region. One then has to require that the wavefunction from each region matches up smoothly at the boundaries and that there are no kinks in the wave function (*i.e.* the slopes match up as well)

Let us look at the wave function in region I when the energy is less than V_0.

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

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m(E - V_0)}{\hbar^2}\psi(x) = 0$$

Since $E<V_0$, the second term will always be negative. The solution to this equation is therefore

$$\psi(x) = e^{\pm kx} \qquad \text{where} \qquad k = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \qquad (k \text{ will always be real})$$

Recall that before we got complex exponentials which are sines and cosines. Now we get real exponentials. Let us match up the wave functions inside the box and outside in the case where the walls are finite.

We will specify the region number as a subscript of the *k* in the exponent, realizing that in a region where $E<V_0$

$$k_I = k_{III} = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}}$$

whereas in region II where $V = 0$ we have $\psi(x) = e^{\pm ik_{II}x}$

$$k_{II} = \sqrt{\frac{2mE}{\hbar^2}}$$

In all cases, *k* is real.

So in regions I - III the solutions are

$$\psi_I(x) = c_1 e^{-k_I x} + c_2 e^{+k_I x}$$

$$\psi_{II}(x) = c_3 e^{-ik_{II}x} + c_4 e^{+ik_{II}x}$$

$$\psi_{III}(x) = c_5 e^{-k_{III}x} + c_6 e^{+k_{III}x}$$

Consider region I for a moment. As $x \to -\infty$, for this function to remain finite, c_1=0.

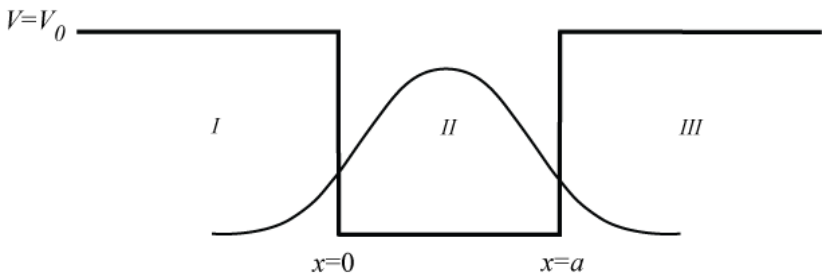
Thus $\psi_I(x) = c_2 e^{k_I x}$

Similarly in region III, to keep the wavefunction finite as $x \to \infty$, $c_6 = 0$.

Thus,

$$\psi_{III}(x) = c_5 e^{-k_{III}x}$$

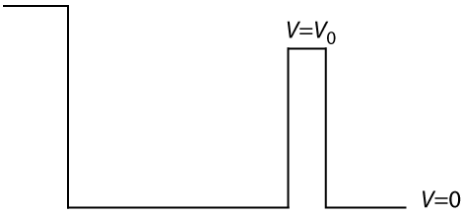
Let us then qualitatively look at the total wavefunction and see how it matches up at the boundaries.



Classically, a particle in such a well would simply undergo elastic collisions with the wall. If the energy is less that the barrier V_0, there is zero probability of finding the particle outside the well.

Quantum mechanically, however, one way to view the situation is that the particle penetrates into the wall somewhat (although it is not precise to talk about trajectories of quantum mechanical particles.)

Consider now the following potential.



Consider a particle with energy less than V_0. In the right hand region, the solutions are sines and cosines. In the well, they are sines and cosines. In the classically forbidden regions, the wavefunction will be real exponentials.

A wave function for a particle initially trapped in this potential well will decay exponentially into the barrier but still be finite when it comes out the other side. So although the particle does not have enough energy to get over the barrier, it can *tunnel* through the barrier and the probability can leak out.

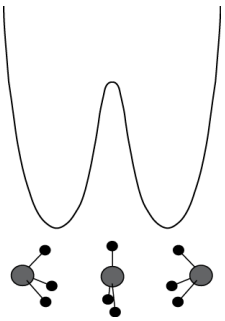
To make this a bit closer to our experience, if tunneling were to occur on a macroscopic scale, one would see the following. If you roll a ball up a hill but do not give it enough energy to reach the crest of the hill and go down the other side, tunneling would correspond to the ball suddenly disappearing from the side you rolled it up and reappearing on the other side of the crest, even though it did not have enough energy to get over the crest.

The probability of tunneling depends on the mass of the particle and the height and shape of the barrier, thus one does not observe it for macroscopic objects. (Although the probability is finite, for all practical purposes it is zero)

It turns out that tunneling phenomena are prevalent in many areas of physics, chemistry and biology. In biological molecules, since electrons are so light, electron transfer among proteins or photosynthetically related compounds involves electron tunneling.

A relatively recent experimental technique is called a tunneling electron microscope. It relies upon electron tunneling to map out the wave functions of atoms on a surface.

A famous example in chemistry is that of the ammonia molecule. Ammonia can undergo an inversion process illustrated in the figure below. At energies below the barrier, it can transform from one side to the other by tunneling.



Tunneling is strictly a consequence of the wave nature of the particle.

We assumed that the potential of a diatomic molecule can be approximated by an harmonic oscillator. This implies that the deviation from the equilibrium distance has to be small. In contrast to the classical case, the quantum mechanical Harmonic oscillator does not have a well-defined vibrational amplitude because of tunneling (the wave function decays exponentially) making it difficult to validate the approximation.

However, one can use the variance or better the standard deviation as a measure. Remember that $\langle x \rangle = 0$ due to symmetry, so we only have to calculate $\langle x^2 \rangle$. We therefore have:

$$x_{rms} = \sqrt{\langle x^2 \rangle}$$

Let us calculate $\langle x^2 \rangle$ for ψ_0

We have $$\psi_0(x) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\alpha x^2}$$

in which case

$$\langle x^2 \rangle = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} x^2 e^{-\alpha x^2} dx$$

From integral tables:

$$\int_0^{\infty} x^2 e^{-\alpha x^2} dx = \frac{1}{4\alpha}\left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}}$$

Thus

$$\langle x^2 \rangle = 2\left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \int_0^{\infty} x^2 e^{-\alpha x^2} dx = \frac{1}{2\alpha}$$

Consequently,

$$x_{rms} = \sqrt{\langle x^2 \rangle} = \frac{1}{\sqrt{2\alpha}}$$

It turns out that this root mean square displacement is small compared to the equilibrium bond length. It is typically on the order of 5% of the bond length.

3.4 Molecular Absorption of Infrared Radiation

For a harmonic oscillator potential I had stated earlier that

$$E_n = \hbar\sqrt{\frac{k}{\mu}}\left(n + \frac{1}{2}\right) = h\nu\left(n + \frac{1}{2}\right) \qquad n = 0,1,2,3,\ldots$$

In absorbing a photon to make a transition from one harmonic oscillator energy level to another, one can show that there is a selection rule in that $\Delta n = \pm 1$.

Consequently, the transition will occur at a photon energy given by

$$\Delta E = E_{n+1} - E_n = \hbar\sqrt{\frac{k}{\mu}} = \hbar\omega = h\nu \qquad \text{for all } n.$$

Please note that infrared absorption occurs when the light is "on resonance" with the classical oscillator frequency.

Because the levels are equally spaced, transitions between any pair of levels will coincide. In general at room temperature one observes only one vibrational transition, Δn=+1 originating in n=0 because most molecules are in n=0.

Thus if we make a measurement of the absorption frequency we can find k and hence the potential curve since we know how to calculate μ from the masses of the atoms. In this way we are able to learn something about the binding between the two atoms in the molecule.

3.5 Solution to the Schrödinger Equation for the Harmonic Oscillator

Recall that the potential energy of the harmonic oscillator is

$$U(x)=\frac{1}{2}kx^2$$

The Schrödinger Equation is then

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

Let us multiply through by $-\frac{2\mu}{\hbar^2}$ and rearrange to get

$$\frac{d^2\psi(x)}{dx^2}+\left(\frac{2\mu E}{\hbar^2}-\frac{\mu k}{\hbar^2}x^2\right)\psi(x)=0$$

Let us now make the following substitutions:

$$\lambda=\frac{2\mu E}{\hbar^2} \qquad \text{and} \qquad \alpha^2=\frac{k\mu}{\hbar^2}$$

We are left with

$$\frac{d^2\psi(x)}{dx^2}+\left(\lambda-\alpha^2x^2\right)\psi(x)=0$$

Let's make one more substitution:

$$\xi=\alpha^{\frac{1}{2}}x$$

And hence

$$\frac{d}{d\xi}=\frac{1}{\sqrt{\alpha}}\frac{d}{dx} \qquad \text{or} \qquad \frac{d}{dx}=\sqrt{\alpha}\frac{d}{d\xi}$$

And thus

$$\frac{d^2}{dx^2}=\alpha\frac{d^2}{d\xi^2}$$

We now have

$$\frac{d^2\psi(\xi)}{d\xi^2}+\left(\frac{\lambda}{\alpha}-\xi^2\right)\psi(\xi)=0 \qquad \text{Eqn. I}$$

This equation still does not have constant coefficients. One way to simplify this equation somewhat is to consider the behavior of the differential equation as $\xi \to \infty$. This will give us the restrictions on ψ when ξ is large.

For large ξ, $\xi^2 >> \lambda/\alpha$ thus

$$\frac{d^2\psi(\xi)}{d\xi^2} - \xi^2\psi(\xi) = 0 \qquad \text{Eqn. II}$$

We want our solutions $\psi(\xi)$ to approach the solutions of this equation at large values of ξ. The solution to Eqn. II is not simple either, since the coefficient of ψ is not constant. If we were to try a solution to Eqn. II of the form $\psi(\xi) = e^{\beta\xi^2}$ one can show that in the limit of large ξ, $\beta = \pm 1/2$ gives a function that has the correct limiting behavior.

We can show this by substituting $\psi(\xi) = e^{\beta\xi^2}$ into Eqn. II to get

$$2\beta e^{\beta\xi^2} + 4\beta^2\xi^2 e^{\beta\xi^2} - \xi^2 e^{\beta\xi^2} = 0$$

or

$$\left[\xi^2\left(4\beta^2 - 1\right) + 2\beta\right]e^{\beta\xi^2} = 0$$

Because ξ is large we can ignore the second term in the brackets and see that $\beta = \pm 1/2$ does give us a solution to Eqn. II. It is therefore a valid solution to Eqn. 1 *in the limit of large* ξ You could easily verify this.

However, note that for β = +1/2, $\psi(\xi)$ blows up, and this is unacceptable for a wave function. Thus, we will look for solutions to the original equation (Eqn. 1), with limiting behavior $\psi(\xi) \to e^{-\frac{1}{2}\xi^2}$ as $\xi \to \infty$. This is a type of boundary condition in a sense, and as we will soon see, application of the boundary condition leads to quantization of the energy.

We will therefore try a solution to the Schrödinger Equation (Eqn. I) of the form

$$\psi(\xi) = H(\xi)e^{-\frac{1}{2}\xi^2}$$

where $H(\xi)$ is a function to be determined. We must impose the requirement on $H(\xi)$ that it falls off quickly enough at large ξ so that the exponential term dominates at large ξ and thus gives ψ the right limiting behavior. We must take the second derivative of $\psi(\xi)$ substitute it back into the differential equation and get an equation for the functions $H(\xi)$.

For the first derivative we have

$$\frac{d\psi(\xi)}{d\xi} = -\xi H(\xi)e^{-\frac{1}{2}\xi^2} + \frac{dH(\xi)}{d\xi}e^{-\frac{1}{2}\xi^2}$$

For the second:

$$\frac{d^2\psi(\xi)}{d\xi^2} = -H(\xi)e^{-\frac{1}{2}\xi^2} - \xi\frac{dH(\xi)}{d\xi}e^{-\frac{1}{2}\xi^2} + \xi^2 H(\xi)e^{-\frac{1}{2}\xi^2} + \frac{d^2H(\xi)}{d\xi^2}e^{-\frac{1}{2}\xi^2} - \xi\frac{dH(\xi)}{d\xi}e^{-\frac{1}{2}\xi^2}$$

$$= \left(\frac{d^2H(\xi)}{d\xi^2} - 2\xi\frac{dH(\xi)}{d\xi} + \left[\xi^2 - 1\right]H(\xi)\right)e^{-\frac{1}{2}\xi^2}$$

Putting this into our differential equation

$$\frac{d^2\psi(\xi)}{d\xi^2}+\left(\frac{\lambda}{\alpha}-\xi^2\right)\psi(\xi)=0$$

gives

$$\left(\frac{d^2H(\xi)}{d\xi^2}-2\xi\frac{dH(\xi)}{d\xi}+\left[\xi^2-1\right]H(\xi)\right)e^{-\frac{1}{2}\xi^2}+\left(\frac{\lambda}{\alpha}-\xi^2\right)H(\xi)e^{-\frac{1}{2}\xi^2}=0$$

$$\left(\frac{d^2H(\xi)}{d\xi^2}-2\xi\frac{dH(\xi)}{d\xi}+\left[\frac{\lambda}{\alpha}-1\right]H(\xi)\right)e^{-\frac{1}{2}\xi^2}=0$$

The exponential term does not equal zero, thus

$$\frac{d^2H(\xi)}{d\xi^2}-2\xi\frac{dH(\xi)}{d\xi}+\left[\frac{\lambda}{\alpha}-1\right]H(\xi)=0$$

Solving this differential equation is then equivalent to solving the original equation since we can find $\psi(x)$ from $H(\xi)$. There are no approximations here.

This equation still does not have constant coefficients; however this differential equation has been studied in detail long before the development of quantum mechanics and is known as Hermite's differential equation. It arises in problems of scattering light of various types of geometrical surfaces.

The standard technique for solving this equation is the so-called ***series method*** or ***power series solution***.

In the series method, we assume $H(\xi)$ can be written as a power series

$$H(\xi)=a_0+a_1\xi+a_2\xi^2+a_3\xi^3+a_4\xi^4+\ldots$$

$$\frac{dH(\xi)}{d\xi}=a_1+2a_2\xi+3a_3\xi^2+4a_4\xi^3+\ldots$$

$$\frac{d^2H(\xi)}{d\xi^2}=1\cdot 2a_2+2\cdot 3a_3\xi+3\cdot 4a_4\xi^2+\ldots$$

We now substitute the series for $H(\xi)$ and its derivatives back into the differential equation and collect terms with like powers of ξ:

$$\left[1\cdot 2a_2+\left(\frac{\lambda}{\alpha}-1\right)a_0\right]+\left[2\cdot 3a_3+\left(\frac{\lambda}{\alpha}-1-2\cdot 1\right)a_1\right]\xi+\left[3\cdot 4a_4+\left(\frac{\lambda}{\alpha}-1-2\cdot 2\right)a_2\right]\xi^2+\left[4\cdot 5a_5+\left(\frac{\lambda}{\alpha}-1-2\cdot 3\right)a_3\right]\xi^3+\ldots=0$$

For this series to vanish for all values of ξ, each of the coefficients must vanish separately

$$1\cdot 2a_2+\left(\frac{\lambda}{\alpha}-1\right)a_0=0$$

$$2\cdot 3a_3+\left(\frac{\lambda}{\alpha}-1-2\cdot 1\right)a_1=0$$

$$3\cdot 4a_4+\left(\frac{\lambda}{\alpha}-1-2\cdot 2\right)a_2=0$$

$$4\cdot 5a_5+\left(\frac{\lambda}{\alpha}-1-2\cdot 3\right)a_3=0$$

or in general

$$(n+2)(n+1)a_{n+2}+\left(\frac{\lambda}{\alpha}-1-2n\right)a_n=0 \qquad n=0,1,2,3,\ldots$$

We can rearrange this to get

$$a_{n+2}=-\frac{\left(\frac{\lambda}{\alpha}-1-2n\right)}{(n+1)(n+2)}a_n$$

This type of formula is called a ***recursion*** formula. It gives one coefficient in the series expansion of $H(\xi)$ in terms of a previous coefficient. This is a two-term recursion formula in that it skips by 2.

Thus, if one chooses two arbitrary constants, a_0 and a_1 then the recursion formula gives two independent sets of coefficients

$$H_{even}=a_0+a_2\xi^2+a_4\xi^4+\ldots$$

$$H_{odd}=a_1\xi+a_3\xi^3+a_5\xi^5+\ldots$$

Recall that the Harmonic oscillator solutions are either odd functions or even functions. This is where this property arises in this system.

We can use the recursion relation to solve for all the even a's in terms of a_0 and all the odd a's in terms of a_1. Thus there are two unknown constants (as you would expect for a second order differential equation). The first few of these relations are (for the even terms):

$$a_2=-\frac{1}{1\cdot 2}\left(\frac{\lambda}{\alpha}-1\right)a_0$$

$$a_4=-\frac{1}{3\cdot 4}\left(\frac{\lambda}{\alpha}-1-2\cdot 2\right)a_2=\frac{1}{4!}\left(\frac{\lambda}{\alpha}-1\right)\left(\frac{\lambda}{\alpha}-1-2\cdot 2\right)a_0$$

and for the odd terms:

$$a_3=-\frac{1}{2\cdot 3}\left(\frac{\lambda}{\alpha}-1-2\cdot 1\right)a_1$$

$$a_5=-\frac{1}{4\cdot 5}\left(\frac{\lambda}{\alpha}-1-2\cdot 3\right)a_3=\frac{1}{5!}\left(\frac{\lambda}{\alpha}-1-2\cdot 1\right)\left(\frac{\lambda}{\alpha}-1-2\cdot 3\right)a_1$$

Now, as I mentioned before, we need to require $H(\xi)$ to converge at large ξ sufficiently rapidly that the exponential part of the wave function, $e^{-\frac{1}{2}\xi^2}$, will dominate. We must make the power series fit the boundary condition $\psi(\xi)\to 0$ as $\xi\to\infty$.

To do this, we need to look at the relative size of successive coefficients, a_n at large n.

From the recursion relation:

$$\lim_{n\to\infty}\frac{a_{n+2}}{a_n}=\frac{2n}{n^2}=\frac{2}{n}$$

For large n, $a_{n+2} << a_n$. This better be true if the series is to converge. But, does this converge fast enough to satisfy the boundary conditions?

To determine this, let us look at a function that we know blows up at large ξ. We can compare the rate of convergence with our power series.

$$e^{\xi^2} = 1 + \xi^2 + \ldots + \frac{\xi^n}{\frac{n}{2}!} + \frac{\xi^{n+2}}{\left(\frac{n}{2}+1\right)!} + \ldots$$

$$= 1 + \xi^2 + \ldots + b_n \xi^n + b_{n+2} \xi^{n+2} + \ldots$$

For large *n* we find

$$\underset{n\to\infty}{Lim} \frac{b_{n+2}}{b_n} = \frac{\frac{n}{2}!}{\left(\frac{n}{2}+1\right)!} = \frac{1}{\left(\frac{n}{2}+1\right)} = \frac{2}{n}$$

So you can see that our power series in ξ has the same asymptotic behavior as e^{ξ^2}

If $\underset{\xi\to\infty}{Lim} H(\xi) = e^{\xi^2}$ then $\underset{\xi\to\infty}{Lim} H(\xi) e^{-\frac{1}{2}\xi^2} = e^{\frac{1}{2}\xi^2}$

This will not satisfy the boundary condition. We therefore must terminate the series after a certain number of terms so that we have a polynomial of finite length as a solution rather than a power series.

Recall that our recursion relation is

$$a_{n+2} = -\frac{\left(\frac{\lambda}{\alpha} - 1 - 2n\right)}{(n+1)(n+2)} a_n$$

You can see that if we let

$$\left(\frac{\lambda}{\alpha} - 1 - 2n\right) = 0$$

And thus

$$\frac{\lambda}{\alpha} = 2n + 1$$

This will make a_{n+2} be the first zero term and hence all the subsequent *a*'s will be zero. It breaks off the series after *n* terms. Thus, we will get an entire set of polynomials that have an increasing number of terms as the integer *n* increases. This will insure that the function $H(\xi)$ converges rapidly enough to satisfy the boundary conditions.

Remember that λ is related to the energy. Thus, to make the solution satisfy the boundary conditions, we have quantized the energy levels. We have seen this same principle in solving the Schrödinger equation for other simple 1-dimensional potentials. Here, the quantization comes in making the series converge.

Note that if *n* is even, the odd terms will not go to zero. However, since we have only even or odd solutions, this is ok.

Let us now look at the energy levels:

$$\frac{\lambda}{\alpha}=\frac{\frac{2\mu E}{\hbar^2}}{\sqrt{\frac{\mu k}{\hbar^2}}}=2n+1$$

This yields

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

$$=h\nu_{classical}\left(n+\frac{1}{2}\right)$$

where $\nu_{classical}=\frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$

The wavefunctions can be given by:

$$\psi_n(\xi)=N_nH_n(\xi)e^{-\frac{1}{2}\xi^2}$$

In making the function $H(\xi)$ have the right behavior at large ξ we got quantization of the energy.

We forced the power series for $H(\xi)$ to be finite by causing one of the coefficients to go to zero at a particular *n*. All the higher terms will also go to zero. We therefore obtained a set of solutions $H(\xi)$ that terminate after different number of terms. You can see that this is what gives us different number of nodes in our wavefunctions as we go to high quantum numbers.

If you look back at the expressions we obtained when we first set the coefficients of each power of ξ equals zero, and let (λ/α) - 1 = 2*n*, we get

$$a_2=-\frac{1}{1\cdot 2}\left(\frac{\lambda}{\alpha}-1\right)a_0=\frac{2n}{1\cdot 2}a_0$$
$$a_4=\frac{1}{4!}\left(\frac{\lambda}{\alpha}-1\right)\left(\frac{\lambda}{\alpha}-1-2\cdot 2\right)a_0=\frac{2n(2n-4)}{4!}a_0$$

One can show that in general

$$H_n(\xi)=a_0\left[1-\frac{2n}{2!}\xi^2+\frac{2^2n(n-2)}{4!}\xi^4-\frac{2^3n(n-2)(n-4)}{6!}\xi^6+\ldots\right]\qquad n=\text{even}$$

$$H_n(\xi)=a_1\left[\xi-\frac{2(n-1)}{3!}\xi^3+\frac{2^2(n-1)(n-3)}{5!}\xi^5-\ldots\right]\qquad n=\text{odd}$$

You can see how each of these truncates after a finite number of terms, which depends on the quantum number *n*.

The values of a_0 and a_1 are arbitrary (since the overall wave function has a normalization constant). By convention, a_0 and a_1 are chosen so that the coefficient of the highest power of ξ, that is ξ^n is 2^n.

This gives

$$a_0 = (-1)^{\frac{n}{2}} \frac{n!}{\frac{n}{2}!}$$

and

$$a_1 = (-1)^{\frac{n-1}{2}} \frac{2n!}{\frac{n-1}{2}!}$$

Remember that I had given you another form for the Hermite polynomials a few lectures ago

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

This is more compact and easier to work with.

It is important not to forget that the full solution to the Schrödinger equation is

$$\psi_n(\xi) = N_n H_n(\xi) e^{-\frac{1}{2}\xi^2} \qquad \text{where} \qquad \xi = \alpha^{\frac{1}{2}} x$$

3.6 Raising and Lowering Operators

If we want to calculate certain properties of the system, we will often encounter integrals of the type:

$$\int_{-\infty}^{\infty} \psi_n^*(x) \hat{A} \psi_m(x) dx$$

We already saw that if we want to determine the average value of the position we have to evaluate:

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi_n^*(x) x \psi_n(x) dx$$

Based upon the symmetry of the harmonic oscillator eigenfunctions it could be shown that this integral equals zero. However if we are interested in the mean deviation or variance of the position we have to evaluate

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} \psi_n^*(x) x^2 \psi_n(x) dx$$

We solved this integral for the ground state eigenfunction. Although not difficult, it is a non-trivial calculation. This is even more true if one were to evaluate this integral using the eigenfunction of higher vibrational levels.

When discussing the absorption spectrum I mentioned that there exists a selection rule $\Delta n = \pm 1$. As you will see in the spectroscopy course of next semester this selection rule follows from the evaluation of the following integral:

$$\int_{-\infty}^{\infty} \psi_n^*(x) x \psi_m(x) dx$$

Although one can calculate all these integrals explicitly, it would be worthwhile to have a way to evaluate them more easily. This can be done by making use of the so-called lowering and raising operators.

Let us start with the Hamiltonian for the harmonic oscillator:

$$\hat{H} = \frac{\hat{p}_x^2}{2\mu} + \frac{1}{2}k\hat{x}^2$$

which can also be written as:

$$\hat{H} = \frac{\hat{p}_x^2}{2\mu} + \frac{1}{2}\mu\omega^2\hat{x}^2$$

with $\omega = \sqrt{\frac{k}{\mu}}$

We see that the Hamiltonian contains only the operators $\hat{p}_x^2$ and $\hat{x}^2$. Since the operators $\hat{p}_x$ and $\hat{x}$ are Hermitian and thus have real eigenvalues it follows that the eigenvalues of the Hamiltonian are non-negative.

Let us at this point introduce the new operators:

$$\hat{a} = \sqrt{\frac{\mu\omega}{2}}\hat{x} + i\frac{\hat{p}_x}{\sqrt{2\mu\omega}}$$

and

$$\hat{a}^\dagger = \sqrt{\frac{\mu\omega}{2}}\hat{x} - i\frac{\hat{p}_x}{\sqrt{2\mu\omega}}:$$

Realizing that these new operators do not commute, *i.e.*

$$\left[\hat{a},\hat{a}^\dagger\right] = \hbar$$

we can write the Hamiltonian as:

$$\hat{H} = \omega\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\hbar\omega$$

One can show that the following commutation relations hold between the newly defined operators and the Hamiltonian:

$$\left[\hat{H},\hat{a}\right] = -h\omega\hat{a}$$

$$\left[\hat{H},\hat{a}^\dagger\right] = h\omega\hat{a}^\dagger$$

Let us now write down the eigenvalue equation, which reads:

$$\hat{H}u_E = Eu_E$$

In the past, whenever we wrote down such an equation the implication was that $\hat{H}$ contained some operators and that the eigenfunction u_E was a function of *x*. In general, this does not have to be the case as we will see later in the course when we will speak about spin.

If we have the commutator $\left[\hat{H},\hat{a}\right]$ acting on this eigenfunction, u_E we find:

$$\left[\hat{H},\hat{a}\right]u_E = -\hbar\omega\hat{a}u_E$$

$$\hat{H}\hat{a}u_E - \hat{a}\hat{H}u_E = -\hbar\omega\hat{a}u_E$$

With the help of the eigenvalue equation, we can rewrite this as:

$$\hat{H}\hat{a}u_E = \left(E - \hbar\omega\right)\hat{a}u_E$$

This equation states that if u_E is an eigenfunction of the Hamiltonian with eigenvalue E then $\hat{a}u_E$ is also an eigenfunction of $\hat{H}$ but with eigenvalue $E - \hbar\omega$, that is , with energy lowered by one unit of

$$\varepsilon = \hbar\omega$$

This implies that when operator $\hat{a}$ is applied to u_E it generates an eigenfunction of the Hamiltonian corresponding to the energy $E - \hbar\omega$.

We can therefore write

$$\hat{a}u_E = C\left(E\right)u_{E-\varepsilon}$$

We need the constant $C\left(E\right)$ here since even if u_E is normalized, $\hat{a}u_E$ need not to be.
If we now apply the same operators to the state $u_{E-\varepsilon}$ we find, in exactly the same way that $\hat{a}u_{E-\varepsilon}$ or, equivalently, $\hat{a}^2 u_{E-\varepsilon}$ gives a state of energy $E - 2\hbar\omega$. Thus by repeated application of operator a to any eigenfunction of the Hamiltonian, u_E, we can generate states of lower and lower energy. Appropriately, $\hat{a}$ is called a *lowering operator*. There is a limit to how many times operator $\hat{a}$ can be applied since the eigenvalues of the Hamiltonian has to be non-negative, as we saw before. Thus, the lowering procedure must end at some point. Consequently, there exists a ground state, which we will denote by u_0, beyond which the lowering ends. So we have:

$$\hat{a}u_0 = 0$$

Analogous to the lowering operator $\hat{a}$, one can show that the operator $\hat{a}^\dagger$ when applied upon u_E increases the energy to $E + \hbar\omega$. One can thus write:

$$\hat{a}^\dagger u_E = C'\left(E\right)u_{E+\varepsilon}$$

The operator $\hat{a}^\dagger$ is appropriately called a *raising operator.*

We saw before that the Hamiltonian expressed in lower and raising operators takes on the form:

$$\hat{H} = \omega\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\hbar\omega$$

If we apply the Hamiltonian onto the ground state eigenfunction u_0 we find the energy of the ground state.

$$\hat{H}u_0 = E_0 u_0$$

Putting in our definition of the Hamiltonian we find:

$$\hat{H}u_0 = \left(\omega\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\hbar\omega\right)u_0$$

$$= \omega\hat{a}^\dagger\hat{a}u_0 + \tfrac{1}{2}\hbar\omega u_0$$

$$= \tfrac{1}{2}\hbar\omega u_0 = E_0 u_0$$

We will now change our notation a little, namely, we will label the state by the number of energy units $\varepsilon = \hbar\omega$ it has over the ground state energy, $E_0 = \frac{1}{2}\hbar\omega$. Thus, we can write:

$$\hat{a}u_n = Cu_{n-1}$$

$$\hat{a}^\dagger u_n = C'u_{n+1}$$

where C and C' are n-dependent constants that can be determined by forcing the wavefunctions to be orthonormal. Doing this (I will not show this here) one finds:

$$\hat{a}u_n = \sqrt{n\hbar}\,u_{n-1}$$

$$\hat{a}^\dagger u_n = \sqrt{(n+1)\hbar}\,u_{n+1}$$

One should note here that $\hat{a}$ and $\hat{a}^\dagger$ move up and down the same "ladder". Hence, these operators are sometimes also referred to as ladder operators.

We can now readily calculate the energy of a given state n. We have:

$$\hat{H}u_n = \left(\omega\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\hbar\omega\right)u_n$$

$$= \omega\hat{a}^\dagger\hat{a}u_n + \tfrac{1}{2}\hbar\omega u_n$$

$$= \omega\hat{a}^\dagger\sqrt{n\hbar}\,u_{n-1} + \tfrac{1}{2}\hbar\omega u_n$$

$$= \omega\sqrt{(n-1+1)\hbar}\sqrt{n\hbar}\,u_n + \tfrac{1}{2}\hbar\omega u_n$$

$$= \omega n\hbar u_n + \tfrac{1}{2}\hbar\omega u_n$$

$$= \hbar\omega(n+\tfrac{1}{2})u_n = E_n u_n$$

This is exactly the same result we found before when we solved the differential equation, except that now we have not solved any complicated mathematical equations. This way of solving a quantum mechanical problem is closely related to the method proposed by Heisenberg.

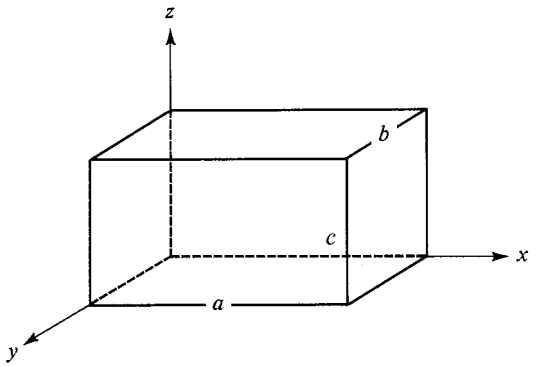
With this at hand, one can now simply solve the integral given at the beginning of this section, something you will do in the exercises.

4 Three Dimensional Systems

We will now move on to the next level of complexity in solving the time independent Schrödinger equation: systems of three dimensions. There are interesting new phenomena that arise in multidimensional systems that are due to the equivalence of the *x*, *y*, and *z* directions.

We will start with the simplest case: a particle in a three dimensional box. We will then move on to the rigid rotor, which is used as a model for the rotational motion of a diatomic molecule. After this, we will spend a little time discussing angular momentum in Quantum Mechanical systems and then treat the hydrogen atom problem.

4.1 Particle in a 3-dimensional box



Let $V = 0$ inside the box and $V = \infty$ outside the box. The Schrödinger equation inside the box is:

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

or

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z)=E\psi(x,y,z)$$

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

is called the Laplacian Operator

We will use the technique of *separation of variables*.

Whenever the Hamiltonian $\hat{H}$ can be written as a sum of parts depending upon each coordinate, *i.e.*

$$\hat{H}(x,y,z)=\hat{H}(x)+\hat{H}(y)+\hat{H}(z)$$

we can assume that

$$\psi(x,y,z)=\psi_x(x)\psi_y(y)\psi_z(z)$$

We can then write:

$$-\frac{\hbar^2}{2m}\left(\psi_y(y)\psi_z(z)\frac{\partial^2\psi_x(x)}{\partial x^2}+\psi_x(x)\psi_z(z)\frac{\partial^2\psi_y(y)}{\partial y^2}+\psi_x(x)\psi_y(y)\frac{\partial^2\psi_z(z)}{\partial z^2}\right)=E\,\psi_x(x)\psi_y(y)\psi_z(z)$$

If we divide both sides by $\psi_x(x)\psi_y(y)\psi_z(z)$ we get

$$-\frac{\hbar^2}{2m\,\psi_x(x)}\frac{\partial^2\psi_x(x)}{\partial x^2}-\frac{\hbar^2}{2m\,\psi_y(y)}\frac{\partial^2\psi_y(y)}{\partial y^2}-\frac{\hbar^2}{2m\,\psi_z(z)}\frac{\partial^2\psi_z(z)}{\partial z^2}=E$$

Each of the terms on the left hand side are independently functions of *x*, *y*, and *z* respectively. For these terms to sum to a constant for all *x*, *y*, and *z*, each term must separately equal a constant. Thus.

$$-\frac{\hbar^2}{2m\,\psi_x(x)}\frac{\partial^2\psi_x(x)}{\partial x^2}=E_x \qquad -\frac{\hbar^2}{2m\,\psi_y(y)}\frac{\partial^2\psi_y(y)}{\partial y^2}=E_y \qquad -\frac{\hbar^2}{2m\,\psi_z(z)}\frac{\partial^2\psi_z(z)}{\partial z^2}=E_z$$

where $E_x+E_y+E_z=E$

Rearranging these we see that we get 3 independent particle-in-a-box equations.

$$\frac{\partial^2\psi_x(x)}{\partial x^2}+\frac{2mE_x}{\hbar^2}\psi_x(x)=0$$
$$\frac{\partial^2\psi_y(y)}{\partial y^2}+\frac{2mE_y}{\hbar^2}\psi_y(y)=0$$
$$\frac{\partial^2\psi_z(z)}{\partial z^2}+\frac{2mE_z(z)}{\hbar^2}\psi_z=0$$

Each of these has boundary conditions that force the wavefunction to zero at either boundary since the potential outside the box is infinite

$$\psi_x(0)=\psi_x(a)=\psi_y(0)=\psi_y(b)=\psi_z(0)=\psi_z(c)=0$$

Applying these boundary conditions as we did with the particle in the box we get:

$$\psi_x(x)=A_x\sin\left(\frac{n_x\pi}{a}x\right) \qquad E_x=\frac{h^2n_x^2}{8ma^2} \qquad n_x=1,2,3,\ldots$$

$$\psi_y(y)=A_y\sin\left(\frac{n_y\pi}{b}y\right) \qquad E_y=\frac{h^2n_y^2}{8mb^2} \qquad n_y=1,2,3,\ldots$$

$$\psi_z(z)=A_z\sin\left(\frac{n_z\pi}{c}z\right) \qquad E_z=\frac{h^2n_z^2}{8mc^2} \qquad n_z=1,2,3,\ldots$$

The total wavefunctions is then

$$\psi(x,y,z)=\psi_x(x)\psi_y(y)\psi_z(z)$$
$$=A_xA_yA_z\sin\left(\frac{n_x\pi}{a}x\right)\sin\left(\frac{n_y\pi}{b}y\right)\sin\left(\frac{n_z\pi}{c}z\right)$$

The normalization constant can be found by:

$$\int_0^a\int_0^b\int_0^c\psi^*(x,y,z)\psi(x,y,z)dxdydz=1$$

One can integrate just over the box since ψ is zero outside box. The integral will yield

$$A_x A_y A_z = \sqrt{\frac{8}{abc}}$$

Another way to do this would be to realize that the particle must independently have unit probability of being found at some *x*, some *y*, and some *z*. (This will be true for any orthogonal coordinate system).

Therefore

$$\int_0^a \psi_x^*(x)\psi_x(x)dx = 1 \qquad \Rightarrow \qquad A_x = \sqrt{\frac{2}{a}}$$

and similarly for *y* and *z*.

To find average properties for a 3-dimensional particle in box (*x* for example) one would write:

$$\langle x \rangle = \int_0^a \int_0^b \int_0^c \psi_x^*(x)\psi_y^*(y)\psi_z^*(z)x\psi_x(x)\psi_y(y)\psi_z(z)dxdydz$$

One can write this as

$$\langle x \rangle = \int_0^a \psi_x^*(x)x\psi_x(x)dx \int_0^b \psi_y^*(y)\psi_y(y)dy \int_0^c \psi_z^*(z)\psi_z(z)dz$$

which is the same as the 1-dimensional problem.

Let's look at the 3-dimensional particle-in-a-box energies:

$$E_{n_x n_y n_z} = E_x + E_y + E_z = \frac{h^2}{8m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2}\right)$$

where we label the energy *E* by the three quantum numbers. The quantum numbers n_x , n_y , n_z vary independently and can be attributed to the 3-dimensional nature of the problem.

Consider the case in which our box has sides of equal length, *a*=*b*=*c*. The energy levels are then given by

$$E_{n_x n_y n_z} = \frac{h^2}{8ma^2}\left(n_x^2 + n_y^2 + n_z^2\right)$$

Notice what this brings about. You get wave functions corresponding to different quantum numbers having the same energy. For example:

$$\psi_{2,1,1}(x,y,z) = \sqrt{\frac{8}{a^3}}\sin\left(\frac{2\pi}{a}x\right)\sin\left(\frac{\pi}{a}y\right)\sin\left(\frac{\pi}{a}z\right)$$

$$\psi_{1,1,2}(x,y,z) = \sqrt{\frac{8}{a^3}}\sin\left(\frac{\pi}{a}x\right)\sin\left(\frac{\pi}{a}y\right)\sin\left(\frac{2\pi}{a}z\right)$$

These have the same energy eigenvalue because $n_x^2 + n_y^2 + n_z^2 = 6$ in both cases.

The situation where two or more wave functions have the same energy is called *degeneracy*, and the levels are called degenerate levels.

If there are *n* different eigenfunctions, which correspond to the same energy, the system is considered to be *n-fold degenerate.*

You can see that this was caused by the symmetry of the system. When we let the 3 sides of the box be the same length, the three directions became completely equivalent.

There are important implications when one has degenerate levels which are related to the following theorem:

THEOREM
If we have an *n*-fold degenerate level in which *n* independent wavefunctions $\psi_1,\psi_2,\ldots,\psi_n$ correspond to the same energy *E*, any linear combination

$$\varphi = c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n$$

of the *n* functions of the degenerate level is also an eigenfunction of the Hamiltonian with eigenvalue *E*.

Proof:
We need to show that if

$$\hat{H}\psi_1 = E\psi_1 \qquad \hat{H}\psi_2 = E\psi_2 \qquad \ldots \qquad \hat{H}\psi_n = E\psi_n$$

then $\hat{H}\varphi = E\varphi$

where $\varphi = c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n$

or $\hat{H}\left[c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\right] = E\left[c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\right]$

Since the Hamiltonian is linear we can write

$$\hat{H}\left[c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\right] = \hat{H}c_1\psi_1 + \hat{H}c_2\psi_2 + \ldots + \hat{H}c_n\psi_n$$

$$= c_1\hat{H}\psi_1 + c_2\hat{H}\psi_2 + \ldots + c_n\hat{H}\psi_n$$

But since ψ_n are eigenfunctions with same eigenvalue *E*,

$$\hat{H}\left[c_1\psi + c_2\psi_2 + \ldots + c_n\psi_n\right] = c_1E\psi_1 + c_2E\psi_2 + \ldots + c_nE\psi_n$$

Thus,

$$\hat{H}\left[c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\right] = E\left[c_1\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\right]$$

or $\hat{H}\varphi = E\varphi$

We will see degeneracy crop up in many other parts of the course, for example when we will discuss the orbitals of the hydrogen atom.

In particular, it has some effects on simultaneous measurements in quantum mechanical systems.

4.2 Separability of the Hamiltonian

The 3-dimensional particle-in-a-box has illustrated another important principle that we have already seen twice before and will see frequently in this course—the use of the technique of separation of variables to solve a partial differential equation.

Essentially it allows us to replace a partial differential equation with a set of ordinary differential equations. We did this for the time dependent Schrödinger equation, the classical mechanical harmonic oscillator, and now for the 3-dimensional particle-in-a-box.

While this is a standard technique for partial differential equations in general, it has particularly important consequences when applied to solving the Schrödinger equation.

In general, if the Hamiltonian can be written in the form

$$\hat{H}(q_1, q_2, q_3) = \hat{H}(q_1) + \hat{H}(q_2) + \hat{H}(q_3)$$

(using a 3-dimensional case as an example) where the *q*'s are any set of orthogonal coordinates, we say that the Hamiltonian is *separable*.

This is true because the Schrödinger equation is a linear, homogeneous differential equation. We can assume solutions to the Schrödinger equation in the form

$$\psi(q_1, q_2, q_3) = \psi(q_1)\psi(q_2)\psi(q_3)$$

We then get 3 ordinary differential equations involving $\psi(q_1)$, $\psi(q_2)$ and $\psi(q_3)$.

The resulting eigenvalues for the energy *E* will be the sum:

$$E = E_1 + E_2 + E_3$$

We can generalize this result to the case where we can write our Hamiltonian in terms of two independent sets of coordinates.

For example, in a many particle system, if we can write the Hamiltonian in a form which divides according to the coordinates of each particle, then we can solve separate Schrödinger equations for each particle. The wave functions will be the product of wave functions of each particle and the energy will be a sum of energies for each particle.

Another way one can divide the Hamiltonian for a system of two particles is to write the total energy in terms of the energy of the motion of the center of mass and the relative motion of the particles. We can do this whenever the potential energy depends only on the relative coordinates. (Remember we had done something like this for the solution to the classical harmonic oscillator).

If we let *X*, *Y*, and *Z* be the coordinates of the center of mass and *x*, *y*, and *z* be the internal or relative coordinates (coordinates relative to the center of mass), we can write the total energy of the system

$$E = \frac{M}{2}\left(\mathrm{V}_X^2 + \mathrm{V}_Y^2 + \mathrm{V}_Z^2\right) + \frac{\mu}{2}\left(\mathrm{v}_x^2 + \mathrm{v}_y^2 + \mathrm{v}_z^2\right) + U(x, y, z)$$

where *M* is the total mass and μ is the reduced mass (as we had defined earlier).

We can write this in terms of momenta by noting that

$$P_X = M\mathrm{V}_X \qquad \text{and} \qquad p_x = \mu \mathrm{v}_x$$

and similarly for y and z.

$$E = \frac{1}{2M}\left(P_X^2 + P_Y^2 + P_Z^2\right) + \frac{1}{2\mu}\left(p_x^2 + p_y^2 + p_z^2\right) + U(x, y, z)$$

We can then write the quantum mechanical Hamiltonian as

$$\hat{H} = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right)-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+U(x,y,z)$$

You can see this falls into the form

$$\hat{H} = \hat{H}_1 + \hat{H}_2$$

where $\hat{H}_1$ depends on the center of mass coordinates only and $\hat{H}_2$ depends on internal coordinates only. The total wavefunction is therefore

$$\psi_{Tot}(X,Y,Z,x,y,z) = \psi_{Trans}(X,Y,Z)\psi_{Int}(x,y,z)$$

and the energies are

$$E_{Tot} = E_{Trans} + E_{Int}$$

If we look at the part of the Hamiltonian that describes the motion of the center of mass

$$\hat{H}_1 = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right)$$

you can see that this just represents the translational motion of a free particle in space. Since there is no potential to impose boundary conditions, *the energy of a free particle is not quantized.*

Thus, the contribution to the total energy is just a constant and is of little interest. If you look at any transitions between energy levels, this constant will drop out since it is the same for all internal states.

We are left with the Schrödinger equation for the relative motion

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+U(x,y,z)\right]\psi_{Int}(x,y,z) = E_{Int}\psi_{Int}(x,y,z)$$

or simply

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2+U(x,y,z)\right]\psi_{Int}(x,y,z) = E_{Int}\psi_{Int}(x,y,z)$$

where ∇^2 is the Laplacian Operator.

Later we will see that this Schrödinger equation can be further separated, although we will not show this explicitly until next semester. For example, when modeling the vibrational and rotational motion of a diatomic, to first order one can separate the vibrational coordinates (relative distance) and the rotational coordinates (angular orientation). This will allow us to write the energy of a vibrating, rotating diatomic as the sum of the vibrational energy and the rotational energy.

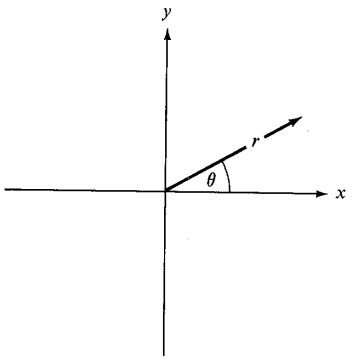
So when we draw vibrational energy levels on a harmonic potential energy curve for example, one can draw the rotational energy levels on top of this.

We will also be able to separate the motions of the electrons from those of the nuclei.

4.3 Choice of Coordinate Systems

For a multi-particle system, the Hamiltonian may be very complicated in Cartesian coordinates with lots of cross terms. However certain problems have an inherent symmetry that can simplify the problem if the coordinate system coincides with that symmetry.

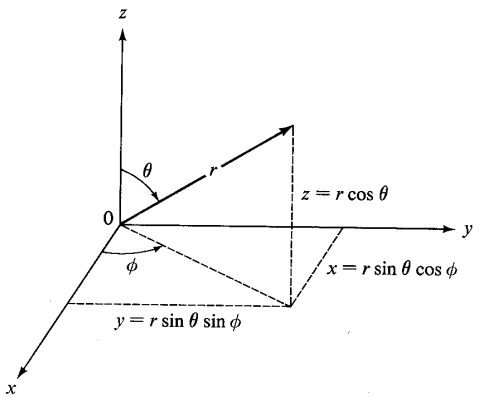
For example consider a diatomic which is constrained to rotate in the *x*-*y* plane:



You can see that upon rotation both *x* and *y* will change, however if you look at what happens in plane polar coordinates (r, θ), only θ changes. The Hamiltonian can be written as a function of θ only (which is the same as particle on a ring).

Consider a rotor in which the distance *r* is fixed (*i.e.*, a *rigid rotor*) but which is not confined to a plane.

We need to convert to spherical polar coordinates.



$$x = r\sin\theta\cos\varphi$$
$$y = r\sin\theta\sin\varphi$$
$$z = r\cos\theta$$

$$0 \le \theta \le \pi, 0 \le \varphi \le 2\pi, 0 \le r \le \infty$$

$$r^2 = x^2 + y^2 + z^2$$
$$\cos\theta = \frac{z}{r} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$
$$\tan\varphi = \frac{y}{x}$$

Here *r* is fixed, and only θ and φ change. In Cartesian coordinates you would have to consider the change in all three coordinates, *x, y*, and *z*.

So choosing the proper coordinate system can help separate (and simplify) the Hamiltonian.

It turns out that for any two- particle problem in which the potential *U* simply depends upon the relative coordinates of the two particles, (*i.e*. *U* = *U*(*r*)), one can greatly simplify the problem by working in spherical polar coordinates.

To convert the Schrödinger equation for internal or relative motion

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + U(x,y,z)\right]\psi(x,y,z) = E_{Int}\psi(x,y,z)$$

into spherical polar coordinates, one must convert the Laplacian operator ∇^2 to spherical polar coordinates. To do this you need to use the chain rule of differentiation to convert all the second partial derivatives from derivatives with respect to *x*, *y*, and *z* to those with respect to *r*, θ, φ.

For example, consider some function $f(r,\theta,\varphi)$ where *r*, θ, and φ are in turn functions of *x*, *y*, and *z*.

$$\left(\frac{\partial f}{\partial x}\right)_{y,z} = \left(\frac{\partial r}{\partial x}\right)_{y,z}\left(\frac{\partial f}{\partial r}\right)_{\vartheta,\varphi} + \left(\frac{\partial \vartheta}{\partial x}\right)_{y,z}\left(\frac{\partial f}{\partial \vartheta}\right)_{r,\varphi} + \left(\frac{\partial \varphi}{\partial x}\right)_{y,z}\left(\frac{\partial f}{\partial \varphi}\right)_{r,\vartheta}$$

and similarly for *y* and *z*.

To make this into an operator equation we need to write:

$$\left(\frac{\partial}{\partial x}\right)_{y,z} f = \left[\left(\frac{\partial r}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial r}\right)_{\vartheta,\varphi} + \left(\frac{\partial \vartheta}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial \vartheta}\right)_{r,\varphi} + \left(\frac{\partial \varphi}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial \varphi}\right)_{r,\vartheta}\right] f$$

The second derivatives come from applying this operator twice.

$$\left(\frac{\partial^2}{\partial x^2}\right)_{y,z} f = \left(\frac{\partial}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial x}\right)_{y,z} f$$

$$= \left(\frac{\partial}{\partial x}\right)_{y,z}\left[\left(\frac{\partial r}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial r}\right)_{\vartheta,\varphi} + \left(\frac{\partial \vartheta}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial \vartheta}\right)_{r,\varphi} + \left(\frac{\partial \varphi}{\partial x}\right)_{y,z}\left(\frac{\partial}{\partial \varphi}\right)_{r,\vartheta}\right] f$$

We now need to substitute in for $\frac{\partial f}{\partial x}$ from the previous equation. Then we need to calculate the partial derivatives of *r*, θ, and φ with respect to *x*, *y*, and *z*.

Then the whole process has to be repeated for the second partials with respect to *y* and *z*. I will skip over the gory details and leave them for the exercises. When we are all done, the Laplacian in spherical polar coordinates is

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

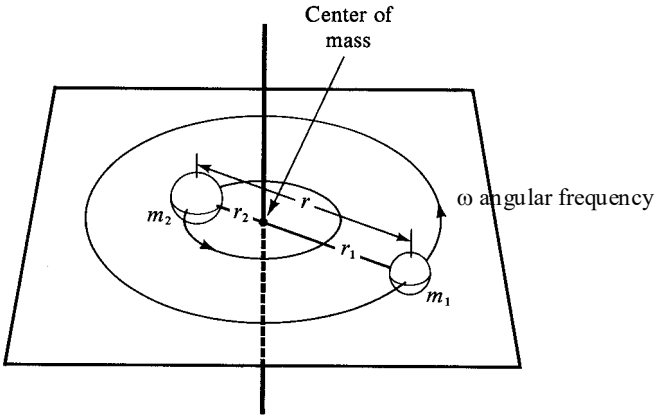
Although this looks quite complicated, it greatly simplifies the Hamiltonian of systems which have spherical symmetry, that is cases where $U = U(r)$.

We will use this form of ∇^2 in our solution of the rigid rotor and hydrogen atom.

4.4 The Rigid Rotor: A Model for a Rotating Diatomic

Consider two particles rotating - rigidly fixed at some distance *r* apart.

For diatomics, this approximation is pretty good because in the ground vibrational state, the root mean square amplitude is approximately 5% of the equilibrium bond distance. For higher vibrational levels this approximation begins to break down.



The system rotates about the center of mass: *i.e.* the center of mass looks stationary.

Center of mass condition:

$$m_1 r_1 = m_2 r_2 \qquad \text{and} \qquad r = r_1 + r_2$$

This gives

$$r_1 = \frac{m_2}{m_1 + m_2} r \qquad \text{and} \qquad r_2 = \frac{m_1}{m_1 + m_2} r$$

The kinetic energy is given by:

$$KE = \frac{1}{2} m_1 \mathrm{v}_1^2 + \frac{1}{2} m_2 \mathrm{v}_2^2$$

but note that

$$\mathrm{v}_1 = r_1 \omega \qquad \text{and} \qquad \mathrm{v}_2 = r_2 \omega$$

We therefore have:

$$KE = \frac{1}{2} m_1 r_1^2 \omega^2 + \frac{1}{2} m_2 r_2^2 \omega^2$$

$$= \frac{1}{2} \left(m_1 r_1^2 + m_2 r_2^2 \right) \omega^2$$

$$= \frac{1}{2} I \omega^2$$

where

$$I = m_1 r_1^2 + m_2 r_2^2 \qquad \text{or} \qquad I = \frac{m_1 m_2}{m_1 + m_2} r^2 = \mu r^2$$

So we find that the moment of inertia of a diatomic is the same as that of a single particle of mass μ rotating about a point.

As an aside, note the symmetry between linear motion and rotational motion.

Linear Motion	***Rotational Motion***
x	θ
$\mathrm{v} = \frac{dx}{dt}$	$\omega = \frac{d\theta}{dt}$
$p = m\mathrm{v}$	$L = I\omega$
$KE = \frac{1}{2}m\mathrm{v}^2 = \frac{p^2}{2m}$	$KE = \frac{1}{2}I\omega^2 = \frac{L^2}{2I}$
m	I

Angular displacement is analogous to linear displacement. Moment of inertia is analogous to mass. Angular velocity is analogous to linear velocity. Angular momentum is analogous to linear momentum.

We now need to write down the Schrödinger equation for the rigid rotor. By our definition of the rigid rotor, the potential term in the Hamiltonian is zero since the distance r is fixed, $U(r) = 0$. We can set the zero of energy wherever we like and we choose it to be zero at the distance r. Since r is fixed, the potential is always zero. Thus, the Hamiltonian consists only of a kinetic energy term.

Because r is fixed for the rigid rotator, this problem has spherical symmetry. That is, $U = 0$ falls into the category of $U = U(r)$ (independent of θ and φ). We would therefore like to write down the Hamiltonian in spherical polar coordinates to take advantage of the symmetry of the system.

We can write down the classical expression for the rotational kinetic energy:

$$KE_{Classical} = \frac{1}{2}I\omega^2 = \frac{L^2}{2I}$$

and then find the Hamiltonian by substituting the $\hat{L}^2$ operator in spherical polar coordinates. This is a fairly lengthy procedure. One writes down

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

and then substitutes in the expressions for the L_x, L_y, and L_z operators (I had given you these when we were discussing postulate 2). Once you substitute in the operators, you then have to convert all the partial derivatives to spherical polar coordinates.

The easier way to do it is to realize that we have already converted the ∇^2 operator into spherical polar coordinates. Recall that the Schrödinger equation for the internal or relative motion of a two-particle system is

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + U\right]\psi = E_{Int}\psi$$

For the rigid rotator, $U = 0$, so we can write

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2$$

Recall that ∇^2 in spherical polar coordinates is

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

For the rigid rotor, *r is constant*, so the first term of this operator operating on a function will just give zero (*i.e.* there will be no *r* dependence of the wave function if *r* is fixed).

The other way to look at it is that the first term of this operator (when multiplied by $-\hbar^2/2\mu$) represents the radial part of the kinetic energy. Since *r* is fixed, there is no radial contribution to the kinetic energy. So we have:

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 = -\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]$$

We can use this to determine the operator for L^2 if we would need it (we will need it shortly when we discuss angular momentum). If we take the $1/r^2$ outside we can write

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 = -\frac{\hbar^2}{2\mu r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]$$

$$= -\frac{\hbar^2}{2I}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]$$

Comparing this to the expression for the kinetic energy in terms of the angular momentum operator we have

$$\hat{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]$$

Getting back to the Schrödinger equation we have

$$\hat{H}\psi(\theta,\varphi) = E\psi(\theta,\varphi)$$

By convention, the wave functions $\psi(\theta,\varphi)$ for the rigid rotor are usually denoted $Y(\theta,\varphi)$:

$$\hat{H}Y(\theta,\varphi) = EY(\theta,\varphi)$$

There is no *r* dependence to the functions because *r* is not a variable.

The eigenfunctions $Y(\theta,\varphi)$ are called ***spherical harmonics*** and arise as solutions to the angular part of all spherically symmetric potentials, including the hydrogen atom.

Writing out the Schrödinger equation in more detail

$$-\frac{\hbar^2}{2I}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]Y(\theta,\varphi) = EY(\theta,\varphi)$$

Let us make the substitution

$$\beta = \frac{2IE}{\hbar^2}$$

and multiply through by $\sin^2\theta$ to get

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right)+\frac{\partial^2 Y(\theta,\varphi)}{\partial\varphi^2}+\beta\sin^2\theta Y(\theta,\varphi) = 0$$

We will once again use the technique of separation of variables to simplify this partial differential equation.

Let

$$Y(\theta,\varphi)=\Theta(\theta)\Phi(\varphi)$$

Substitute this into the differential equation and divide both sides by $\Theta(\theta)\Phi(\varphi)$:

$$\frac{\sin\theta}{\Theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta(\theta)}{\partial\theta}\right)+\beta\sin^2\theta+\frac{1}{\Phi(\varphi)}\frac{\partial^2\Phi(\varphi)}{\partial\varphi^2}=0$$

You can see that the first two terms are only functions of θ and the last is only a function of φ. For this equation to hold for all θ and φ each part must separately equal a constant.

Let us call the separation constant m^2.

$$\frac{\sin\theta}{\Theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta(\theta)}{\partial\theta}\right)+\beta\sin^2\theta=m^2$$

$$\frac{1}{\Phi(\varphi)}\frac{\partial^2\Phi(\varphi)}{\partial\varphi^2}=-m^2 \qquad \text{or} \qquad \frac{\partial^2\Phi(\varphi)}{\partial\varphi^2}+m^2\Phi(\varphi)=0$$

The solution to the Φ part is straightforward; in fact we have done it before.

By inspection

$$\Phi(\varphi)=A_m e^{im\varphi} \qquad \text{or} \qquad \Phi(\varphi)=A_m e^{-im\varphi}$$

These solutions represent clockwise and counterclockwise rotation in the angle φ respectively.

Since the wave function must be single valued and continuous, we must require that $\Phi(\varphi)=\Phi(\varphi+2\pi)$

$$A_m e^{im\varphi}=A_m e^{im(\varphi+2\pi)}$$

consequently

$$e^{im2\pi}=1$$

This will be true when m = 0, ±1, ±2, . . .

We could have applied the boundary conditions to the $e^{-im\varphi}$ term as well and gotten the same result (or we could have applied it to a linear combination of the two solutions).

The constant A_m is determined by the normalization condition:

$$\int_0^{2\pi}\Phi^*(\varphi)\Phi(\varphi)d\varphi=1$$

$$A_m^2\int_0^{2\pi}e^{-im\varphi}e^{im\varphi}\,d\varphi=A_m^2\int_0^{2\pi}d\varphi=1$$

$$A_m^2 2\pi=1 \qquad \Rightarrow \qquad A_m=\frac{1}{\sqrt{2\pi}}$$

So the solution to the Φ part of the equation is

$$\Phi(\varphi) = \frac{1}{\sqrt{2\pi}} e^{im\varphi} \qquad m = 0, \pm 1, \pm 2, \ldots$$

It is clear that this is exactly the same as the particle on a ring. It should be. It is a one-dimensional rotational problem with fixed distance.

We have now solved the φ part of the rigid rotator problem. We still need to solve the differential equation for the θ part

$$\frac{\sin\theta}{\Theta(\theta)} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial \Theta(\theta)}{\partial\theta} \right) + \beta \sin^2\theta = m^2$$

If we take this equation and let $x = \cos\theta$ and $\Theta(\theta) = P(x)$ one could show that the differential equation becomes

$$\left(1 - x^2\right) \frac{d^2P(x)}{dx^2} - 2x \frac{dP(x)}{dx} + \left[\beta - \frac{m^2}{1 - x^2} \right] P(x) = 0$$

This differential equation is called **Legendre's equation** and is well known in classical physics. It occurs in a variety of problems which have spherical symmetry. The solution to this equation is similar to our solution to the differential equation that arose from the Harmonic Oscillator problem. The standard technique to solving this is to expand the function in a power series, generate a recursion relation, and then break off the power series after a finite number of terms to force it to be well behaved at the boundaries.

We will not go through the details but in an analogous way to the Harmonic oscillator problem, in breaking off the power series, one gets the quantization of the energy:

$$\beta = l(l+1)$$

but recall that

$$\beta = \frac{2IE}{\hbar^2}$$

so

$$E = \frac{\hbar^2}{2I} l(l+1) \qquad l = 0, 1, 2, \ldots$$

I will first discuss the wave functions and will then come back and discuss the implications of these energy eigenvalues.

The solutions to Legendre equation

$$\left(1 - x^2\right) \frac{d^2P(x)}{dx^2} - 2x \frac{dP(x)}{dx} + \left[l(l+1) - \frac{m^2}{1 - x^2} \right] P(x) = 0$$

when m=0 are called the **Legendre Polynomials** and are designated by the value of l as a subscript, *i.e.* $P_0(x)$, $P_1(x)$, etc.

These are listed in many standard texts in mathematics and physics. The first few Legendre Polynomials are:

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}\left(3x^2 - 1\right)$$

$$P_3(x) = \frac{1}{2}\left(5x^3 - 3x\right)$$

$$P_4(x) = \frac{1}{8}\left(35x^4 - 30x^2 + 3\right)$$

Note that there is some similarity to the Hermite polynomials from the harmonic oscillator problem

To generate the function $\Theta(\theta)$ you must substitute $x = \cos\theta$.

The solutions for $m \neq 0$ are called the **Associated Legendre Functions**, and these are related to the Legendre Polynomials. They are designated by a subscript for *l* and a superscript indicating the absolute value of *m*,

$$P_l^{|m|}(x)$$

The Associated Legendre functions can be related to the Legendre Polynomials by

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

Since only the absolute value of *m* is included here, then the Θ equation for +*m* and -*m* is the same. (The overall functions are different however because the Φ part is different.)

Note that the Legendre Polynomials are also associated Legendre functions; hence the solution to the problem is usually stated in terms of the latter (*i.e.* the superscript for $|m|$ is retained even when *m*=0).

The first few associated Legendre functions are:

	l=0	*l*=1	*l*=2	*l*=3
m=0	$P_0^0 = 1$	$P_1^0 = \cos\theta$	$P_2^0 = \frac{1}{2}\left(3\cos^2\theta - 1\right)$	$P_3^0 = \frac{1}{2}\left(5\cos^3\theta - 3\cos\theta\right)$
m=±1		$P_1^1 = \sin\theta$	$P_2^1 = 3\cos\theta\sin\theta$	$P_3^1 = \frac{3}{2}\left(5\cos^3\theta - 1\right)\sin\theta$
m=±2			$P_2^2 = 3\sin^2\theta$	$P_3^2 = 15\cos\theta\sin^2\theta$
m=±3				$P_3^3 = 15\sin^3\theta$

Recall that our solutions to the full Schrödinger equation for the Rigid Rotator are of the form

$$Y(\theta,\varphi) = \Theta(\theta)\Phi(\varphi)$$

Combining the Θ and Φ solutions we then have

$$Y_l^m(\theta,\varphi) = N_{lm}P_l^{|m|}(\cos\theta)e^{im\varphi}$$

From the properties of the Associated Legendre functions and from our normalization of the Φ equation one can show that the normalization constant is

$$N_{lm} = \left[\frac{(2l+1)}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{\frac{1}{2}}$$

The Rigid Rotator wave functions can then be written:

$$Y_l^m(\theta,\varphi) = \left[\frac{(2l+1)}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{\frac{1}{2}} P_l^{|m|}(\cos\theta) e^{im\varphi}$$

Recall that these functions are called **Spherical Harmonics.**

From the definition of the associated Legendre polynomials one finds that the quantum number *m* can only go as high as ±*l*. Thus $m = 0, \pm 1, \pm 2, \ldots, \pm l$ We will derive this explicitly later in the course.

Rigid Rotor Wave Functions

The first few spherical harmonics are

	l=0	*l*=1	*l*=2
m=0	$Y_0^0 = \frac{1}{\sqrt{4\pi}}$	$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$	$Y_2^0 = \sqrt{\frac{5}{16\pi}}(3\cos\theta - 1)$
m=+1		$Y_1^1 = \sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}$	$Y_2^1 = \sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta e^{i\varphi}$
m=-1		$Y_1^{-1} = \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\varphi}$	$Y_2^{-1} = \sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta e^{-i\varphi}$
m=+2			$Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{i2\varphi}$
m=-2			$Y_2^{-2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-i2\varphi}$

As the energy is given by:

$$E = \frac{\hbar^2}{2I} l(l+1) \qquad l = 0,1,2,\ldots$$

and does not depend upon *m* there is some degeneracy in that different eigenfunctions will have the same energy.
Thus, each column represents a set of degenerate functions. Think about the physical interpretation of these functions in terms of probabilities. What do they mean?

These functions form an orthonormal set. The orthonormality of these eigenfunctions is expressed by

$$\int_0^{2\pi}\int_0^{\pi} Y_l^m(\theta,\varphi)^* Y_l^m(\theta,\varphi)\sin\theta\, d\varphi\, d\theta$$

It is important to note the $\sin\theta$ in the integral and understand where it comes from. The volume element in spherical polar coordinates is $dx\,dy\,dz = r^2 \sin\theta\, dr\, d\theta\, d\varphi$. r is a constant for the rigid rotator. Integration of the wave function over r must just yield 1, since there must be unit probability of finding the particle at some r. Therefore the orthonormality integral reduces to an integral over only θ and φ. The θ and φ part of the volume element is $\sin\theta\, d\theta\, d\varphi$. Physically this represents the differential element of surface area of a sphere of unit radius. This is where the name *spherical harmonics* arises.

4.5 Spectroscopic Implications of the Rigid Rotor Energy Levels

Let's now consider implications of the energy eigenvalues:

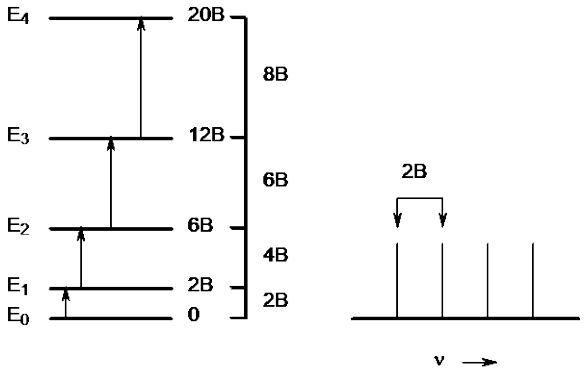
$$E = \frac{\hbar^2}{2I} l(l+1) = Bl(l+1) \qquad l = 0,1,2,3,\ldots$$

Here $$B = \frac{\hbar^2}{2I} = \frac{\hbar^2}{2\mu r^2}$$

is known as the rotational constant and characteristic for every molecule.

You can see that the energies are independent of the m quantum number. This means that there is degeneracy.

The first level (l=0) occurs at an energy of 0. The second level occurs at 2B. The next at 6B. This is shown in the figure below.



Energy Levels Spectrum

In considering how a rigid rotator interacts with light, we must consider the selection rules, just as we did for the harmonic oscillator. In the course "Spectroscopy" we will show that transitions between rotational levels must follow *the selection rule* $\Delta l = \pm 1$, similar to the Harmonic oscillator problem.

As we will then also see that in addition to this selection rule, a molecule must possess a *permanent dipole moment* to undergo a transition from one rotational level to another. Thus, homonuclear diatomic molecules will not show a pure rotational spectrum because they do not have a permanent dipole moment.

Using the selection rule one finds for the energy differences:

$$\Delta E = E_{l+1} - E_l = B(l+1)(l+2) - Bl(l+1)$$

$$= 2B(l+1)$$

where we have designated l as the quantum number of the lower level.
So the first transition is at a frequency corresponding to 2B, the second at 4B, third at 6B, etc.

In the rigid rotator approximation, we expect transitions to be equally spaced by units of 2*B*. Deviations from this approximation will show up as transitions that are not quite equally spaced.

From the value of *B* measured spectroscopically, one can determine the moment of inertia and hence the equilibrium bond length since $I = \mu r^2$.

For $H^{35}Cl$, for example *B* is about 10.6 cm^{-1}. This value of *B* would yield a bond length of 1.29Å. The most accurate geometries of molecules are determined from microwave spectroscopy.

I would like to now return to our discussion of the rigid rotor (and 3-dimensional systems in general) and consider the subject of angular momentum in more detail.

5 Angular Momentum

Molecules are 3-dimensional objects that rotate in space. To understand the rotational motion of molecules, we must first understand how angular momentum is treated in quantum mechanics. To do this we need to look at the properties of angular momentum operators.

Let us look back at the Schrödinger equation for the rigid rotor

$$\hat{H}Y_l^m(\theta,\varphi) = EY_l^m(\theta,\varphi) = \frac{\hbar^2 l(l+1)}{2I}Y_l^m(\theta,\varphi)$$

Recall that since classically we find for the energy

$$E = \frac{L^2}{2I}$$

we could represent the Hamiltonian for the rigid rotor in terms of the $\hat{L}^2$ operator:

$$\hat{H} = \frac{\hat{L}^2}{2I}$$

Since the $\hat{H}$ and $\hat{L}^2$ operators only differ by a constant, $\frac{1}{2I}$, we can also write

$$\hat{L}^2 Y_l^m(\theta,\varphi) = \hbar^2 l(l+1)Y_l^m(\theta,\varphi)$$

This says that the spherical harmonics are also eigenfunctions of the $\hat{L}^2$ operator and that the square of the angular momentum can only have quantized values given by

$$L^2 = \hbar^2 l(l+1) \qquad l = 0,1,2,3,\ldots$$

I would like to consider further the implications of the quantization of the square of the magnitude of the angular momentum, $\hat{L}^2$. But first, I would like to make a brief digression to give you some commutator identities that we will need.

Digression: Commutator Identities

$$\left[\hat{A},\hat{A}^n\right] = 0 \qquad n = 0,1,2,3,\ldots$$

$$\left[k\hat{A},\hat{B}\right] = \left[\hat{A},k\hat{B}\right] = k\left[\hat{A},\hat{B}\right]$$

$$\left[\hat{A},\hat{B}+\hat{C}\right] = \left[\hat{A},\hat{B}\right] + \left[\hat{A},\hat{C}\right]$$

$$\left[\hat{A},\hat{B}\hat{C}\right] = \left[\hat{A},\hat{B}\right]\hat{C} + \hat{B}\left[\hat{A},\hat{C}\right]$$

$$\left[\hat{A}\hat{B},\hat{C}\right] = \left[\hat{A},\hat{C}\right]\hat{B} + \hat{A}\left[\hat{B},\hat{C}\right]$$

Since angular momentum is a vector quantity, we need to think about vectors as operators.

Consider the position vector operator:

$$\hat{\mathbf{r}} = (\hat{x}, \hat{y}, \hat{z})$$

or the momentum vector operator:

$$\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$$

where $\hat{x} \to x$ $\qquad$ $\hat{p}_x \to -i\hbar \frac{\partial}{\partial x}$

Each component of the vector is an operator. When we make a measurement of a vector quantity, we can measure its magnitude or one of its components (projection).

It will be helpful to know the commutators of the coordinates and the different components of the momenta with respect to one another:

$$[\hat{x}, \hat{y}] = 0$$

$$[\hat{p}_x, \hat{p}_y] = 0 \qquad \text{(because mixed partial derivatives commute)}$$

$$[\hat{x}, \hat{p}_y] = 0 \qquad \text{(all mixed terms like this commute)}$$

$$[\hat{x}, \hat{p}_x] = -i\hbar \left[x, \frac{\partial}{\partial x} \right]$$

To evaluate this, must operate on a function:

$$\left[x, \frac{\partial}{\partial x} \right] f(x) = x \frac{\partial f(x)}{\partial x} - \frac{\partial x f(x)}{\partial x} = x \frac{\partial f(x)}{\partial x} - f(x) - x \frac{\partial f(x)}{\partial x} = -f(x)$$

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

One could also show

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

and similarly for the other coordinates.

The angular momentum is a vector containing the components

$$\hat{L} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$$

and is defined in classical mechanics as:

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

We showed earlier in the course while discussing postulate 2 that

$$\mathbf{r} \times \mathbf{p} = (yp_z - zp_y)\mathbf{i} + (zp_x - xp_z)\mathbf{j} + (xp_y - yp_x)\mathbf{k}$$

Thus

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

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

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

It is important to know how these operators commute with each other.

$$\begin{aligned}
\left[\hat{L}_x, \hat{L}_y\right] &= \left[\left(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right), \left(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right)\right] \\
&= \left[\hat{y}\hat{p}_z, \hat{z}\hat{p}_x\right] - \left[\hat{y}\hat{p}_z, \hat{x}\hat{p}_z\right] - \left[\hat{z}\hat{p}_y, \hat{z}\hat{p}_x\right] + \left[\hat{z}\hat{p}_y, \hat{x}\hat{p}_z\right] \\
&= \hat{y}\left[\hat{p}_z, \hat{z}\hat{p}_x\right] + \left[\hat{y}, \hat{z}\hat{p}_x\right]\hat{p}_z + \hat{z}\left[\hat{p}_y, \hat{x}\hat{p}_z\right] + \left[\hat{z}, \hat{x}\hat{p}_z\right]\hat{p}_y \\
&= \hat{y}\hat{z}\left[\hat{p}_z, \hat{p}_x\right] + \hat{y}\left[\hat{p}_z, \hat{z}\right]\hat{p}_x + \hat{x}\left[\hat{z}, \hat{p}_z\right]\hat{p}_y + \left[\hat{z}, \hat{x}\right]\hat{p}_z\hat{p}_y \\
&= -i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y \\
&= i\hbar\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right) \\
&= i\hbar\hat{L}_z
\end{aligned}$$

Note that several terms are omitted above because certain commutators are zero. Remember that mixed partial derivatives commute, different position operators commute, and position operators commute with momenta in other coordinates

So

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z$$

$$\left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x$$

$$\left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y$$

Note that it is cyclic (*xyz*, *yzx*, *zxy*). This cyclic pattern is a property of all types of angular momentum.

Since these operators do not commute, they do not have a complete set of common eigenfunctions. We therefore cannot simultaneously define precise values to these quantities**.**

Remember: $\Delta A^2 \Delta B^2 = \sigma_A^2 \sigma_B^2 \geq -\frac{1}{4}\left(\int \psi^*(x)\left[\hat{\mathrm{A}}, \hat{\mathrm{B}}\right]\psi(x)dx\right)^2$

However, consider the operator

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

This represents the square of the magnitude of the total angular momentum.

Remember, the magnitude of the angular momentum is quantized.

$$\hat{L}^2 Y_l^m(\theta,\varphi) = \hbar^2 l(l+1) Y_l^m(\theta,\varphi)$$

We want to find the commutator of $\hat{L}^2$ with $\hat{L}_x$, $\hat{L}_y$, or $\hat{L}_z$.

$$\begin{aligned}
\left[\hat{L}^2, \hat{L}_z\right] &= \left[\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_z\right] \\
&= \left[\hat{L}_x^2, \hat{L}_z\right] + \left[\hat{L}_y^2, \hat{L}_z\right] + \left[\hat{L}_z^2, \hat{L}_z\right] \\
&= \hat{L}_x\left[\hat{L}_x, \hat{L}_z\right] + \left[\hat{L}_x, \hat{L}_z\right]\hat{L}_x + \hat{L}_y\left[\hat{L}_y, \hat{L}_z\right] + \left[\hat{L}_y, \hat{L}_z\right]\hat{L}_y \\
&= \hat{L}_x\left(-i\hbar\hat{L}_y\right) + \left(-i\hbar\hat{L}_y\right)\hat{L}_x + \hat{L}_y\left(i\hbar\hat{L}_x\right) + \left(i\hbar\hat{L}_x\right)\hat{L}_y \\
&= 0
\end{aligned}$$

We could similarly show that

$$\left[\hat{L}^2, \hat{L}_x\right] = 0 \qquad \left[\hat{L}^2, \hat{L}_y\right] = 0$$

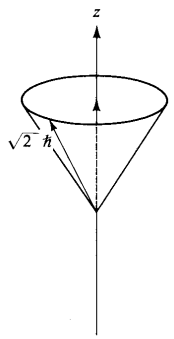
So each component of the angular momentum commutes with the $\hat{L}^2$ operator.

Since each component commutes with $\hat{L}^2$ but no component commutes with each other, we can only simultaneously specify one component along with $\hat{L}^2$ (with infinite precision, that is). We can therefore choose the eigenfunctions of $\hat{L}^2$ to also be eigenfunctions of one of these operators, and by convention this is chosen to be $\hat{L}_z$.

One could take L_x or L_y as well, but the set of eigenfunctions they share is different. I will say more on this later. Note that in specifying $L^2 = |\mathbf{L}|^2$ we are not specifying **L**, only its magnitude. A complete specification of **L** requires specification of its components, and in general we cannot do that (because all the operators don't commute).

In classical mechanics, when **L** is conserved each component has a definite value. In quantum mechanics, when **L** is conserved, we can only specify its magnitude and one of its components.

This is an important concept to understand.



So what we are doing is constraining the vector **L** to lie anywhere on a cone, since we know its magnitude and its L_z component (or any one of the three components).

We will often call the *z*-component of the angular momentum vector the projection of **L** on the *z*-axis or merely the projection of **L**.

I have already shown that the magnitude of the angular momentum vector is quantized, that is we showed that the eigenvalues of $\hat{L}^2$ were $\hbar^2 l(l+1)$. We can also show that the *z*-projection of the angular momentum is quantized. The easiest way to see this is to look at the $\hat{L}_z$ operator in spherical polar coordinates. To find this we simply need to take the expression in Cartesian coordinates and substitute the polar equivalent for the coordinates and the partial derivatives.

The result is

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

By inspection we can see that the eigenfunctions of this operator are $\propto e^{im\varphi}$

$$\hat{L}_z\left(e^{im\varphi}\right) = -i\hbar\frac{\partial}{\partial\varphi}\left(e^{im\varphi}\right) = m\hbar\left(e^{im\varphi}\right) \qquad m = 0,\pm1,\pm2,\ldots$$

You should recognize that these are just the functions describing the Φ part of the spherical harmonics. Consequently the eigenfunctions of the of the $\hat{L}^2$ operator are also eigenfunctions of the $\hat{L}_z$ operator.

Recall the definition of the spherical harmonics:

$$Y_l^m(\theta,\varphi) = N_{lm}P_l^{|m|}(\cos\theta)e^{im\varphi}$$

$$N_{lm} = \left[\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}\right]^{\frac{1}{2}}$$

Thus,

$$\hat{L}_z\,Y_l^m(\theta,\varphi) = -i\hbar\frac{\partial\left(N_{lm}P_l^{|m|}(\cos\theta)e^{im\varphi}\right)}{\partial\varphi} = m\hbar\left(N_{lm}P_l^{|m|}(\cos\theta)e^{im\varphi}\right) = m\hbar\,Y_l^m(\theta,\varphi)$$

The last thing we need to do is to see the connection between the allowable values of *m* and those of *l*.

Consider the following equation

$$\left(\hat{L}^2 - \hat{L}_z^2\right)Y_l^m(\theta,\varphi) = \left(\hat{L}_x^2 + \hat{L}_y^2\right)Y_l^m(\theta,\varphi) = \left[l(l+1) - m^2\right]\hbar^2 Y_l^m(\theta,\varphi)$$

I operated on the $Y_l^m(\theta,\varphi)$ functions to get the eigenvalues indicated, but I also used the relationship

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

to convert $\left(\hat{L}^2 - \hat{L}_z^2\right)$ to $\left(\hat{L}_x^2 + \hat{L}_y^2\right)$

Since $\hat{L}_x^2 + \hat{L}_y^2$ is the sum of two squared real terms it cannot be negative, so

$$\left[l(l+1) - m^2\right]\hbar^2 \geq 0$$

or $$l(l+1) \geq m^2$$

This says that

$$|m| \le l$$

or that the only possible values of m are:

$$m = 0, \pm 1, \pm 2, \ldots, \pm l$$

Consequently, there are $2l+1$ values of m for each value of l (l positive values, l negative values and 0). Thus,

$$\hat{L}^2 Y_l^m(\theta,\varphi) = l(l+1)\hbar^2 Y_l^m(\theta,\varphi) \qquad l = 0,1,2,3,\ldots$$

$$\hat{L}_z Y_l^m(\theta,\varphi) = m\hbar\, Y_l^m(\theta,\varphi) \qquad m = 0,\pm 1,\pm 2,\ldots,\pm l$$

Let us now consider the implications of the commutators of these operators on the measurement process. Recall that $\hat{H}_{rr}$ and $\hat{L}^2$ are simply related by a constant and thus commute.

The fact that the three operators

$$\hat{H}_{rr}, \hat{L}^2 \text{ and } \hat{L}_z$$

all commute has important implications on the measurement process. It says that the stationary states of the system (the eigenfunctions of $\hat{H}_{rr}$) can simultaneously have definite total energy, angular momentum (in magnitude), and have a well-defined projection on one axis (chosen by convention to be the z axis). They cannot, however, simultaneously have well defined values of L_x and L_y.

The fact that we can choose one axis to be special arises from the fact that the eigenvalues of $\hat{H}$ (the rigid rotor Hamiltonian) or of $\hat{L}^2$ only depend upon the quantum number l. Remember the theorem we introduced when we started multi-dimensional systems. Any linear combination of functions with the same eigenvalue with respect to a particular operator is also and eigenfunction of that operator.

For example, we said that there are $2l+1$ allowable values of m for each l,

$$m = 0, \pm 1, \pm 2, \ldots \pm l$$

but we know that the energy of the Rigid Rotor depends only upon l:

$$E = \frac{\hbar^2}{2I} l(l+1) \qquad l = 0,1,2,3,\ldots$$

This says that each level is $(2l+1)$-fold degenerate.

Any linear combination of the $2l+1$ degenerate eigenfunctions of $\hat{H}_{rr}$ and $\hat{L}^2$ are also eigenfunctions of $\hat{H}_{rr}$ and $\hat{L}^2$. By choosing z to be the special axis, we choose the linear combinations in such a way that they are eigenfunctions of $\hat{L}_z$. We could have chosen those linear combinations of $\hat{H}_{rr}$ and $\hat{L}^2$ eigenfunctions to be eigenfunctions of $\hat{L}_x$ or $\hat{L}_y$. However, since $\hat{L}_z$ is a simple operator in spherical polar coordinates, the mathematical form of the eigenfunctions is simplest if we choose them in as eigenfunctions of $\hat{L}_z$.

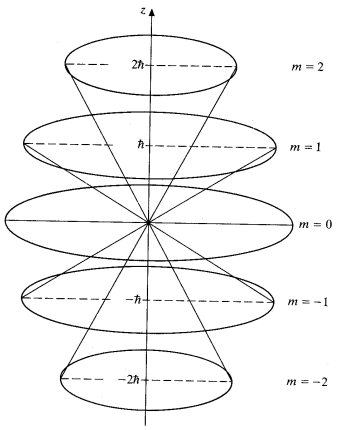
Consider an example where $l=2$. Because $l=2$, m can only have the values 0, ±1, ±2

Thus the magnitude of the vector **L**, that is |L| is

$$|L| = \sqrt{L^2} = \sqrt{l(l+1)}\hbar = \sqrt{6}\hbar$$

and the projection $\hat{L}_z$ can have the values: $-2\hbar$, $-\hbar$, 0, $\hbar$, $2\hbar$

We therefore have the following picture:



$$|L| = \sqrt{l(l+1)}\hbar = \sqrt{6}\hbar$$

$$m = -2,-1,0,1,2$$

$$L_z = m\hbar$$

If we specify the magnitude of the angular momentum vector and only one component, L_z, then we restrict the vector **L** to lie somewhere on a cone.

Remember that we cannot simultaneously specify L^2, L_x, L_y, and L_z because the components $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do not commute. Any one commutes with $\hat{L}^2$, so we can choose the eigenfunctions of $\hat{L}^2$ to simultaneously be eigenfunctions of one of those components.

By convention we chose $\hat{L}_z$. The vector can lie anywhere on a cone, since if we specify $\hat{L}_z$, we know nothing of $\hat{L}_x$ and $\hat{L}_y$ except that $\hat{L}_x^2 + \hat{L}_y^2 = \hat{L}^2 - \hat{L}_z^2$. (Since the eigenvalues of $\hat{L}^2$ and $\hat{L}_z$ are a constant $\hat{L}_x^2 + \hat{L}_y^2$ = constant just defines the circle at the top of the cone.)

Note that the maximum value of L_z is less than |**L**|, which says that L_z cannot point in the same direction as **L**. If it did, it would violate the Heisenberg Uncertainty Principle in that we could simultaneously know L_x, L_y, and L_z with infinite precision.

Remember, each vector that I have drawn represents the angular momentum of a particular eigenfunction. There are 5 vectors here since for l=2, there are 2l+1=5 eigenfunctions. All of these functions are degenerate (they have the same energy eigenvalue). I will speak more about degeneracy later. This is an important picture to keep in your mind. We will run into it again.

It is also important to realize that this picture is not specific to the Rigid Rotor. It just depends upon the angular momentum operators, their commutators and their eigenvalues, which is independent of the system. All angular momentum in quantum mechanics can be viewed this way. That is, the relationship between the angular momentum operators is the same for all systems. However, the fact that $\hat{H}$ commutes with the angular momentum operators need not be true. It will be true whenever the potential is spherically symmetric.

To summarize:

Since $\hat{H}_{rr}$, $\hat{L}^2$ and $\hat{L}_z$ all commute they have a common set of eigenfunctions, the spherical harmonics, and we can make simultaneous measurements of any of these and not affect the other. Each of these quantities has quantized eigenvalues, and we determined what the allowed values were.

Also we determined that since the energy doesn't depend upon *m*, all states with the same *l* but different *m* have the same energy. These states are (2*l*+1)-fold degenerate. The 2*l*+1 degenerate eigenfunctions correspond, for example, to the different positions of the vector **L** in the figure above.

6 The Hydrogen Atom

We are now ready to move on to a slightly more difficult quantum mechanical problem whose solution pervades much of the way we think about chemistry -- ***the hydrogen atom***. As we will see, the eigenfunctions of the hydrogen atom Hamiltonian serve as prototypes for more complex atoms and for molecules. The concepts we will develop will be familiar from first year chemistry, in that it provides the basis for talking about orbitals and their properties.

We will consider the hydrogen atom as a proton fixed at the origin and an electron with mass m_e interacting with the proton through a coulomb potential

$$U(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$$

Here, e is the electron charge, ε_0, the permittivity of free space, and r the distance between the proton and electron.

Recall that we can reduce a two-particle problem to a single particle problem with a reduced mass of μ by separating off the center of mass motion. In this case, the electron is so light compared to the proton, that the reduced mass of the system is not very different from the mass of the electron.

$$\mu = \frac{m_p m_e}{m_p + m_e} \cong \frac{m_p m_e}{m_p} = m_e$$

Nevertheless, one can calculate the reduced mass simply enough that we will use it.

The spherical symmetry of the system (*i.e.* the fact that the potential depends only upon the proton-electron distance and not the angles θ and φ) suggests that we should use spherical polar coordinates.

We can write down the Schrödinger equation for a hydrogen atom as

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi(r,\theta,\varphi) + U(r)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi)$$

If we substitute ∇^2 in spherical polar coordinates we get

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2}{\partial\varphi^2}\right)\right]\psi(r,\theta,\varphi) + U(r)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi)$$

Let's multiply through by $2\mu r^2$ and bring everything over to the left hand side:

$$-\hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi(r,\theta,\varphi)}{\partial r}\right) - \hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi(r,\theta,\varphi)}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\left(\frac{\partial^2\psi(r,\theta,\varphi)}{\partial\varphi^2}\right)\right] + 2\mu r^2\left[U(r) - E\right]\psi(r,\theta,\varphi) = 0$$

The second term here depends only upon θ and φ, and if you compare it to the expression we had for the $\hat{L}^2$ operator, you can see that we can write

$$-\hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi(r,\theta,\varphi)}{\partial r}\right) + \hat{L}^2\psi(r,\theta,\varphi) + 2\mu r^2\left[U(r) - E\right]\psi(r,\theta,\varphi) = 0$$

Since all the θ and φ dependence is carried by the $\hat{L}^2$ term, we can use the method of separation of variables once again. We can let the wave function be a function of r times a function of θ and φ.

Because the θ and φ part of the equation is simply an eigenvalue equation for $\hat{L}^2$, then we know what the θ and φ part of the wave function will be, the eigenfunctions of the $\hat{L}^2$ operator.

$$\hat{L}^2 Y_l^m(\theta,\varphi) = \hbar^2 l(l+1) Y_l^m(\theta,\varphi) \qquad l = 0,1,2,3,\ldots$$

The total wave function can be written

$$\psi(r,\theta,\varphi) = R(r) Y_l^m(\theta,\varphi)$$

So we had already solved the angular part of the hydrogen atom when we found the eigenfunctions of the $\hat{L}^2$ operator.

Notice that this result is perfectly general for any case when the potential U is spherically symmetrical (that is only a function of *r*). We saw this in the case of the rigid rotor where *U*(*r*)=0 since *r* was constant.

If we substitute this form for the wave function back into the differential equation, divide by $2\mu r^2$ and use the fact that we know the eigenvalues of $\hat{L}^2$ we get the following equation for *R*(*r*):

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)+\left[\frac{\hbar^2 l(l+1)}{2\mu r^2}+U(r)-E\right]R(r)=0$$

This is called the radial equation for the hydrogen atom because its solutions yield *R*(*r*), the radial part of the wavefunction. Solving this equation is the only new part of the hydrogen atom problem, since we already know the solution to the angular part.

It is important to have some feel for the meaning of the terms in this equation.

- The first term represents the radial kinetic energy of the hydrogen atom system, *i.e*. the energy due to the change in *r*.
- The second term is the angular kinetic energy term (recall that it originates from the $\hat{L}^2$ term).
- The third is the potential energy.

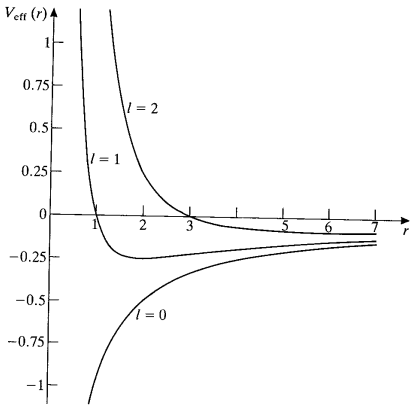
The angular kinetic energy term $\frac{l(l+1)\hbar^2}{2\mu r^2}$ is called the centrifugal potential. It is frequently grouped together with the coulomb potential term to yield an effective potential

$$U_{eff}(r) = U(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} = -\frac{e^2}{4\pi\varepsilon_0 r} + \frac{l(l+1)\hbar^2}{2\mu r^2}$$

Let's take a look at this potential

The coulomb potential alone is purely attractive (and hence negative), going to -∞ at *r*=0. The centrifugal term is purely repulsive (positive). The positive $1/r^2$ term wins out over the negative $1/r$ term at *r*=0, and hence the effective potential goes to ∞ at *r*=0 for nonzero *l*.

The radial equation for the hydrogen atom must be solved by the series method similar to our solution for the Harmonic Oscillator. I will not repeat this treatment, but trust that you recall the general approach.



In a manner similar to these previous problems, when one forces to the wave functions to obey the boundary conditions, quantization of the energy arises. In this case, the quantization condition is

$$E_n = -\frac{\mu e^4}{8\varepsilon_0^2 h^2 n^2} \qquad n = 1,2,3,\ldots$$

This is often written in terms of the Bohr radius, a_0

$$a_0 = \frac{\varepsilon_0 h^2}{\pi \mu e^2} \qquad \text{(Recall the Bohr planetary model)}$$

The energy equation then becomes

$$E_n = -\frac{e^2}{8\pi\varepsilon_0 a_0 n^2} \qquad n = 1,2,3,\ldots$$

Notice that although the quantum number *l* appeared in the radial equation, the energy does not directly depend upon *l* (although we will see that the wave function certainly does). Rather, the energy depends inversely upon the square of a total quantum number *n*.

The restrictions on the quantum number *n* that arises in solving the radial equation is $n \geq l + 1$. Since , as we saw previously, the smallest value of $l = 0$, this is usually written as

$$0 \leq l \leq n-1 \qquad n = 1,2,3,\ldots$$

Note that the levels get closer together due to the $1/n^2$ dependence as one approaches zero energy. For positive values of the energy (as measured relative to the energy of the electron at infinite separation) the energies of the hydrogen atom are no longer quantized but continuous. This is because there is no boundary condition to keep the electron from going to $r=\infty$.

The Hamiltonian for any one-electron atom (such as He^+, Li^{2+}, Be^{3+}) is exactly the same as for the hydrogen atom, except that the nuclear charge changes. If we had solved the problem for the more general case of one-electron atoms, we would have had a factor of *Z* in the numerator of the Hamiltonian. In this case the energy is given by:

$$E_n = -\frac{\mu e^4 Z^2}{8\varepsilon_0^2 h^2 n^2} = -\frac{RZ^2}{n^2} \qquad n = 1,2,3,\ldots$$

where $R = R_H = 109{,}677.4212\ \text{cm}^{-1}$ for the hydrogen atom. (It is very slightly different for other nuclei because μ, the reduced mass of the electron changes as the mass of the nucleus changes). This is simply the Rydberg expression that we discussed in the early part of the course. Using this expression, one can therefore predict the spectra of one-electron atoms.

Let us now look at the hydrogen atom wavefunctions. Recall that because the potential energy term in the Hamiltonian, $U(r)$, is spherically symmetric, the wave function will be of the form

$$\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_l^m(\theta,\varphi)$$

	l=0	l=1	l=2
m=0	$Y_0^0 = \frac{1}{\sqrt{4\pi}}$	$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$	$Y_2^0 = \sqrt{\frac{5}{16\pi}}\left(3\cos\theta - 1\right)$
m=+1		$Y_1^1 = \sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}$	$Y_2^1 = \sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta e^{i\varphi}$
m=-1		$Y_1^{-1} = \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\varphi}$	$Y_2^{-1} = \sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta e^{-i\varphi}$
m=+2			$Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{i2\varphi}$
m=-2			$Y_2^{-2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-i2\varphi}$

where the $Y_l^m(\theta,\varphi)$ are the spherical harmonics. These were the angular solutions to the rigid rotor problem as well. The first few are shown in the table above.

Recall that the radial equation (given a few pages back) can be solved by using a power series solution similar to our approach for the harmonic oscillator.

The radial wave functions for the hydrogen atom can be given in terms of well-known polynomials called the **associated Laguerre functions**.

The general formula for the radial wave function is

$$R_{nl}(r) = \left[\frac{(n-l-1)!}{2n\left[(n+l)!\right]^3}\right]^{\frac{1}{2}}\left(\frac{2}{na_0}\right)^{l+\frac{3}{2}} r^l e^{-\frac{r}{na_0}} L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right)$$

where the $L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right)$ are the associated Laguerre functions and $a_0 = \frac{\varepsilon_0 h^2}{\pi\mu e^2}$ is called the Bohr radius.

The first few Laguerre functions are:

n=1	l=0	$L_1^1(x) = -1$	$x = \frac{2r}{a_0}$
n=2	l=0	$L_2^1(x) = -2!(2-x)$	$x = \frac{r}{a_0}$
	l=1	$L_3^3(x) = -3!$	
n=3	l=0	$L_3^1(x) = -3!\left(3 - 3x + \frac{1}{2}x^2\right)$	$x = \frac{2r}{3a_0}$
	l=1	$L_4^3(x) = -4!(4-x)$	
	l=2	$L_5^5(x) = -5!$	

The total wave function, $\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_l^m(\theta,\varphi)$, depends upon three quantum numbers. Note that the radial wave function depends both upon *n* and *l*, and the angular part of the wave function depends upon *l* and *m*.

***n* - the total quantum number** (note that *n* alone determines the energy)

This comes from the solution of the radial part of the wave function. Recall that the energy of the hydrogen atom is determined by the total quantum number *n*.

$$E_n = -\frac{\mu e^4}{8\varepsilon_0^2 h^2 n^2} \qquad n = 1,2,3,\ldots$$

***l* - the azimuthal quantum number or angular momentum quantum number**

The *l* quantum number arose from solving the θ part of the angular equations. It is related to the magnitude of the angular momentum vector.

l can take on the values $0 \le l \le n-1$ $\qquad n = 1,2,3,\ldots$

All the functions with the same value of *n* but different *l* have the same energy.

***m* - the magnetic quantum number**

It gets its name from the fact that in a magnetic field, states of different *m* which are degenerate split. It represents the projection of the angular momentum vector on the *z*-axis.

m can take on any of the 2*l*+1 values ranging from

-*l*, -*l*+1, -*l*+2, . . , 0, . . , *l* -2, *l* -1, *l*

Each *n*, *l* state is 2 *l* +1 degenerate.

For historical reasons, the *l* quantum number is usually denoted by letters instead of numbers.

l = 0 1 2 3 4 5 6 7 (after *f* they go alphabetically except for *j*)
s p d f g h i k

Also, when *l*=0 and *m*=0, *m* is dropped as a subscript of ψ, since if *l*=0, *m* must equal zero. Thus the first few wavefunctions are designated

ψ_{1s}	*n*=1	*l* =0	*m*=0
ψ_{2s}	*n*=2	*l* =0	*m*=0
$\psi_{2p\text{-}1}$	*n*=2	*l* =1	*m*=-1
ψ_{2p0}	*n*=2	*l* =1	*m*=0
ψ_{2p1}	*n*=2	*l* =1	*m*=1

Let us put everything together and look at the first few total wave functions.

$$\psi_{1s}(r,\theta,\varphi) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}$$

$$\psi_{2s}(r,\theta,\varphi)=\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(2-\frac{Z}{a_0}r\right)e^{-\frac{Zr}{2a_0}}$$

$$\psi_{2p\pm1}(r,\theta,\varphi)=\frac{1}{\sqrt{64\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}}re^{-\frac{Zr}{2a_0}}\sin\theta e^{\pm i\varphi}$$

$$\psi_{2p0}(r,\theta,\varphi)=\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}}re^{-\frac{Zr}{2a_0}}\cos\theta$$

The parameter *Z* in these functions is the charge of the nucleus. These equations are valid not only for the hydrogen atom, but for any one electron atom such as He^{+} or Li^{2+}. Only *Z* will differ.

Recall that the energy of the H atom is independent of *l* and *m*, so all the levels *n*=2 levels have the same energy. For a given *n* and *l*, the (2 *l* +1) *m* levels are degenerate, but as we will see, this degeneracy can be broken by a magnetic field.

There are several things I would like us to consider regarding these wave functions. First note that since the wave functions are the eigenfunctions of a Hermitian operator, they are orthogonal.

This is expressed by the following integral.

$$\int_0^\infty\int_0^\pi\int_0^{2\pi}\psi^*_{nlm}(r,\theta,\varphi)\psi_{n'l'm'}(r,\theta,\varphi)d\tau=\delta_{nn'}\delta_{ll'}\delta m_{mm'}$$

where $d\tau$ is the volume element in spherical polar coordinates.

This says unless *n*, *l*, and *m* quantum numbers are all the same the functions with be orthogonal.

The wave functions as I have given them to you are normalized and thus one can calculate the probability of finding the electron in any range of *r* or θ or φ, however one must be careful to do this properly.

The normalization integral is:

$$\int_0^\infty\int_0^\pi\int_0^{2\pi}\left(R_{nl}(r)Y_l^m(\theta,\varphi)\right)^*\left(R_{nl}(r)Y_l^m(\theta,\varphi)\right)d\tau=1$$

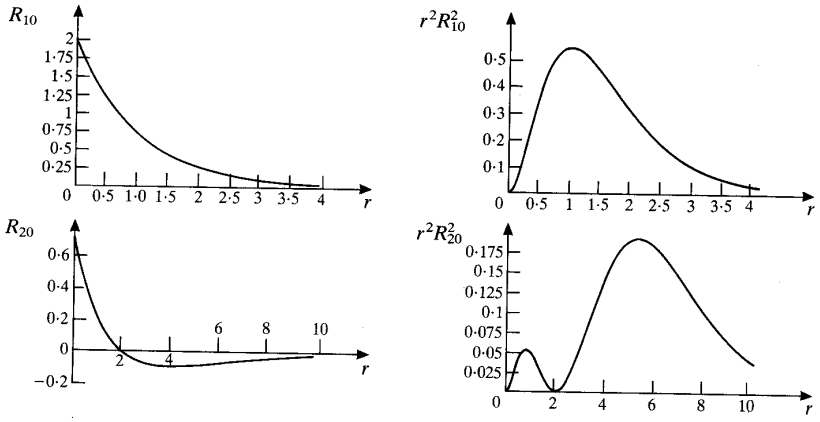
We can normalize each part separately, since the probability of finding the system with *r*=0 to ∞ must = 1, and similarly for finding θ between 0 and π and for φ between 0 and 2π.
However you must be careful to include the proper terms in the volume element for each coordinate.

$$\int_0^\infty\int_0^\pi\int_0^{2\pi}\left(R_{nl}(r)Y_l^m(\theta,\varphi)\right)^*\left(R_{nl}(r)Y_l^m(\theta,\varphi)\right)r^2\sin\theta dr d\theta d\varphi=\int_0^\infty R^*_{nl}(r)R_{nl}(r)r^2dr\int_0^\pi\int_0^{2\pi}Y_l^{m*}(\theta,\varphi)Y_l^m(\theta,\varphi)\sin\theta d\theta d\varphi$$

The probability of finding the electron between *r* and *r*+*dr* $\left|R_{nl}(r)\right|^2r^2dr$ since the functions $Y_l^m(\theta,\varphi)$ are normalized.

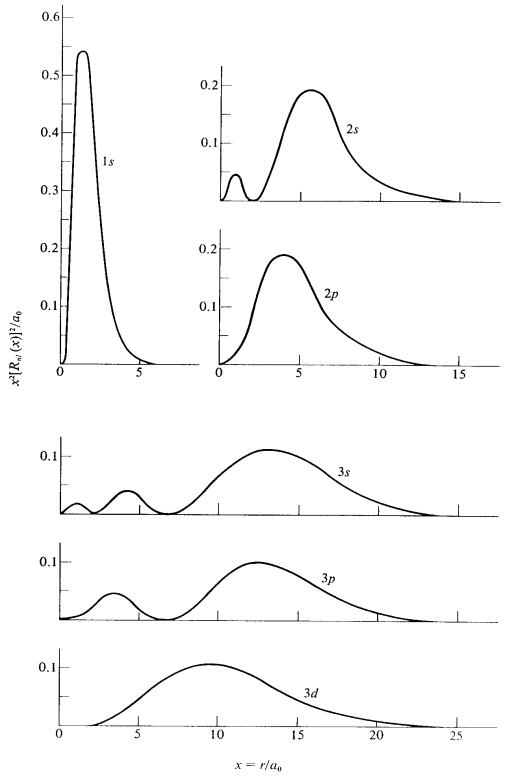
So to find the electron between r_1 and r_2 one must integrate $\left|R_{nl}(r)\right|^2r^2dr$ between these limits. Physically, the r^2 comes in that we are calculating the probability of finding the electron in a spherical shell between *r* and *r*+*dr.* The volume of that shell gets smaller as *r* gets smaller, so even though the 1*s* wave function is exponentially

decaying and peaks at r = 0, the probability density function $\left|R_{nl}(r)\right|^2 r^2$ peaks at some non-zero value of r and is zero at the origin.



Because the wave functions are three dimensional, it is difficult to visualize them and have physical intuition into their meaning. One way to do that is to plot the radial part of the wave function separately. What is probably more informative than the wave function itself is the radial probability distribution (as shown above for the 1*s* function), since that has a physical interpretation.

The radial probability distribution for the first functions is shown below.

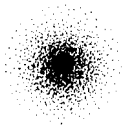


Note that the higher the energy, the larger the average r of the electron. Also note that there are n- l-1 nodes.

The angular parts of the wave functions are more difficult to display. However the wave functions with l=0 are somewhat easier to display. The angular part of the wave function for l=0, m=0 is

$$Y_0^0(\theta,\varphi) = \frac{1}{\sqrt{4\pi}}$$

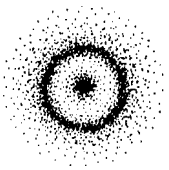
You can see that there is no explicit dependence upon θ and φ. Therefore, the wave function is spherically symmetric. Combining this angular dependence with the radial probability distribution we showed earlier, the picture is something like



$1s$

where the probability density is shown by the density of the dots.

A 2*s* wavefunction has the same angular part but with a different radial distribution

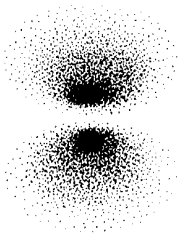


$2s$

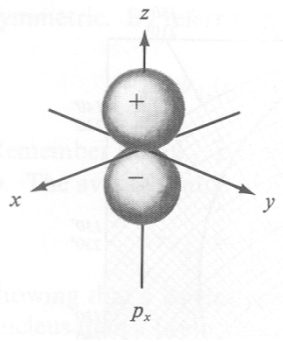
The wave function with *n*=2, *l*=1, *m*=0, that is the ψ_{2p0} function has somewhat of a different angular dependence.

$$\psi_{2p0}(r,\theta,\varphi)=\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\cos\theta$$

It looks something like this



$2p_0$



Note that the type of figure on the right which one commonly sees in textbooks only shows the angular part of the wavefunction. A shell is drawn containing a certain percentage of the probability density. One must combine this with the radial portion to get a good view of the electron density.

One way to help visualize this function is to make a substitution for the term *r* cosθ. If we realize that the Cartesian coordinate *z* in polar coordinates is *z* = *r* cosθ we could write the ψ_{2p0} wave function as

$$\psi_{2p0}(r,\theta,\varphi)=\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} ze^{-\frac{Zr}{2a_0}}$$

Even though it is a little awkward to mix coordinates here, this serves to emphasize the fact that the wave function has a ***nodal plane***, which in this case is the *xy* plane. The *z* in the wave function emphasizes this since when *z*=0 then the wave function and the probability equal zero. Because of this, ψ_{2p0} is usually called ψ_{2pz}

The wave functions corresponding to $m \neq 0$ (for example $m = \pm 1$) are somewhat more difficult to visualize because they are imaginary (due to the $e^{i\varphi}$ factor). We would like to work with real wave functions so we can easily plot them.
We can use the imaginary wave functions to construct real wave functions by taking linear combinations of the degenerate imaginary solutions. Remember, if two wavefunctions are solutions to the Schrödinger equation with the same eigenvalue (*i.e*. they are degenerate), any linear combination of them is also a solution with the same eigenvalue.

We can take an equally weighted linear combination of the *m*=±1 orbitals:

$$\psi_{2px}(r,\theta,\varphi) = \frac{1}{\sqrt{2}}\left(\psi_{2p-1}(r,\theta,\varphi) + \psi_{2p+1}(r,\theta,\varphi)\right)$$

You will see why we call it ψ_{2px} in a moment

Recall that $$\psi_{2p\pm1}(r,\theta,\varphi) = \frac{1}{\sqrt{64\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}} \sin\theta e^{\pm i\varphi}$$

Thus, $$\psi_{2px}(r,\theta,\varphi) = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{64\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}} \sin\theta\left(e^{-i\varphi} + e^{+i\varphi}\right)$$

By manipulating Euler's formula we can see that

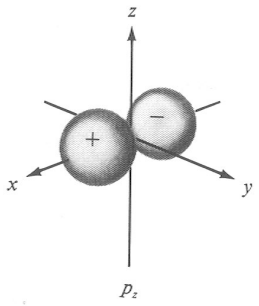
$$\cos\varphi = \frac{e^{-i\varphi} + e^{+i\varphi}}{2}$$

So, $$\psi_{2px}(r,\theta,\varphi) = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}} \sin\theta\cos\varphi$$

But recall the conversion from the Cartesian coordinate *x*, to spherical polar coordinates:

$$x = r\sin\theta\cos\varphi$$

Thus $$\psi_{2px}(r,\theta,\varphi) = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} xe^{-\frac{Zr}{2a_0}}$$



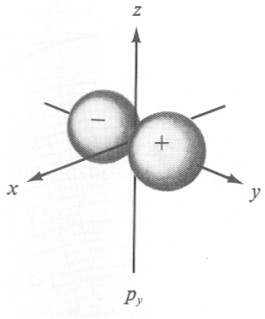
You can see that the *yz*-plane (*x*=0) is a nodal plane

One can similarly show that if we take the linear combination:

$$\psi_{2py}(r,\theta,\varphi)=\frac{1}{i\sqrt{2}}\left(\psi_{2p+1}(r,\theta,\varphi)-\psi_{2p-1}(r,\theta,\varphi)\right)$$

$$=\frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} ye^{-\frac{Zr}{2a_0}}$$

the nodal plane is the *xz*-plane (*y*=0).



So we now have three 2*p* functions, $\psi_{2px},\psi_{2py},\psi_{2pz}$ which are all real. We could demonstrate that $\psi_{2px},\psi_{2py},\psi_{2pz}$ are mutually orthogonal. Now, because $\psi_{2p-1},\psi_{2p0},\psi_{2p+1}$ all have the same energy eigenvalue, any linear combination will also have the same energy eigenvalue. The same holds true with respect to the $\hat{L}^2$ operator. $\psi_{2p-1},\psi_{2p0},\psi_{2p+1}$ all have the same eigenvalues when operated on by $\hat{L}^2$, hence their linear combinations will also. Hence $\psi_{2px},\psi_{2py},\psi_{2pz}$ are eigenfunctions of $\hat{L}^2$.

However, $\psi_{2p\pm1}$ have different eigenvalues ***with respect to*** $\hat{L}_z$ (different *m*'s), *thus their linear combinations are not eigenfunctions of* $\hat{L}_z$ (in general). The *m*=0 function will be since we didn't change it. Hence ψ_{2px},ψ_{2py} are not eigenfunctions of $\hat{L}_z$ (that is, *m* is not a good quantum number for these wave functions). It turns out that ψ_{2px} and ψ_{2py} are eigenfunctions of the $\hat{L}_x$ and $\hat{L}_y$ operators respectively. By looking at the functions you can see this intuitively. They have the exact same form as ψ_{2pz} but with different axis labels.

We can also take linear combinations of degenerate eigenfunctions for states of higher *l* to get real wavefunctions.

For a given *l*, *m* ranges from *-l . . . .l*

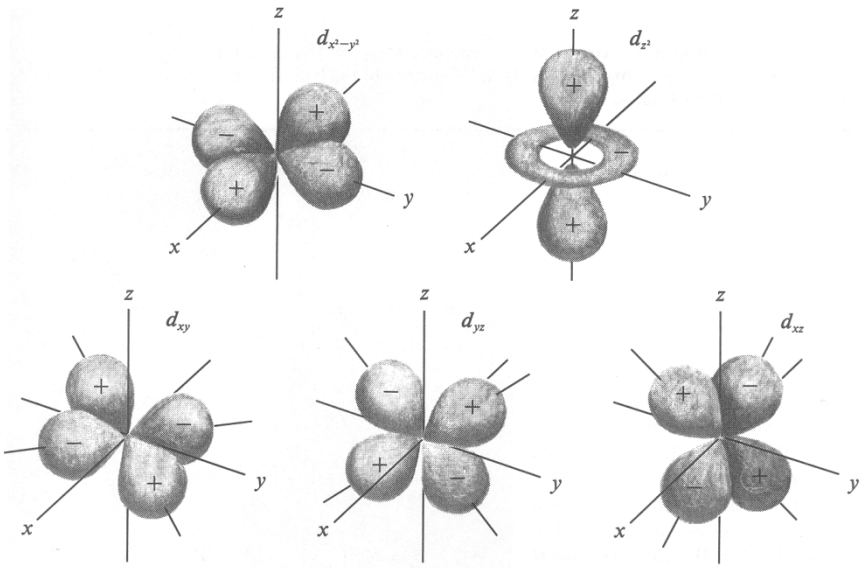
One will find that for every function containing an $e^{im\varphi}$, there is a corresponding one containing $e^{-im\varphi}$. Addition and subtraction of these functions will give two real functions. The subscripts on these new functions come from substituting Cartesian coordinates as we did with the ψ_{2px} and ψ_{2py} functions.

For example, for *n*=3, *l*=2, *m*=±2

$$\psi_{3d_{x^2-y^2}}(r,\theta,\varphi)=\frac{1}{\sqrt{2}}\left(\psi_{3d2}(r,\theta,\varphi)+\psi_{3d-2}(r,\theta,\varphi)\right)$$

$$=\frac{1}{81\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}}e^{-\frac{Zr}{3a_0}}\left(x^2-y^2\right)$$

Hence, this is called a $3d_{x^2-y^2}$ function.



Hydrogen-like wavefunctions are also called ***hydrogen-like orbitals.***

In first year chemistry you may have wondered where the strange labels for these hydrogen atom orbitals came from. It is precisely from the mathematical form of the wave functions after we take linear combinations of the imaginary functions with $\pm m$.

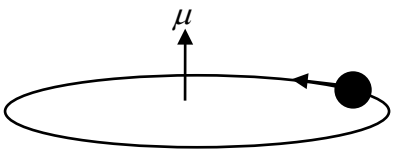
These hydrogen atom wave functions derive additional importance because when we get to many electron atoms, products of hydrogen atom functions are used as a first approximation to the eigenfunctions.

Also, linear combinations of hydrogen-like orbitals are used to understand molecular bonding. The terms *sp*, sp^2, and sp^3 hybridization originates from linear combinations of *s* and *p* orbitals. Remember the energy depends only upon *n*. So for example, a 2*s* and 2*p* have the same energy to first order.

6.1 The Zeeman effect

I would briefly like to consider how the hydrogen atom energy levels change when the atom is placed in a magnetic field. This will give you a bit more insight into the meaning of the magnetic quantum number, *m*. This effect is called the **Zeeman Effect**.

We must first consider the nature of the interaction of a moving charge with a magnetic field.



The motion of an electric charge around a closed loop produces a magnetic dipole μ whose magnitude is

$$\mu = iA$$

where i is the current in amps/sec and A is the area of the loop in m^2.

For a circular loop,

$$i = \frac{q\mathrm{v}}{2\pi r}$$

where q is the charge, v is the linear velocity and r is the radius of the loop.

The area of a circular loop $A = \pi r^2$ so

$$\mu = \frac{q\mathrm{v}r}{2}$$

However we must consider the fact that the magnetic dipole moment is a vector property with both magnitude and direction. This tells us the magnitude but not the direction of the magnetic dipole.

In general,

$$\boldsymbol{\mu} = \frac{q(\mathbf{r}\times\mathbf{v})}{2}$$

This is true whether or not you have a circular loop. You can show that if the loop is circular, then you will get the result we had previously.

This says that the direction of the magnetic moment is perpendicular to the plane of rotation as I have drawn above. Remember when doing a cross product you use the right hand rule to determine the direction of the resultant vector.

Note that

$$\mathbf{p} = m\mathbf{v},$$

so

$$\boldsymbol{\mu} = \frac{q(\mathbf{r}\times\mathbf{p})}{2m} = \frac{q}{2m}\mathbf{L}$$

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

For an electron q = -|e|, so

$$\boldsymbol{\mu} = -\frac{|e|}{2m_e}\mathbf{L}$$

Even though we cannot think about the electron in a hydrogen atom moving in a circular orbit, we know that it does have angular momentum and hence will have a magnetic moment. To understand how the magnetic field affects the energy levels of the hydrogen atom, we must add the energy of interaction between a magnetic field and a magnetic dipole into the Hamiltonian and solve the Schrödinger equation. The potential energy of interaction between a magnetic dipole and a magnetic field is given by

$$U = -\boldsymbol{\mu}\cdot\mathbf{B}$$

where **B** is the magnetic field. If we choose the magnetic field to be aligned along the *z*-direction, we have

$$U = -\boldsymbol{\mu}\cdot\mathbf{B} = \frac{|e|}{2m_e}\mathbf{L}\cdot\mathbf{B} = \frac{|e|}{2m_e}B_z L_z$$

We can therefore write the Hamiltonian for the H atom in a magnetic field as

$$\hat{H} = \hat{H}_0 + \frac{|e|B_z}{2m_e}\hat{L}_z$$

where $\hat{H}_0$, is the Hamiltonian in the absence of the field and $\hat{L}_z$ is the operator corresponding to the *z*-component of the angular momentum.

The Schrödinger equation is therefore

$$\hat{H}_0\psi + \frac{|e|B_z}{2m_e}\hat{L}_z\psi = E\psi$$

Recall that the hydrogen atom wavefunctions are eigenfunctions of both the $\hat{H}_0$, and the $\hat{L}_z$ operators. That is, we can write separate eigenvalue equations:

$$\hat{H}_0 R_{nl}(r)Y_l^m(\theta,\varphi) = -\frac{\mu e^4}{8\varepsilon_0^2 h^2 n^2}R_{nl}(r)Y_l^m(\theta,\varphi)$$

and $$\hat{L}_z R_{nl}(r)Y_l^m(\theta,\varphi) = m\hbar R_{nl}(r)Y_l^m(\theta,\varphi)$$

You can use these relations to see that:

$$E = -\frac{\mu e^4}{8\varepsilon_0^2 h^2 n^2} + \frac{|e|B_z}{2m_e}m\hbar \qquad \begin{array}{l} n = 1,2,3,\ldots \\ m = 0,\pm1,\pm2,\ldots\pm l \end{array}$$

Let's consider the implications of this.

In the absence of the field the hydrogen atom energies are just as we had previously calculated. However in the presence of the field, an extra term is added which depends on the value of *m*.

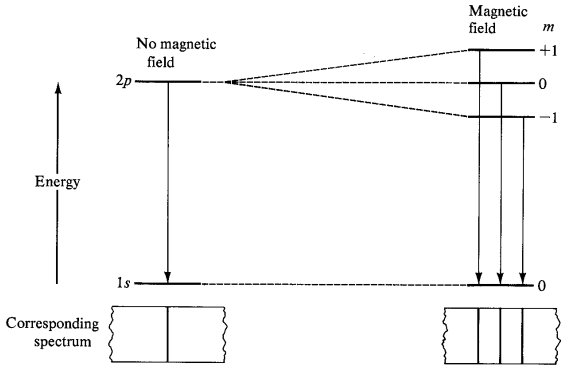
Consider the 2*p* level for example. *m* can take on the values *m* = -1, 0, 1

For *m*=0 there is no energy due to the magnetic field.

For *m*=1 there is a positive term added to the energy.

For *m* = -1 there is a negative term.

The energy that is added to each level is directly proportional to the magnetic field. Schematically this looks like the following:



The 2*p* level is a single level in the absence of the field splits up into three different levels in the absence of the field. The magnitude and direction of the shift depends upon the value of *m* and the magnetic field strength.

The splitting can be observed experimentally by looking at the transitions of the hydrogen atom. In the absence of the field, there will be a single transition from 1*s* to 2*p*, whereas in the presence of the field, the single line will be split into a triplet.

The splitting is relatively small relative to the 1*s*→2*p* transition energy. If we were considering a 3*d* level rather than a 2*p*, the level would be split into a quintet, since *m* can range from -2 to +2. One can use this effect to determine the value of *l* in the terminal level.

Pieter Zeeman

7 Approximation Methods

In our logical progression from simple systems to more complex systems of chemical interest, the next system to study after the hydrogen atom is the helium atom.

The Schrödinger equation for the helium atom is:

$$\left(-\frac{\hbar^2}{2M}\nabla^2-\frac{\hbar^2}{2m_e}\nabla_1^2-\frac{\hbar^2}{2m_e}\nabla_2^2\right)\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)+\left(-\frac{2e^2}{4\pi\varepsilon_0|\mathbf{R}-\mathbf{r}_1|}-\frac{2e^2}{4\pi\varepsilon_0|\mathbf{R}-\mathbf{r}_2|}+\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\right)\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)=E\psi(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2)$$

R is the position of the helium nucleus, $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the two electrons, M is the mass of the nucleus and m_e is the mass of the electrons. ∇^2 is the Laplacian operator with respect to the position of the nucleus (*i.e.* the derivatives are with respect to the nucleus position). ∇_1^2 and ∇_2^2 are the Laplacian operators with respect to the position of the electrons.

Because this is a three-body problem, the analytical separation between relative coordinates and center of mass coordinates is more complicated. However, because the nucleus of helium (with two protons and two neutrons) is 7300 times more massive than the electron, it is a good approximation to assume that the center of mass is at the nucleus. The center of mass motion is then the motion of this point through space. For the internal problem of the electrons in the center of mass coordinate system, this amounts to ignoring the kinetic energy of the nucleus, since its position is fixed at the origin. This is essentially equivalent to using m_e instead of μ in the hydrogen atom problem, but in this case it is even a better approximation because the helium nucleus is 4 times the mass of a proton.

We can therefore write (to a very high degree of approximation):

$$-\frac{\hbar^2}{2m_e}\left(\nabla_1^2+\nabla_2^2\right)\psi(\mathbf{r}_1,\mathbf{r}_2)-\frac{2e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\psi(\mathbf{r}_1,\mathbf{r}_2)+\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\psi(\mathbf{r}_1,\mathbf{r}_2)=E\psi(\mathbf{r}_1,\mathbf{r}_2)$$

Even though this is simplified by ignoring the nuclear kinetic energy, this Schrödinger equation cannot be solved exactly due to the term:

$$\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}$$

This term depends upon the relative separation of the two electrons and represents the potential energy due to inter-electronic repulsion.

If this term were gone, the Hamiltonian would just be a sum of two hydrogen atom Hamiltonians. In this case, the wave functions would simply be a product of H atom wave functions and the energy would be a sum of H atom energies. The presence of the inter-electronic repulsion term makes such a separation impossible, however. We must therefore turn to approximation techniques to deal with this problem. I will discuss approximation techniques in general, using several different systems as examples, and afterwards I will return specifically to the problem of the helium atom.

You can see that in any three-(or more)-particle system, we will always run into the problem of the Hamiltonian not being separable if the particles interact. Thus, the Schrödinger equation for many-electron atoms or molecules cannot be solved exactly.

However, there are two different approximation techniques which can yield extremely good results and are widely used in quantum mechanics.

The first of these approaches is called perturbation theory.

7.1 Perturbation Theory

Let me rewrite the Hamiltonian for the Helium atom in scalar form using the coordinate r_{12} to denote $|\mathbf{r_1}\text{-}\mathbf{r_2}|$.

$$\hat{H} = -\frac{\hbar^2}{2m_e}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{4\pi\varepsilon_0 r_{12}}$$

We can see that this Hamiltonian has the form

$$\hat{H} = \hat{H}_{H1} + \hat{H}_{H2} + \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r_{12}}$$

where

$$\hat{H}_{Hi} = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{2e^2}{4\pi\varepsilon_0}\frac{1}{r_i} \qquad i = 1,2$$

$\hat{H}_{Hi}$ is just the Hamiltonian for a hydrogen-like atom (a one-electron atom) in which the nuclear charge $Z = 2$. So the Hamiltonian for the Helium atom looks like the sum of two hydrogen atom Hamiltonians plus a term which represents the inter-electronic repulsion.

If the inter-electronic repulsion term,

$$\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r_{12}}$$

weren't there, we could separate the Hamiltonian into two Hamiltonians which depend on coordinates of different electrons. We would then know how to solve that problem. The wave functions would be the product of one-electron wave functions and the energies would be the sum of energies of each electron. However, the inter-electronic repulsion term prevents us from doing this.

Let's think back to another problem that we could not solve exactly – the anharmonic oscillator problem. Recall that the vibrational motions of real molecules are not exactly harmonic, but for small amplitude vibrations a harmonic approximation works pretty well. We demonstrated this by writing the potential energy as a Taylor series expansion about the bottom of the well:

$$U(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4 + \ldots$$

The Hamiltonian is then

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4 + \ldots$$

We then pointed out that for small amplitude vibrations, the latter two terms would be small and we could neglect them. The Hamiltonian for the *anharmonic* oscillator then reduces to that for the harmonic oscillator.

The problem of the Helium atom and that of the anharmonic oscillator have an important property in common -- they both can be written in the form

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $\hat{H}^{(0)}$ is the Hamiltonian for a problem that we know how to solve exactly, and $\hat{H}'$ is a small additional term that prevents an exact solution.

$\hat{H}^{(0)}$ is called the unperturbed Hamiltonian and $\hat{H}'$ is called the perturbation.

In each case we know how to solve the Schrödinger equation for $\hat{H}^{(0)}$, that is we know the solution of

$$\hat{H}^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)}$$

The basic approach of *perturbation theory* is to write the Hamiltonian in the form

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

If the perturbation to the Hamiltonian is sufficiently small, the energies and the wave functions will not be very different from the unperturbed problem. Perturbation theory then formulates the solution to the perturbed problem in the form of corrections to the energy and wave functions of the unperturbed problem.

I will call the unperturbed wave functions $\psi^{(0)}$ the *zeroth-order wave functions* and the energies $E^{(0)}$ the *zeroth-order energies*.

In the case of the Helium atom, the zeroth-order energy is the sum of two hydrogen atom energies and the zeroth-order wavefunctions the product of two hydrogen atom wave functions.

In the case of the anharmonic oscillator, the zeroth-order energies and wave functions are simply those of the harmonic oscillator problem.

We will solve these problems after we develop the theory.

7.1.1 Derivation of the Perturbation Theory

The basic problem is that we need to find solutions to the Schrödinger equation

$$\hat{H}\psi = E\psi$$

but many times we cannot solve it exactly. This means that the Hamiltonian is such that we cannot find analytic solutions for the energies and wave functions. However, we can often write $\hat{H}$ in the form:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $\hat{H}^{(0)}$ is a Hamiltonian for which we can solve the Schrödinger Equation exactly and $\hat{H}'$ is a small *perturbation* to that Hamiltonian.

For Perturbation Theory to yield reasonable results, the perturbation $\hat{H}'$ must be small compared to $\hat{H}^{(0)}$. To emphasize this we will write the perturbation as $\lambda\hat{H}'$ where λ is some small number <<1.

So we will in general write:

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}'$$

We never need to know what the parameter λ is besides to know that it is a small number. It is a useful construct that will drop out.

The problem we are trying to solve is then

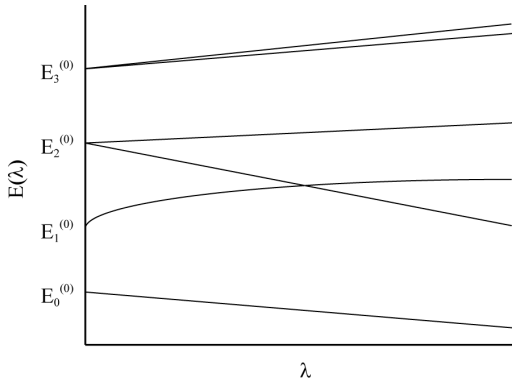
$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)\psi_n = E_n\psi_n$$

Since $\hat{H}$ depends on the parameter λ, both E_n and ψ_n will depend on λ, that is

$$\psi_n = \psi_n(\lambda, q) \qquad \text{and} \qquad E_n = E_n(\lambda)$$

where q are the coordinates.

We can look at the problem in the following way:



The eigenvalues will depend upon the value of the parameter λ, that is, how large the perturbation is. As $\lambda \to 0$, they will approach the eigenvalues of the unperturbed system.

There is at least one wave function for each energy, and for some there is more than one (degenerate states). As we turn on the perturbation by making λ different from zero, both the energies and eigenfunctions will change. One can actually do this with magnetic or electric fields. In some cases the perturbation will split the degeneracy ($E_3^{(0)}$), in others it may not ($E_4^{(0)}$). The perturbation treatment of degenerate and non-degenerate states differs. I will deal only with non-degenerate states.

Since $\hat{H}$ depends on λ, E_n and ψ_n should as well. We will therefore assume that we can expand the wave function ψ and the energy E_n in powers of λ,

$$\psi_n = \psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + \lambda^k E_n^{(k)} + \ldots$$

where we call $E_n^{(k)}$ and $\psi_n^{(k)}$ the k^{th} order corrections to the energy and wave function.

For small perturbations this series should converge, and the energy and wave function should be fairly well approximated by carrying just a few of the correction terms. Putting these expressions in to the Schrödinger equation:

$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots\right) =$$

$$\left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + \lambda^k E_n^{(k)} + \ldots\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \ldots + \lambda^k\psi_n^{(k)} + \ldots\right)$$

If we multiply out all the terms and group similar powers of λ we get:

$$H^{(0)}\psi_n^{(0)} + \lambda\left(\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}'\psi_n^{(0)}\right) + \lambda^2\left(\hat{H}^{(0)}\psi_n^{(2)} + \hat{H}'\psi_n^{(1)}\right) + \ldots$$

$$= E_n^{(0)}\psi_n^{(0)} + \lambda\left(E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}\right) + \lambda^2\left(E_n^{(2)}\psi_n^{(0)} + E_n^{(1)}\psi_n^{(1)} + E_n^{(0)}\psi_n^{(2)}\right) + \ldots$$

For this equality to hold, the coefficient of each power of λ must be equal.

For the λ^0 terms, we get:

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$$

There is no new information here, since this is simply the Schrödinger equation for the unperturbed system, the solution to which we already know.

For the λ^1 terms we get:

$$\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}'\psi_n^{(0)} = E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}$$

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)\psi_n^{(1)} = \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)}$$

We need to find $E_n^{(1)}$ and $\psi_n^{(1)}$. Since $H^{(0)}$ is Hermitian, its eigenfunctions form a complete orthonormal set. We can then expand $\psi_n^{(1)}$ in terms of these functions:

$$\psi_n^{(1)} = \sum_j a_j \psi_j^{(0)}$$

Our equation then becomes:

$$\sum_j a_j \left(\hat{H}^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} = \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)}$$

but we know that

$$\hat{H}^{(0)}\psi_j^{(0)} = E_j^{(0)}\psi_j^{(0)}$$

so we can write

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} = \left(E_n^{(1)} - H'\right)\psi_n^{(0)}$$

If we now multiply both sides by $\psi_n^{(0)*}$ and integrate, one has:

$$\int \psi_n^{(0)*} \sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\psi_j^{(0)} d\tau = \int \psi_n^{(0)*} \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)} d\tau$$

We can rewrite the left side of this equation

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\int \psi_n^{(0)*}\psi_j^{(0)} d\tau = \sum_j a_j \left(E_j^{(0)} - E_n^{(0)}\right)\delta_{nj}$$

You can see that the summation yields zero since when we do it, the only term that survives is when *j*=*n*, and

$$\left(E_n^{(0)} - E_n^{(0)}\right) = 0$$

Thus the right side of our equation must also equal zero

$$\int \psi_n^{(0)*} \left(E_n^{(1)} - \hat{H}'\right)\psi_n^{(0)} d\tau = 0$$

Solving for $E_n^{(1)}$

$$E_n^{(1)} = \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau$$

Remember that the 1st order correction term to the energy was given by $\lambda E_n^{(1)}$

$$\lambda E_n^{(1)} = \lambda \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau = \int \psi_n^{(0)*} \lambda \hat{H}' \psi_n^{(0)} d\tau$$

We define H'_{nn} as the integral of $\lambda H'$

$$H'_{nn} = \int \psi_n^{(0)*} \lambda \hat{H}' \psi_n^{(0)} d\tau$$

Recall that $\lambda H'$ is just the perturbation part of the Hamiltonian. Thus, the first order correction to the energy, $\lambda H'$, represents the *average value of the perturbation*.

We can now do away with the parameter λ. What we called $\lambda H'$ before we will now call H' – the perturbation part of the Hamiltonian. We can also absorb λ into the definition of the first order correction $E_n^{(1)}$ as well.

We can then write to first order:

$$E_n = E_n^{(0)} + E_n^{(1)} = E_n^{(0)} + H'_{nn}$$

where the first order correction to the energy, H'_{nn}, represents the average value of the perturbation.

We now need to find the first order correction to the wave function, $\psi_n^{(1)}$. Remember we had expanded $\psi_n^{(1)}$ in basis functions of the eigenfunctions of the unperturbed Hamiltonian:

$$\psi_n^{(1)} = \sum_j a_j \psi_j^{(0)}$$

Finding the correction to the wave function consists of finding the coefficients a_j in this expansion.

If we go back to our equation

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \psi_j^{(0)} = \left(E_n^{(1)} - \hat{H}' \right) \psi_n^{(0)}$$

and multiply by $\psi_m^{(0)*}$ and integrate (where $m \neq n$)

$$\int \psi_m^{(0)*} \sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \psi_j^{(0)} d\tau = \int \psi_m^{(0)*} \left(E_n^{(1)} - \hat{H}' \right) \psi_n^{(0)} d\tau$$

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \int \psi_m^{(0)*} \psi_j^{(0)} d\tau = E_n^{(1)} \int \psi_m^{(0)*} \psi_n^{(0)} d\tau - \int \psi_m^{(0)*} \hat{H}' \psi_n^{(0)} d\tau$$

$$\sum_j a_j \left(E_j^{(0)} - E_n^{(0)} \right) \delta_{mj} = E_n^{(1)} \delta_{mn} - H'_{mn}$$

$\delta_{mn} = 0$ since we are restricting ourselves to $m \neq n$. Also, we can evaluate the sum over *j* since only the term in which $j = m$ will survive:

$$a_m \left(E_m^{(0)} - E_n^{(0)} \right) = -H'_{mn}$$

Since $m \neq n$, we can divide both sides by $\left(E_m^{(0)} - E_n^{(0)}\right)$ (we can do this because states *m* and *n* are non-degenerate)

$$a_m = \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$ (Note the change of order and the loss of the minus sign)

So our first order correction to the wave function is

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)} \psi_m^{(0)}$$

where the sum is for $m \neq n$.

Note that although there is no $\psi_n^{(0)}$ in the sum, the zeroth-order wave function does appear in the expression for the corrected ψ_n.

The wave function ψ_n is then

$$\psi_n = \psi_n^{(0)} + \sum_{m \neq n} \frac{H'_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)} \psi_m^{(0)}$$

Note that once again, we have absorbed λ into our definition of H'_{mn}.

This quantity H'_{mn} is also an integral involving the perturbation part of the Hamiltonian, but now the integral involves two different functions. We call both H'_{nn} and H'_{mn} *matrix elements*, in that we can form a matrix of all the integrals of the perturbation with all of the different wave functions. The first order correction to the energy, H'_{nn} would lie along the diagonal, and the H'_{mn} would be the various off-diagonal elements.

Looking back at the expression for the first order correction to the wave function, you can see that the effect of a perturbation is to "mix" the unperturbed (or "zeroth-order") state with contributions from other states. In general, those contributions are larger the closer they are to the level under consideration because of the denominator of this expression.

Our treatment thus far has been concerned with finding only the first order correction to the wave function and energy. If we go back to our original Schrödinger equation in which we had substituted expansions in powers of λ for *E* and ψ and had grouped powers of λ, the terms involving λ^2 will give us the second order corrections. I will simply quote the result:

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left|H'_{mn}\right|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$

The energy to second order can therefore be written

$$E_n = E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{\left|H'_{mn}\right|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)}$$

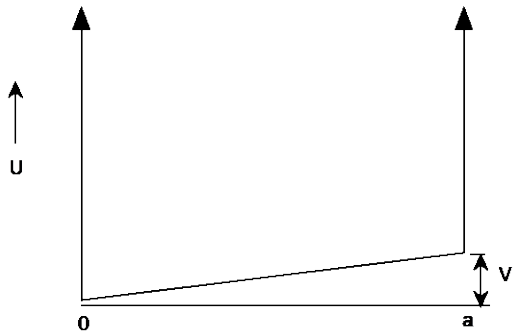
To get the first order correction to the energy, only a single integral needs to be evaluated, whereas for the second order correction to the energy and the first order correction to the wave function, we must evaluate all the integrals in the infinite sum. It is often difficult to evaluate these sums. For higher corrections it is more difficult still.

However, the most important terms come from states that are closest in energy to the one being calculated because of the $\frac{1}{\left(E_n^{(0)} - E_m^{(0)}\right)}$ term.

You should know what all the terms in the expression for the 1st order correction to the wave function and second order correction to the energy mean. Often, many of the matrix elements that you need to evaluate will be zero on the basis of symmetry. You should be aware of this and look for it.

7.1.2 Examples of Using Perturbation Theory

Example 1: Particle in a box with slanted bottom.



We need to formulate the problem as an zeroth-order part plus a perturbation. The zeroth-order part must be one that we know how to solve.

If we let $H^{(0)}$ be the particle in the box Hamiltonian, then we can write

$$\hat{H} = \hat{H}^{(0)} + \hat{H}'$$

where $$\hat{H}' = \frac{Vx}{a}$$

We know the zeroth-order energy is just the particle in a box energy. Let us then find the first order correction to the energies, $E_n^{(1)}$.

$$E_n^{(1)} = H'_{nn} = \int_0^a \psi_n^{(0)*}(x)\left(\frac{Vx}{a}\right)\psi_n^{(0)}(x)dx$$

$$= \frac{2}{a}\frac{V}{a}\int_0^a x\sin^2\left(\frac{n\pi}{a}x\right)dx$$

$$= \frac{2}{a}\frac{V}{a}\frac{a^2}{4}$$

So $$E_n^{(1)} = \frac{V}{2}$$

Note that this is independent of *n*. In general, the effect of the perturbation may depend upon the quantum numbers. In this case, however, each level is shifted by the same amount.

Thus, $$E_n = E_n^{(0)} + E_n^{(1)} = \frac{n^2h^2}{8ma^2} + \frac{V}{2}$$

Example 2: Anharmonic oscillator.

Recall that in discussing the harmonic oscillator problem I indicated that real molecules are somewhat anharmonic (that is, the potential contains terms other than $\frac{1}{2}kx^2$). Consider an anharmonic oscillator whose potential is given by

$$U(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4$$

Let us determine the first-order correction to the ground state energy of an anharmonic oscillator with a potential as given above. Comparing $U(x)$ above to the harmonic oscillator potential, we can see that

$$\hat{H}' = \frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4$$

Recall that the ground state wave function for the harmonic oscillator is

$$\psi_0(x) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2}x^2}$$

where $$\alpha = \sqrt{\frac{k\mu}{\hbar^2}}$$

The first order correction to the ground state energy is given by

$$E_0^{(1)} = H'_{00} = \int_{-\infty}^{\infty} \psi_0^{(0)*}(x)\hat{H}'\psi_0^{(0)}(x)dx$$

$$= \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} \left(\frac{1}{6}\gamma x^3 + \frac{1}{24}bx^4\right) e^{-\alpha x^2} dx$$

$$= \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \left[\int_{-\infty}^{\infty} \left(\frac{1}{6}\gamma x^3\right) e^{-\alpha x^2} dx + \int_{-\infty}^{\infty} \left(\frac{1}{24}bx^4\right) e^{-\alpha x^2} dx\right]$$

The first of these integrals equals zero since the function under the integral is odd. The second term gives

$$E_0^{(1)} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{2}{24} b \int_0^{\infty} x^4 e^{-\alpha x^2} dx = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \frac{b}{12} \frac{3\sqrt{\pi}}{8\alpha^{\frac{5}{2}}}$$

Thus, $$E_0^{(1)} = \frac{b}{32\alpha^2} = \frac{\hbar^2 b}{32k\mu}$$

The total ground state energy is then

$$E_0 = \frac{1}{2}h\nu + \frac{\hbar^2 b}{32k\mu} + \ldots$$

7.2 The Variational Principle

I would now like to introduce a second approximation method for solving the Schrödinger equation – the *variational principle*. This method does not require that a similar problem has been solved exactly as in the case of perturbation theory. It is widely used for solving the *electronic* Schrödinger equation (that is, the Schrödinger equation for electronic motion in molecules).

Consider the ground (lowest energy) state of some arbitrary system. The wave function ψ_0, and energy E_0, satisfy the relation

$$\hat{H}\psi_0 = E_0\psi_0$$

If we multiply this equation by ψ_0^* and integrate over all space we get

$$E_0 = \frac{\int \psi_0^* \hat{H} \psi_0 \, d\tau}{\int \psi_0^* \psi_0 \, d\tau}$$

where $d\tau$ is the volume element.

Note that ψ_0 is the true wave function for the lowest energy state, even though we don't know what it is.

We could set the denominator equal to one if we choose normalized functions, but if we leave it we can allow for the possibility that the functions are not normalized.

If we substitute any other function for ψ into this expression (any function that we might guess to be the true wave function), the variation principle says that energy we would calculate from this expression will be greater than or equal to the true ground state energy.

That is, if we calculate

$$E_\varphi = \frac{\int \Phi^* \hat{H} \Phi \, d\tau}{\int \Phi^* \Phi \, d\tau}$$

then

$$E_\varphi \geq E_0$$

The expression for E_φ is called the *variational function*, and the wave function Φ is called a *trial function*.

The only requirements on the function Φ are that it satisfies the proper boundary conditions. We can determine this without solving the problem.

7.2.1 Proof of the Variational Principle

Let us expand our function Φ as a linear combination of eigenfunctions of the Hamiltonian for the system of interest. That is:

$$\Phi = \sum_n c_n \psi_n$$

where $\hat{H}\psi_n = E_n\psi_n$

We can see that

$$\hat{H}\Phi = \sum_n \hat{H} c_n \psi_n = \sum_n c_n E_n \psi_n$$

Substituting these expressions back into the variational function we get

$$E_\varphi = \frac{\int \left(\sum_m c_m^* \psi_m^* \right) \left(\sum_n c_n E_n \psi_n \right) d\tau}{\int \left(\sum_m c_m^* \psi_m^* \right) \left(\sum_n c_n \psi_n \right) d\tau}$$

We can rearrange this to get

$$E_\varphi = \frac{\sum_m \sum_n c_m^* c_n E_n \int \psi_m^* \psi_n d\tau}{\sum_m \sum_n c_m^* c_n \int \psi_m^* \psi_n d\tau}$$

Because the eigenfunctions form a complete orthonormal set we can write:

$$E_\varphi = \frac{\sum_n c_n^* c_n E_n}{\sum_n c_n^* c_n}$$

If we now subtract E_0 from the left side and the quantity below (which is equal to E_0) from the right side

$$\frac{\sum_n c_n^* c_n E_0}{\sum_n c_n^* c_n} = E_0$$

we get

$$E_\varphi - E_0 = \frac{\sum_n c_n^* c_n \left(E_n - E_0 \right)}{\sum_n c_n^* c_n}$$

We know that $E_n \geq E_0$ for all n since E_0 is the lowest energy (the ground state).

Also, $c_n^* c_n \geq 0$. Thus, the right hand side of the equation is positive.

Thus we have

$$E_\varphi - E_0 \geq 0$$

or $$E_\varphi \geq E_0$$

Basically, *any trial function will always yield an energy greater than the exact solution.* We can then choose the function to contain "variational" parameters, and we can minimize E_φ as a function of the parameters. The value will be the best value for the adjustable parameter.

7.2.2 Examples of Variational Calculations

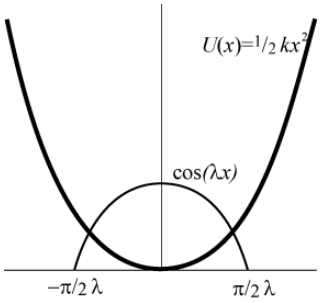
Example 1:

Consider the ground state of a harmonic oscillator. Assume for the sake of this example that we did not know how to solve it by exact methods. We might expect, however, that the ground state wave function would be symmetric about x=0 (by analogy to problems like the particle in a box).

Let us try the function

$$\Phi = \cos(\lambda x) \qquad -\frac{\pi}{2\lambda} \le x \le \frac{\pi}{2\lambda}$$

where λ is a variable parameter. We can vary the parameter λ to get the best estimate of the ground state energy.



We will only consider the function between $-\pi/2\lambda$ and $\pi/2\lambda$ since we know the function must damp out at x = $\pm\infty$ and not oscillate as this does. If we consider it just to the point where it goes to zero it should be ok.

Remember the variational function has the form

$$E_\varphi = \frac{\int \Phi^* \hat{H} \Phi \, d\tau}{\int \Phi^* \Phi \, d\tau}$$

Recall that the Hamiltonian for the harmonic oscillator is

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2$$

Substituting for Φ and $\hat{H}$

$$E_\varphi = \frac{\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right]\cos(\lambda x)\,dx}{\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\cos(\lambda x)\,dx}$$

Let's do the denominator first:

$$\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos^2(\lambda x)\,dx = \frac{\pi}{2\lambda}$$

Next, the numerator:

$$\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}} \cos(\lambda x)\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2}+\frac{1}{2}kx^2\right]\cos(\lambda x)dx = \frac{\hbar^2\lambda^2}{2\mu}\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}}\cos^2(\lambda x)dx+\frac{k}{2}\int_{-\frac{\pi}{2\lambda}}^{\frac{\pi}{2\lambda}}x^2\cos^2(\lambda x)dx$$

$$=\frac{\hbar^2\lambda^2}{2\mu}\left(\frac{\pi}{2\lambda}\right)+\frac{k}{\lambda^3}\left(\frac{\pi^3}{48}-\frac{\pi}{8}\right)$$

The function E_φ is then

$$E_\varphi = \frac{\frac{\hbar^2\lambda^2}{2\mu}\left(\frac{\pi}{2\lambda}\right)+\frac{k}{\lambda^3}\left(\frac{\pi^3}{48}-\frac{\pi}{8}\right)}{\frac{\pi}{2\lambda}}$$

$$=\frac{\hbar^2\lambda^2}{2\mu}+\frac{k}{\lambda^2}\left(\frac{\pi^2}{24}-\frac{1}{4}\right)$$

We now want to minimize this function with respect to the parameter λ. So we need to take the first derivative with respect to λ and set it equal to zero (recall that this is the prescription for finding a minimum or maximum of a function).

$$\frac{dE_\varphi}{d\lambda}=\frac{2\hbar^2\lambda}{2\mu}-\frac{2k}{\lambda^3}\left(\frac{\pi^2}{24}-\frac{1}{4}\right)=0$$

Multiplying both sides by λ^3 we get:

$$\frac{\hbar^2}{\mu}\lambda^4-2k\left(\frac{\pi^2}{24}-\frac{1}{4}\right)=0$$

The minimum is then at

$$\lambda_{\min}^2=+\frac{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}{\hbar}$$

Note that we took the positive root only as λ^2 must be positive.

Substituting λ^2 into E_φ:

$$E_{\varphi\min}=\frac{\hbar^2}{2\mu}\frac{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}{\hbar}+\left(\frac{\pi^2}{24}-\frac{1}{4}\right)\frac{k\hbar}{\sqrt{2k\mu\left(\frac{\pi^2}{24}-\frac{1}{4}\right)}}$$

This gives

$$E_{\varphi\min}=\sqrt{2}\sqrt{\frac{\pi^2}{24}-\frac{1}{4}}\hbar\sqrt{\frac{k}{\mu}}=0.57\hbar\omega=0.57h\nu$$

Compare this to the exact solution:

$$E_0 = 0.5\ h\nu$$

Note that $E_\varphi > E_0$. Also note that even with this crude approximation we get excellent results.

Example 2:

Now let us repeat the variational calculation using the trial function

$$\Phi = e^{-\frac{\alpha}{2}x^2}$$

Then

$$E_\varphi = \frac{\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2}+\frac{1}{2}kx^2\right]e^{-\frac{\alpha}{2}x^2}\,dx}{\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}e^{-\frac{\alpha}{2}x^2}\,dx}$$

Numerator:

$$\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}\left[-\frac{\hbar^2}{2\mu}\left(\alpha^2x^2e^{-\frac{\alpha}{2}x^2}-\alpha e^{-\frac{\alpha}{2}x^2}\right)+\frac{1}{2}kx^2e^{-\frac{\alpha}{2}x^2}\right]dx = \int_{-\infty}^{\infty}\frac{\hbar^2\alpha}{2\mu}e^{-\alpha x^2}dx+\int_{-\infty}^{\infty}\left(\frac{1}{2}k-\frac{\hbar^2\alpha^2}{2\mu}\right)x^2e^{-\alpha x^2}dx$$

$$= \frac{\hbar^2\alpha}{2\mu}\sqrt{\frac{\pi}{\alpha}}+\left(\frac{1}{2}k-\frac{\hbar^2\alpha^2}{2\mu}\right)\frac{1}{2\alpha}\sqrt{\frac{\pi}{\alpha}}$$

Denominator:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}dx = \sqrt{\frac{\pi}{\alpha}}$$

$$E_\varphi = \frac{\hbar^2\alpha}{2\mu}+\left(\frac{1}{2}k-\frac{\hbar^2\alpha^2}{2\mu}\right)\frac{1}{2\alpha}$$

So,

$$= \frac{\hbar^2\alpha}{2\mu}+\frac{1}{4\alpha}k-\frac{\hbar^2\alpha}{4\mu}$$

$$= \frac{\hbar^2\alpha}{4\mu}+\frac{1}{4\alpha}k$$

We now have to take the derivative and set it equal to zero to find the best value of α.

$$\frac{dE_\varphi}{d\alpha} = \frac{\hbar^2}{4\mu}-\frac{1}{4\alpha^2}k = 0$$

yielding:

$$\alpha^2 = \frac{k\mu}{\hbar^2}$$

Now substitute this back into the expression for E_φ to get the minimum value.

$$E_{\varphi\min} = \frac{\hbar^2}{4\mu}\sqrt{\frac{k\mu}{\hbar^2}} + \frac{k}{4}\sqrt{\frac{\hbar^2}{k\mu}}$$

$$= \frac{\hbar}{4}\sqrt{\frac{k}{\mu}} + \frac{\hbar}{4}\sqrt{\frac{k}{\mu}}$$

$$= \frac{\hbar}{2}\sqrt{\frac{k}{\mu}} = \frac{1}{2}\hbar\omega = \frac{1}{2}h\nu$$

Note: this is the same as the exact result!!

The better agreement (exact in this case) using this trial function stems from the fact that the function is closer to (in this case identical to) the true wave function. It came out exact because we chose a function that was in the exact same form as the ground state solution. The α we found was simply the definition for α we had used previously.

However, the general principle holds true: The closer the wave function is to the true one, the better the estimate of the energy. The general approach is to include many variable parameters in the trial function and vary them in a systematic manner to obtain the lowest energy possible. The wave function becomes closer to the exact solution.

In principle one can use this approach to find not only the ground state energy but also the ground state wave function, however the estimate of the energy approaches the true energy faster than the wave function approaches the true wave function. You therefore need an energy very close to the true energy before the wave function is close to the true one.

Up to now, we have only used a single variable parameter in our trial function. You can imagine that if we choose a trial function with several variable parameters, calculate the variational function, and then vary each of the parameters to minimize the energy, one could get an even better estimate of the energy.

7.2.3 Variational Principle with more than one variable parameter

A systematic way to handle a trial function with more than one variational parameter is to write it as a linear combination of functions

$$\Phi = \sum_{n=1}^{N} c_n f_n$$

where the c_n are the parameters which we are going to vary, and the f_n are arbitrary known functions that are not necessarily orthonormal.

The more terms we take in the sum, the better our estimate of the ground state energy will be. This accuracy comes at the cost of computational complexity, not in the sense of conceptual difficulty, but in the sense of the raw computer power required.

Consider the simplest case where we have a linear combination of only two functions

$$\Phi = c_1 f_1 + c_2 f_2$$

We will choose the c_n and f_n to be real.

Let us plug this into the variational function E_{φ}.

First evaluate the numerator:

$$\int \Phi^* \hat{H} \Phi \, d\tau = \int \left(c_1 f_1 + c_2 f_2\right)^* \hat{H} \left(c_1 f_1 + c_2 f_2\right) d\tau$$

$$= c_1^2 \int f_1 \hat{H} f_1 \, d\tau + c_1 c_2 \int f_1 \hat{H} f_2 \, d\tau + c_2 c_1 \int f_2 \hat{H} f_1 \, d\tau + c_2^2 \int f_2 \hat{H} f_2 \, d\tau$$

(I dropped the * since the *c*'s and *f*'s are real)

Let us define a matrix element H_{ij} as

$$H_{ij} = \int f_i \hat{H} f_j \, d\tau$$

Note that since $\hat{H}$ is Hermitian $H_{ij} = H_{ji}$

We can therefore write

$$\int \Phi^* \hat{H} \Phi \, d\tau = c_1^2 H_{11} + 2c_1 c_2 H_{12} + c_2^2 H_{22}$$

Now the denominator:

$$\int \Phi^* \Phi \, d\tau = \int \left(c_1 f_1 + c_2 f_2\right)^* \left(c_1 f_1 + c_2 f_2\right) d\tau$$

$$= c_1^2 \int f_1 f_1 \, d\tau + 2c_1 c_2 \int f_1 f_2 \, d\tau + c_2^2 \int f_2 f_2 \, d\tau$$

Let's define an overlap integral S_{ij}

$$S_{ij} = \int f_i f_j \, d\tau$$

Since the functions f_i and f_j are not by definition eigenfunctions of an operator, they are not necessarily orthonormal and thus we can write:

$$\int \Phi^* \Phi \, d\tau = c_1^2 S_{11} + 2c_1 c_2 S_{12} + c_2^2 S_{22}$$

This gives us

$$E_\varphi(c_1, c_2) = \frac{c_1^2 H_{11} + 2c_1 c_2 H_{12} + c_2^2 H_{22}}{c_1^2 S_{11} + 2c_1 c_2 S_{12} + c_2^2 S_{22}}$$

We now want to differentiate this with respect to c_1 and c_2. Before we do this let's write this expression in the following form

$$E_\varphi(c_1, c_2)\left(c_1^2 S_{11} + 2c_1 c_2 S_{12} + c_2^2 S_{22}\right) = c_1^2 H_{11} + 2c_1 c_2 H_{12} + c_2^2 H_{22}$$

Now let us differentiate it with respect to c_1. We need to use the chain rule on the left side:

$$\frac{\partial E_\varphi(c_1, c_2)}{\partial c_1}\left(c_1^2 S_{11} + 2c_1 c_2 S_{12} + c_2^2 S_{22}\right) + E_\varphi(c_1, c_2)\left(2c_1 S_{11} + 2c_2 S_{12}\right) = 2c_1 H_{11} + 2c_2 H_{12}$$

To find the minimum we set the derivative to zero:

$$\frac{\partial E_\varphi(c_1,c_2)}{\partial c_1}=0$$

This gives

$$E_\varphi(c_1,c_2)\left(2c_1S_{11}+2c_2S_{12}\right)=2c_1H_{11}+2c_2H_{12}$$

and consequently we get:

$$c_1\left(H_{11}-E_\varphi(c_1,c_2)S_{11}\right)+c_2\left(H_{12}-E_\varphi(c_1,c_2)S_{12}\right)=0$$

Similarly, if we differentiate E with respect to c_2 and set the derivative equal to zero we get

$$c_1\left(H_{12}-E_\varphi(c_1,c_2)S_{12}\right)+c_2\left(H_{22}-E_\varphi(c_1,c_2)S_{22}\right)=0$$

These two equations constitute a pair of linear algebraic equations for the variables c_1 and c_2.

You should know from linear algebra that this set of equations will only have a non-trivial solution if

$$\begin{vmatrix} H_{11}-E_\varphi S_{11} & H_{12}-E_\varphi S_{12} \\ H_{12}-E_\varphi S_{12} & H_{22}-E_\varphi S_{22} \end{vmatrix}=0$$

This determinant is called a ***Secular Determinant.***

When we multiply out the determinant, we will get a quadratic equation in E which is called a ***Secular Equation***. Solving the secular equation will give us two values of E. The lowest one will be the approximation to the ground state energy. The next highest root will be an upper limit to the next energy level.

If one takes the value of E obtained from this procedure and substitutes it back into the set of linear equations, one can then solve for the values of c_1 and c_2 which gives the best estimate of Φ.

This whole procedure can be generalized to the case in which we use N functions rather than just 2 functions. In that case the secular determinant is

$$\begin{vmatrix} H_{11}-E_\varphi S_{11} & H_{12}-E_\varphi S_{12} & \dots & H_{1N}-E_\varphi S_{1N} \\ H_{12}-E_\varphi S_{12} & H_{22}-E_\varphi S_{22} & & \vdots \\ \vdots & & & \vdots \\ H_{1N}-E_\varphi S_{1N} & \cdots & \dots & H_{NN}-E_\varphi S_{NN} \end{vmatrix}=0$$

In this case, when we multiply out the determinant, we get an N^{th} order secular equation with N roots for the energy E. The lowest root is an upper limit to the ground state energy. The estimate of the energy gets better as you add more functions, but you can see that the needed computational power also increases.

If we choose our functions f_n to be an orthonormal set, then all the off-diagonal overlap integrals will equal zero and all the diagonal ones will equal one.

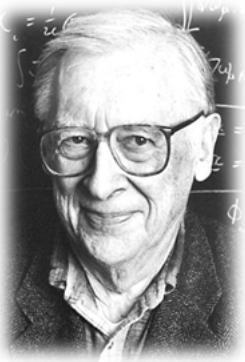
That is: $S_{ij}=\delta_{ij}$

The secular determinant then becomes

$$\begin{vmatrix} H_{11}-E_\varphi & H_{12} & \dots & H_{1N} \\ H_{12} & H_{22}-E_\varphi & & \vdots \\ \vdots & & & \vdots \\ H_{1N} & \cdots & \dots & H_{NN}-E_\varphi \end{vmatrix} = 0$$

This approach is one of the two major techniques used in computational chemistry for solving the electronic Schrödinger equation; we will discuss this later. The key to making it work well is to choose the right trial functions.

John Pople and Walter Kohn won the 1998 Nobel Prize in chemistry for their work developing and using these two different techniques.



John Pople

Walter Kohn

8 Many Electron Atoms

When we come to the point of dealing with the Helium atom, you will find that we cannot solve the problem exactly due to the inter-electronic repulsion term. We must therefore use approximation techniques such as perturbation theory and the variational principle.

The helium atom problem takes on special importance because the methods that are used to treat helium and the concepts that emerge will be applicable to atoms with more electrons.

Before I apply the techniques of perturbation theory and the variational principle to helium, I would like to make a brief digression and introduce a system of units called ***atomic units***. The Schrödinger equation for atoms becomes quite simple if we write it in atomic units.

First, let us write out the helium atom Hamiltonian explicitly (here we explicitly assume that we can neglect the kinetic energy associated with the nuclear motion):

$$\hat{H} = -\frac{\hbar^2}{2m_e}\nabla_1^2 - \frac{\hbar^2}{2m_e}\nabla_2^2 - \frac{Ze^2}{4\pi\varepsilon_0 r_1} - \frac{Ze^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 r_{12}}$$

Now let us choose a set of units such that

$$m_e = 1 \quad , \quad \hbar = 1 \quad , \quad e = 1 \quad , \quad 4\pi\varepsilon_0 = 1$$

If we do this the helium atom Hamiltonian becomes

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}$$

We no longer have to include any physical constants in our Schrödinger Equation. These units are called **atomic units**, and all other physical quantities can be expressed in terms of these four. Certain quantities related to atoms become "natural units".

Quantity	**Natural Unit**	**SI unit equivalent**
Mass	m_e = 1	9.1 x 10^{-31} kg
Charge	$\|e\|$ = 1	1.6 x 10^{-19} C
Angular momentum	$\hbar$=1	1.05 x 10^{-34} Js
Permittivity constant	$4\pi\varepsilon_0$ = 1	1.1126 x 10^{-10} C^2 J^{-1} m^{-1}

These are the four basic units. Others follow from these:

Quantity	**Natural Unit**	**SI unit equivalent**
Length	$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2}$ = 1 (Bohr)	5.3 x 10^{-11} m (about 0.5 Å)
Energy	$\frac{e^2}{4\pi\varepsilon_0 a_0}$ = 1 (Hartree)	4.35 x 10^{-18} J

Note that one Bohr is the radius of the hydrogen atom in its lowest state. Also note that one Hartree is twice the ionization energy of the ground state energy of the hydrogen atom.

8.1 Perturbation Treatment of Helium

We want to solve the Schrödinger equation for helium. If we use perturbation theory, we can write the zeroth-order Hamiltonian as

$$\hat{H}^{(0)} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2}$$

As we have done before, we could further break this up into two one-electron Hamiltonians.

We can write the zeroth-order wave function for the ground state as

$$\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)$$

Where the ψ_{1s} functions are the hydrogen-like 1*s* wave functions for electron 1 and electron 2.

$$\psi_{1s}(\mathbf{r}_j) = \sqrt{\frac{Z^3}{\pi}} e^{-Zr_j}$$

In atomic units, the ground state energy of the hydrogen-like atom is:

$$E_1 = -\frac{Z^2}{2}$$

You can take the formula that I gave you for the hydrogen atom energy levels and put it in atomic units to verify this. The zeroth-order energy can therefore be given by

$$E_1^{(0)} = -\frac{Z^2}{2} - \frac{Z^2}{2} = -Z^2$$

The perturbation part of the Hamiltonian is just

$$\hat{H}' = \frac{1}{r_{12}}$$

so the first order correction to the energy is

$$E_1^{(1)} = H'_{11} = \int\int \psi_1^{(0)*}(\mathbf{r}_1, \mathbf{r}_2)\hat{H}'\psi_1^{(0)*}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2$$

$$= \int\int \psi_{1s}^*(\mathbf{r}_1)\psi_{1s}^*(\mathbf{r}_2)\left(\frac{1}{r_{12}}\right)\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2$$

Note that the integration is over the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Thus each of these integral signs above represents a triple integral, one for each coordinate of each particle (that is $d\mathbf{r}_1$ is $dx_1\, dy_1\, dz_1$ and similarly for $d\mathbf{r}_2$)

Remember, that the first order correction in perturbation theory is simply the average value of the perturbation. Thus, this integral represents the average inter-electronic repulsion (calculated with the zeroth-order wave functions).

I will not take the time to do this integral explicitly here. The result from the integral is

$$E_1^{(1)} = \frac{5}{8}Z$$

So corrected to first order, the energy of the helium atom is

$$E_1 = E_1^{(0)} + E_1^{(1)} = -Z^2 + \frac{5}{8}Z$$

For helium, $Z = 2$, so

$$E_1 = -\frac{11}{4} = -2.75$$

The actual value is -2.9033 au, so this result is about 5% in error.

If we were to take the is to second order we would find E_1 = - 2.91 au

To third order one would get E_1 = -2.9037 au, in excellent agreement with the experimental value.

Recall that there is no restriction that the perturbation theory estimate approaches the true energy as an upper limit. This only holds for the variational method.

8.2 Variational Treatment of Helium

We can also use the variational method to estimate the energy of the ground state of helium. We will take our trial function to be the same as our zeroth-order perturbation theory wave function

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)$$

Where

$$\psi_{1s}(\mathbf{r}_j) = \sqrt{\frac{Z^3}{\pi}}e^{-Zr_j}$$

except that we can treat the nuclear charge Z as a variational parameter.

Thus

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z^3}{\pi}e^{-Z(r_1+r_2)}$$

We now have to evaluate the variational function given by

$$E_\varphi = \int \Phi^*(\mathbf{r}_1, \mathbf{r}_2)\hat{H}\Phi(\mathbf{r}_1, \mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

where we are omitting the normalization integral in the denominator because we chose 1*s* functions which were initially normalized.

The Hamiltonian for the helium atom is the same whether you use perturbation theory or the variational principle, however in the latter, you must include the entire Hamiltonian in the integral rather than just the perturbation part when calculating the integrals.

Recall that the Hamiltonian for helium is

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

We can rewrite this as

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + \frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}}$$

(It should be clear that this is the same as the original.)

If we put the Hamiltonian and the wave functions into the variational function we get

$$E_\varphi(Z) = \int \frac{Z^3}{\pi} e^{-Z(r_1+r_2)} \left(-\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + \frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}} \right) \frac{Z^3}{\pi} e^{-Z(r_1+r_2)}\, d\mathbf{r}_1\, d\mathbf{r}_2$$

We can now simplify this by noting that

$$\left(-\frac{1}{2}\nabla_j^2 - \frac{Z}{r_j} \right) \sqrt{\frac{Z^3}{\pi}} e^{-Zr_j} = -\frac{Z^2}{2} \sqrt{\frac{Z^3}{\pi}} e^{-Zr_j}$$

This is just the Schrödinger Equation for a one-electron atom with charge Z.

We can write one of these equations for electron 1 and another for electron 2. The fact that the wave function is a product of functions for electron 1 and 2 makes no difference, since the electron 1 Laplacian does not operate on the electron two part of the wave function and vice versa. We can therefore simplify the variational function to obtain:

$$E_\varphi(Z) = -\frac{Z^2}{2} - \frac{Z^2}{2} + \int \frac{Z^3}{\pi} e^{-Z(r_1+r_2)} \left(\frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}} \right) \frac{Z^3}{\pi} e^{-Z(r_1+r_2)}\, d\mathbf{r}_1\, d\mathbf{r}_2$$

We can simplify the remaining integral in the following way

$$\int \frac{Z^3}{\pi} e^{-Z(r_1+r_2)} \left(\frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}} \right) \frac{Z^3}{\pi} e^{-Z(r_1+r_2)}\, d\mathbf{r}_1\, d\mathbf{r}_2 = 2\frac{Z^3}{\pi} \int \frac{Z-2}{r_1} e^{-2Zr_1} d\mathbf{r}_1 \int \frac{Z^3}{\pi} e^{-2Zr_2} d\mathbf{r}_2 + \left(\frac{Z^3}{\pi} \right)^2 \int \frac{1}{r_{12}} e^{-2Z(r_1+r_2)} d\mathbf{r}_1 d\mathbf{r}_2$$

(factor of 2 comes from two identical terms.)

Notice that the last integral is the same integral we had to evaluate in first order perturbation theory. The result equaled 5/8 *Z*.

The second integral in the first term is just the normalization integral for a 1*s* wave function and equals 1. We are left with

$$2\frac{Z^3}{\pi} \int \frac{Z-2}{r_1} e^{-2Zr_1} d\mathbf{r}_1 = 2(Z-2)\frac{Z^3}{\pi} \int \frac{1}{r_1} e^{-2Zr_1} d\mathbf{r}_1 = 2(Z-2)Z$$

where the last step was evaluated by looking up the integral.

Combining all the terms in our variational function one gets

$$E_\varphi(Z) = -\frac{Z^2}{2} - \frac{Z^2}{2} + 2Z\left(Z-2\right) + \frac{5}{8}Z$$

$$= Z^2 - \frac{27}{8}Z$$

If we now take the derivative and set it equal to zero (to minimize the energy with respect to our variational parameter *Z*, we find that

$$\frac{\partial E_\varphi(Z)}{\partial Z} = 2Z - \frac{27}{8} = 0$$

Consequently we have

$$Z_{\min} = \frac{27}{16}$$

Putting this back into the expression for $E_\varphi(Z)$ gives

$$E_{\min} = \left(\frac{27}{16}\right)^2 - \frac{27}{8}\frac{27}{16} = -2.848$$

Note that $Z_{\min} < 2$, the charge on the nucleus. This can be interpreted as partial screening of the full nuclear charge by the other electron (that is each electron partially screens the other). This is a physically reasonable result.

At first glance, the value for the energy, -2.848 au, seems to be in excellent agreement with the experimental value, *i.e.*, within 1.8%. One must take a closer look at this agreement, however. While in a relative sense, the percent agreement with the experimental value is good, even a small percentage of the total energy is still a significant amount of energy in an absolute sense, particularly compared to typical chemical bond energies.

The difference equals:

-2.848 - (-2.9033) =0.0553 au

Knowing that

1 au (hartree) = 2625.5 kJ/mol

this difference of 0.0553 au equals 145.19 kJ/mol

This amount of energy is comparable to a weak chemical bond (it is about 1/3 of a typical CH bond energy). Clearly, this will not do if we want to use quantum mechanics for more complicated molecular systems.

Part of the problem is that we restricted ourselves to a trial function, which was a product of one-electron functions. We did this for two reasons:

- This type of function is the zeroth-order function that we used in perturbation theory. Physically this means that it would be the true wave function if there were not any inter-electronic repulsion. We then consider the inter-electronic repulsion as a small perturbation on the system and might not expect the true functions to be much different.

- We do this because the concepts of electron orbitals shape the way we think about chemistry. (This is the way we like to think about chemistry.)

Think back to first year chemistry. We write the electronic configuration for the ground state of an atom like carbon as $\mathbf{1}s^2\mathbf{2}s^2\mathbf{2}p^2$ for example.

The periodic behavior of the elements in the periodic table results from filling electrons into such one-electron orbitals. This is the way the periodic table is organized. What are we doing when we write that?

Basically we are saying that the wave function for the carbon atom is the product of 2 1*s* wave functions, 2 2*s* wave functions and 2 2*p* wave functions.
The whole concept of electrons going into orbitals assumes the wavefunction is a product of one electron wavefunctions.

What we are in effect doing when we write down an electronic configuration such as this is writing down the zeroth-order perturbation theory wave function. Thus, it helps us think about electrons going into individual orbitals. This, in turn, has great predictive power about chemical bonding! This notation, which you may have thought about in first year chemistry as being somewhat mysterious, really has its root in quantum mechanics.

One more related point:
Remember from first year chemistry that the sequence in which orbitals fill as you go to atoms of higher atomic number is not quite the same as the energies of the H atom orbitals.

The order of the H atom wave functions depend only upon the quantum number *n*

So **1*s* < 2*s*, 2*p* < 3*s*, 3*p*, 3*d* < 4*s*, 4*p*, 4*d*, 4*f* < 5*s* . . .**

But remember the order of orbitals in many electron atoms is

1*s* < 2*s* < 2*p* < 3*s* < 3*p* < 4*s* < 3*d* < 4*p* < 5*s* < 4*d* . . .

This is because in writing individual one-electron orbitals, we are neglecting inter-electronic repulsion. We are neglecting shielding of one electron from others.

When we include inter-electronic repulsion, the energies of the different orbitals change since some are screened more than others. Recall the radial distributions that I showed for the hydrogen atom wavefunctions. Some orbitals had little lumps of probability near the nucleus. This influences the amount of screening.

Let us now get back to our discussion of the variational estimate for the ground state of helium. If we don't restrict our functions to products of one electron functions we can get essentially the exact answer for the energy of helium, however, we lose the physical interpretation of one electron orbitals which so much shapes our thinking about chemistry.

However, even if thinking about wavefunctions as products as one-electron functions is only approximate, it has great predictive power and thus is practical. There are ways that we can keep this useful picture of electrons in individual orbitals and get a better estimate of the energy. One way to do this is to introduce one-electron orbitals that are more general functions than the hydrogen atom one-electron functions. One set of such orbitals are called **Slater Orbitals** since they were introduced by John Slater. They have the form

$$S_{nlm}(r,\vartheta,\varphi) = N_{nl} r^{n-1} e^{-\xi r} Y_l^m(\vartheta,\varphi)$$

I will not discuss these in detail, but the angular parts are spherical harmonics and the radial parts are similar to the H atom eigenfunctions. The parameter ξ is taken to be arbitrary and is not necessarily equal to *Z*/*n* as in the hydrogen atom wave functions.

Even if we have an extremely flexible one-electron function, there is still a theoretical limit to the accuracy that one can achieve in a variational calculation if one assumes the form of the wave function to be a product of one-electron functions.

A procedure called the ***Hartree-Fock Self-Consistent Field method*** gives a procedure by which we can calculate the best wave function which is restricted to a product of one electron functions. In another words, if we are going to restrict ourselves to one-electron functions, a Hartree-Fock calculation is the best we can do.

8.3 Hartree-Fock SCF Method

I will first explain the application of this approach for Helium. We will generalize it later.

The Hartree-Fock approach starts by writing the wavefunction as a product of one-electron functions or one-electron orbitals.

$$\psi(\mathbf{r}_1,\mathbf{r}_2)=\varphi(\mathbf{r}_1)\varphi(\mathbf{r}_2)$$

One typically chooses these one electron functions to be hydrogenlike orbitals but with somewhat more flexibility. In practice one chooses a linear combination of Slater type orbitals for each function. No matter what the functional form of these orbitals, they retain their identity as one-electron functions.

The basic physics of the Hartree-Fock method is that it takes into consideration the inter-electronic repulsion in an average way.

Consider the fact that the potential between two point charges is given by

$$U(r)=\frac{q_1q_2}{4\pi\varepsilon_0 r}=\frac{e^2}{4\pi\varepsilon_0 r}$$

Rather than considering the inter-electronic repulsion as an interaction between two point charges, consider it as the interaction of a single point charge with a continuous charge distribution. This approach makes sense inasmuch as the quantum mechanical wave function for an electron gives us its probability distribution.

The potential energy for the interaction of a point charge with some continuous charge distribution is

$$dU(r_{12})=\frac{q_1\rho_2}{4\pi\varepsilon_0 r_{12}}d\mathbf{r}_2$$

where ρ_2 is the charge per unit volume for electron 2 and dr_2 is the volume element. $\rho_2\, dr_2$ is the differential charge, dq_2

To get the full potential one must integrate over the volume element $d\mathbf{r}_2$.

$$U(r_{12})=\int\frac{q_1\rho_2}{4\pi\varepsilon_0 r_{12}}d\mathbf{r}_2$$

However, we know the probability distributions from the one-electron wave functions.

If $\varphi(\mathbf{r}_2)$ is the orbital (wave function) for electron 2, then the probability distribution for electron 2 is given by

$$\rho_2 d\mathbf{r}_2=-e\varphi^*(\mathbf{r}_2)\varphi(\mathbf{r}_2)d\mathbf{r}_2$$

where $d\mathbf{r}_2$ is the volume element of electron.

Since this is a probability distribution for an electron we can also interpret this as a charge distribution for electron 2. We can therefore write the potential energy that electron 1 experiences from its interaction with electron 2 as

$$U_1^{eff}(\mathbf{r}_1) = \int \varphi^*(\mathbf{r}_2)\frac{1}{r_{12}}\varphi(\mathbf{r}_2)d\mathbf{r}_2$$ (Note: the change to atomic units)

We can therefore write an effective one electron potential

$$\hat{H}_1^{eff}(\mathbf{r}_1) = -\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} + U_1^{eff}(\mathbf{r}_1)$$

The Schrödinger equation corresponding to this effective Hamiltonian is

$$H_1^{eff}(\mathbf{r}_1)\varphi(\mathbf{r}_1) = \varepsilon_1\varphi(\mathbf{r}_1)$$

where $\varphi(\mathbf{r}_1)$ is a one electron orbital and ε_1 is the orbital energy.

This is called the ***Hartree-Fock Equation*** for helium.

One can use the variational principle to solve this equation and get the best values of the parameters in $\varphi(\mathbf{r}_1)$ that will minimize the energy ε_1.

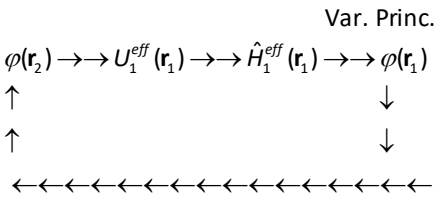
Here one is applying the variational principle to a one-electron problem!! That is, one calculates the variational function, $E_{\text{Ø}}$, using the one electron functions and the *effective* one-electron Hamiltonian. One then minimizes the variational function by varying the parameters in the trial functions. This gives an estimate of the energy ε_1, and a new wavefunction (new variable parameters in the linear combination of Slater orbitals).

Let me first talk first about these wave functions and then about the orbital energy ε_1.

Recall that we started with a wave function $\psi(\mathbf{r}_1,\mathbf{r}_2) = \varphi(\mathbf{r}_1)\varphi(\mathbf{r}_2)$

The one-electron functions $\varphi(\mathbf{r}_1)$ and $\varphi(\mathbf{r}_2)$ have the same functional form and have one or several variable parameters. We then calculated the average effective potential $U_1^{eff}(\mathbf{r}_1)$ that electron 1 experiences using the function $\varphi(\mathbf{r}_2)$. That in turn gives us an effective one-electron Hamiltonian $\hat{H}_1^{eff}(\mathbf{r}_1)$ which in turn gives us a new function $\varphi(\mathbf{r}_1)$ and thus also $\varphi(\mathbf{r}_2)$ as the parameters in both $\varphi(\mathbf{r}_1)$ and $\varphi(\mathbf{r}_2)$ vary in the same way. We then use this $\varphi(\mathbf{r}_2)$ to calculate a new effective potential and then another effective one-electron Hamiltonian which gives us another function $\varphi(\mathbf{r}_2)$.

Schematically we can write this



One continues in this cyclic process until the functions $\varphi(\mathbf{r})$ no longer change. When that is the case, the resulting orbitals are called ***Hartree-Fock Orbitals.***

This is why the procedure is called the ***Hartree-Fock Self Consistent Field Method*** because the wave functions which are calculated are self consistent - *i.e.* the average field one would calculate from them will produce a Hamiltonian which will yield a wave function which is the same as the original.

This procedure is fairly simple to visualize for helium, since although there are two electrons and two one-electron wave functions, the functions are the same. Remember from first year chemistry that one can put two electrons in each orbital. For many electron atoms, one can still have a wave function that is the product of one-electron functions, but there must be one distinct function for every two electrons.
In the case of many electrons, one calculates the average potential experienced by electron 1 from all the other electrons in the atom. One then constructs an effective one electron Hamiltonian and solves for a new one-electron function for electron 1. One then moves on to electron 2. In calculating the average potential for electron 2, one uses the new function for electron 1. One then gets a new function for electron 2. You move down the line repeating this for each of the electrons. When one is done you go back to the beginning and start all over again. You continue to iterate until the functions change no more.

Let us get back to the case of Helium. When you have found the optimum Hartree-Fock Orbitals, you find the energy by taking its expectation (average) value using the full Hamiltonian and the full wave function (which is a product of the optimized Hartree-Fock Orbitals)

$$E = \int\int \varphi^*(\mathbf{r}_1)\varphi^*(\mathbf{r}_2)\hat{H}\varphi(\mathbf{r}_1)\varphi(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

where you may recall that for helium,

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}$$

If we substitute this into the variational function we get

$$E = I_1 + I_2 + J_{12}$$

where

$$I_j = \int \varphi^*(\mathbf{r}_j)\left[-\frac{1}{2}\nabla_j^2 - \frac{Z}{r_j}\right]\varphi(\mathbf{r}_j)d\mathbf{r}_j$$

Note that this is just the average energy of an electron in the Hartree-Fock orbitals neglecting inter-electronic repulsion.

and

$$J_{12} = \int\int \varphi^*(\mathbf{r}_1)\varphi^*(\mathbf{r}_2)\left[\frac{1}{r_{12}}\right]\varphi(\mathbf{r}_1)\varphi(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

which is called the coulomb integral. This looks like the average value of the inter-electronic repulsion in the Hartree-Fock orbitals.

Note that in this determination of the total energy, one simply uses the optimized Hartree-Fock orbitals and does not vary their parameters. The optimization has already been done.

It turns out that the total energy of the Helium atom is not simply the sum of the orbital energies. We can show this if we go back to the one-electron Schrödinger equation that we had written

$$H_1^{eff}(\mathbf{r}_1)\varphi(\mathbf{r}_1) = \varepsilon_1\varphi(\mathbf{r}_1)$$

If we multiply this on both sides by $\varphi^*(\mathbf{r}_1)$ and integrate we get

$$\varepsilon_1 = \int \varphi^*(\mathbf{r}_1) H_1^{eff}(\mathbf{r}_1) \varphi(\mathbf{r}_1)\, d\mathbf{r}_1$$

If we were to substitute in our expression for the effective one-electron Hamiltonian we could easily show that

$$\varepsilon_1 = I_1 + J_{12}$$

If we then take the sum of the orbital energies we get

$$\varepsilon_1 + \varepsilon_2 = I_1 + J_{12} + I_2 + J_{12} \neq E \qquad \text{(this counts } J_{12} \text{ twice)}$$

However, if we compare the orbital energy to the expression we had for *E* before we can see that

$$\varepsilon_1 = E - I_2$$

If we look at the expression for the integral I_2 we had earlier, you can see that it simply looks like the average energy of a one-electron atom with charge *Z* (in this case *Z*=2 for helium) calculated with the Hartree-Fock orbital. This is just an approximation to the energy of a helium ion. ε_1 is then the difference between the energy of the helium atom and the energy of the helium ion. That is it the energy needed to remove an electron from that particular orbital. This is what we know as the ***Ionization energy***.

Thus $IE \cong -\varepsilon_1$

This is called **Koopman's Theorem** and can be (and has been) verified experimentally using spectroscopy.

If we were to go through all the mechanics of the Hartree-Fock Procedure and calculate the energy of helium we would get

E = -2.8617 au.

This is to be compared to the exact energy

E = -2.9037 au (experimental).

Remember, this procedure yields the best estimate of the energy in the approximation that we can represent the wave function as a product of one-electron functions. The closeness of this result to the exact energy tells us that it is ok to think of electrons in separate orbitals to a fairly high degree of approximation. While this result is almost within 1% of the exact energy, as we stated earlier, even a small percentage error can be a big absolute error.

The question is why isn't this better. What have we left out which might contribute to the energy?Well, in the basic approach of the Hartree-Fock method, the electrons are assumed to be independent of each other, interacting through some average or effective potential. This is the assumption that the motion of the electrons is uncorrelated. We know in reality, the motion of the electrons must be correlated, that is the motion of one electron will effect that of the other.

We can define a **correlation energy** as

$$E_{corr} = E_{exact} - E_{HF}$$

Although Hartree-Fock gets almost 99% of the exact energy, the difference is about 100 kJ/mole, which is comparable to chemical bond energies.

The calculation of correlation energies and the inclusion of electron correlations into the wave functions is an area of active interest.

Let us now move on to Lithium, which has three electrons. If we follow along the lines that we have been proceeding, it might seem *natural* to start by writing the wave function as a product of 1s orbitals

$$\psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\psi_{1s}(\mathbf{r}_3)$$

However, most of you probably recall from first year chemistry that you cannot put 3 electrons in a 1*s* orbital. To understand why, I need to discuss the subject of Electron spin and Pauli Principle.

Tjalling Koopmans

9 Electron Spin and the Pauli Principle

I will digress and talk about electron spin in general--how it was discovered and what some of its properties are. Then I will go back and treat the lithium atom and see what role electron spin plays.

9.1 Electron Spin

Recall that at the end of our treatment of the hydrogen atom we discussed the Zeeman Effect. If you remember, the hydrogen atom can have a magnetic moment due to the orbital motion of the electron

$$\boldsymbol{\mu} = -\frac{|e|}{2m_e}\mathbf{L}$$

In our discussion of the Zeeman effect, I explained that this magnetic moment would interact with an external magnetic field **B** and add a term to the energy

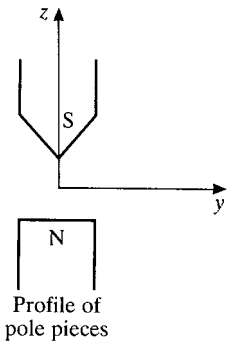
$$U = -\boldsymbol{\mu}\cdot\mathbf{B} = \frac{|e|}{2m_e}\mathbf{L}\cdot\mathbf{B}$$

If we take the magnetic field to be along the *z* axis **B** = B_z then the interaction is

$$U = \frac{|e|B_z}{2m_e}L_z$$

We know that the eigenvalues of the $\hat{L}_z$ operator are $m\hbar$, so the energy levels are split according to their value of the *m* quantum number. (Remember the *m* quantum number represents the projection of **L** on a the space fixed *z*-axis. We know that there are 2*l*+1 values of *m* for every *l*, so a given level is split into 2*l*+1 levels. *l* can take on the values *l* = 0, 1, 2, .)

The magnetic field in this case was assumed to be homogeneous (the same everywhere so each atom in a sample experiences the same field). However, let us consider the case in which we have two magnetic plates shaped something like this.



Let us say that we pass a beam of neutral atoms through such an inhomogeneous magnetic field. Since the magnetic dipole moment is proportional to $\hat{L}_z$ and its eigenvalues are quantized we can write:

$$U = \frac{|e|B_z}{2m_e}m\hbar$$

Using the definition of the force

$$\mathbf{F} = -\nabla U$$

We find that the exerted force in the *z*-direction is given by:

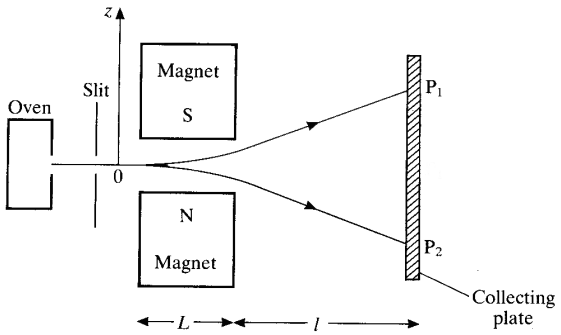
$$F_z = -\frac{\partial U}{\partial z} = -\frac{|e|m\hbar}{2m_e}\frac{\partial B_z}{\partial z}$$

One thus expect that the force on the atoms will split the one beam up into several discrete beams of atoms depending on their *m* quantum number. The one beam should split into 2*l*+1 beams, since that is how many *m* states there are.

In 1922, Otto Stern and Walter Gerlach performed an experiment which was designed to do just that. In the so-called ***Stern-Gerlach experiment*** they used a beam of silver atoms from an oven beam source and sent it through an inhomogeneous magnetic field like the one I have shown schematically, with the field oriented in the *z*-direction.

They expected it to split states with different projections of the angular momentum in the *z*-direction, that is different *m* states.

What they observed was the following:



They observed that the original beam of silver atoms only split into <u>*two*</u> beams.

To be able to split into an even number of *m* states, *l* must be half integral, and for there to be exactly two states, *m* must be ± 1/2, since *m* ranges from + *l* to -*l*. The Stern-Gerlach experiment therefore suggests the *existence of half-integral values of the angular momentum.* Recall that our treatment of orbital angular momentum indicated that the only allowable values of *l* were integral values.

[Otto Stern]

[Walther Gerlach]

From these experiments Stern and Gerlach postulated the existence of another type of angular momentum, called ***spin angular momentum***, that could take on half-integral values.

Note that a silver atom has many electrons, so this experiment did not prove that the spin angular momentum of an electron is 1/2. However, it did prove that it must involve half-integral values. A collection of particles each with angular momentum will have some overall angular momentum that is the vector sum of the individual values. There is no way to get a net spin of 1/2 if at least one of the particles had a spin of half-integral value.

Within the quantum mechanical framework we have developed in this course, the existence of spin angular momentum cannot be explained without the introduction of additional hypotheses.

Note: A more sophisticated treatment of quantum mechanics that takes into consideration relativistic effects was developed by Dirac in the early nineteen thirties. From this treatment the concept of spin angular momentum arises naturally. However such a treatment is beyond the scope of this course.

Wolfgang Pauli introduced a way in which spin angular momentum could be incorporated into a non-relativistic treatment of quantum mechanics by taking a few additional postulates. We will follow his approach.

The additional postulates we must include are:

1. ***The spin operator is an angular momentum operator.***

   This may sound trivial, but we are making assumptions based on the results of experiments which deflect atoms in magnetic fields. Such observations are consistent with this postulate but don't prove it. Hence it must be taken as a postulate.

   The definition of what is and isn't an angular momentum is based on the commutator relations. The definition of angular momentum is an operator $\hat{S}$ whose components satisfy the relations

   $$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z$$

   $$\left[\hat{S}_x, \hat{S}^2\right] = 0$$

   and the cyclic permutations of this commutator.
   This implies that $\hat{S}^2$ and $\hat{S}_z$ commute and have a common set of eigenfunctions. However, the spin angular momentum in contrast to the orbital angular momentum cannot be represented by the position and momentum operators. It is therefore not possible to write down the spin eigenfunctions in the usual way. In order to represent these functions one has to use the concepts developed by Heisenberg (in the beginning of this course I already mentioned that quantum mechanics can be represented in two different but equivalent ways)

   In the bra-ket notation as introduced by Dirac the orbital angular eigenfunctions are represented by:

   $$|l, m\rangle$$

   which is equivalent to the Schrödinger eigenfunction:

   $$Y_l^m(\theta, \varphi) = N_{nl} P_l^{|m|}(\cos\theta) e^{im\varphi}$$

   The eigenvalue problem as we have seen before

$$\hat{L}^2 Y_l^m(\theta,\varphi) = l(l+1)\hbar^2 Y_l^m(\theta,\varphi)$$

$$\hat{L}_z Y_l^m(\theta,\varphi) = m\hbar Y_l^m(\theta,\varphi)$$

is in Dirac's notation written as

$$\hat{L}^2 |l,m\rangle = l(l+1)\hbar^2 |l,m\rangle$$

$$\hat{L}_z |l,m\rangle = m\hbar |l,m\rangle$$

One could show that the eigenvalues simply depend on the commutation relations and hence will hold true for any angular momentum operator.

In the case for spin angular momentum one can therefore write:

$$\hat{S}^2 |s,m_s\rangle = s(s+1)\hbar^2 |s,m_s\rangle$$

$$\hat{S}_z |s,m_s\rangle = m_s\hbar |s,m_s\rangle$$

As was the case of orbital angular momentum, the value of m_s can range from *-s* to *+s*. However, the allowable values of the quantum number *s* are different than those of orbital angular momentum.

2. A given particle has a unique value of the quantum number s, and is said to have a spin S. The electron is a spin 1/2 particle (s = 1/2).

We therefore have a fixed spin quantum number. This is different from orbital angular momentum where we can have many *l* values. The projection of the spin angular momentum can still vary, but each particle has a unique value of *s*.

Also, its intrinsic magnetic moment is given by

$$\boldsymbol{\mu}_S = -\frac{|e|}{m_e}\mathbf{S}$$

as compared to

$$\boldsymbol{\mu}_L = -\frac{|e|}{2m_e}\mathbf{L}$$

for the orbital moment.

At the present time the existence of particles with spin ranging from 0, 1/2, 1, 3/2, 2, . . . up to higher values such as 11/2 are known. Protons and neutrons also have half-integral spin.

3. All spin operators commute with all orbital operators.

Thus they depend on different variables. One might think of explaining spin classically as the motion of the electron spinning on its axis which would give rise to an intrinsic angular momentum. To do this, we would have to describe the electron as a solid body with spatial extent and would need 3 more coordinates to describe its orientation (in addition to 3 for it's position)

The theory we are considering postulates that the electron behaves as a point that requires only 3 coordinates to fix its position. Thus spin angular momentum is not derived from any position or momentum variable. We cannot write a classical mechanical expression for this operator.

That is, Spin has no classical analog.

Another way to view this is that for other quantum mechanical quantities that we have dealt with, we know from the Bohr Correspondence principle that as the quantum number gets large, the behavior of a quantum mechanical system approaches that of its corresponding classical system. Because the spin angular momentum of electron is limited to 1/2, this can never happen and the Bohr Correspondence Principle does not apply. Spin is therefore entirely a quantum mechanical quantity.

Spin Eigenfunctions

Because spin operator cannot be written as a function of classical mechanical variables, its eigenfunctions do not depend upon classical mechanical variables. It can therefore be difficult to get a feel for the eigenfunctions of the spin operators.

Since there are only two eigenvalues of $\hat{S}_z$ for an electron, there must be only two eigenfunctions, one for each eigenvalue.

$$\hat{S}^2\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle \qquad \hat{S}^2\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle$$

$$\hat{S}_z\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle = \frac{1}{2}\hbar\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle \qquad \hat{S}_z\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle = -\frac{1}{2}\hbar\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle$$

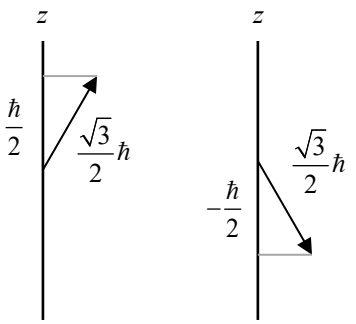
The magnitude of the spin vector is:

$$|S| = \sqrt{s(s+1)}\hbar$$

For $s = \frac{1}{2}$ we find

$$|S| = \frac{\sqrt{3}}{2}\hbar$$

and since the projection of this vector on the *z*-axis can have magnitudes from -*s* to *s*, the only possible values of m_s is ± 1/2 ħ



It can be a little difficult to get a good physical feel for the nature of the eigenfunctions. These functions are discrete functions. They simply represent a spin up or spin down (*i.e*. projection +1/2 or -1/2). In stead of using braket notation the two eigenfunctions of the electron spin are therefore often simplified as spin up, $\alpha = \left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle$, and spin down, $\beta = \left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle$.

Since the spin operators are Hermitian (one could show this from the commutation rules), then the spin eigenfunctions form a complete orthonormal set. In a formal sense we can write the orthonormality as

$$\int \alpha^* \alpha \, d\sigma = \int \beta^* \beta \, d\sigma = 1$$

$$\int \alpha^* \beta \, d\sigma = \int \beta^* \alpha \, d\sigma = 0$$

where σ is called the spin variable and has no classical mechanical analog.

Now that we know about spin eigenfunctions, we should have been using them all along in our discussion of hydrogen and helium.

We postulated that the spin and spatial parts of the wave function are independent. This seems reasonable since the spin eigenfunctions are independent of spatial coordinates. In the Hamiltonian, the spatial operators don't do anything to spin coordinates. In the absence of a field, there are no spin terms in the Hamiltonian (to a high degree of approximation). We can therefore write

$$\psi(x,y,z,\sigma) = \begin{cases} \psi(x,y,z)\,\alpha(\sigma) \\ \psi(x,y,z)\,\beta(\sigma) \end{cases}$$

The complete wave function ψ is called a ***spin-orbital***.

For a hydrogen-like atom, for example, a spin-orbital would be

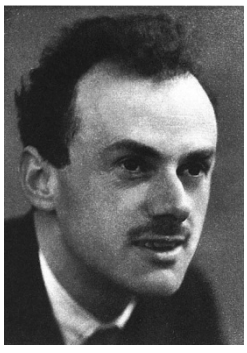
$$\psi_{1,0,0,\frac{1}{2}} = \sqrt{\frac{Z^3}{\pi}} e^{-Zr} \alpha$$

(or the equivalent function with β) in which the forth quantum number in the subscript represents m_s.

Having introduced the concept of electron spin, we are now ready to introduce the ***Pauli Exclusion Principle*.**

You probably remember the Pauli Exclusion Principle in the form of *no two electrons can have all the same quantum numbers*, however it is actually more general than that. I will introduce it in a general form from which the usual statement of the principle results.

[Wolfgang Pauli]



[Paul Dirac]

9.2 The Pauli Principle

The Pauli Principle arises from considering how to treat identical particles in quantum systems. You will see in a moment that there is an important relationship to spin angular momentum.

In classical mechanics, two indistinguishable particles cause no special problems. If we know the initial conditions of the particles and the forces acting on them, we can follow the trajectories of each one of them and thus tell them apart.

In Quantum Mechanics, the Heisenberg Uncertainty Principle tells us we cannot follow the trajectories of individual particles. Thus if two particles have all the same intrinsic properties (mass, charge, spin), we cannot tell them apart. Thus, the wavefunction cannot distinguish between identical particles. This leads to certain very important restrictions.

The implications of this simple principle are enormous - it is responsible for the periodic behavior of the elements, and hence is at the core of chemistry!!!

Consider a wave function for *N* identical particles. Let the symbol q_1 represent the space *and* spin coordinates of particle 1, q_2 those for particle 2, etc.

We can write our wave function

$$\psi = \psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N)$$

I will define the permutation operator $\hat{P}_{ij}$ as the operator which exchanges all the coordinates of particles *i* and *j*.

$$\hat{P}_{ij}\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N) = \psi(q_1, q_2, \ldots, q_j, \ldots, q_i, \ldots q_N)$$

We need to find the eigenvalues of $\hat{P}_{ij}$, i.e.:

$$\hat{P}_{ij}\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N) = c\,\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N)$$

If we operate twice on our wavefunction,

$$\hat{P}_{ij}\hat{P}_{ij}\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N) = \hat{P}_{ij}\psi(q_1, q_2, \ldots, q_j, \ldots, q_i, \ldots q_N) = \psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N)$$

it leaves the function unchanged.

We find:

$$\hat{P}_{ij}\hat{P}_{ij}\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N) = \hat{P}_{ij}\, c\,\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N) = c^2\,\psi(q_1, q_2, \ldots, q_i, \ldots, q_j, \ldots q_N)$$

Thus $c^2 = 1$

You can see that the eigenvalues of $\hat{P}_{ij}$ will be ±1.

If ψ is an eigenfunction of $\hat{P}_{ij}$ with eigenvalue +1, ψ is unchanged upon interchange of the coordinates of particles *i* and *j*. We call an eigenfunction with eigenvalue of 1 *symmetric* with respect to interchange of particles *i* and *j*.

If ψ is an eigenfunction of $\hat{P}_{12}$ with eigenvalue -1, we call it *antisymmetric* with respect to interchange of particles *i* and *j*.

Thus, a wave function for a system of identical particles must be *symmetric* or *antisymmetric* with respect to interchange of *any two* particles. Since the particles are identical, it can't matter which pair you interchange.

The wavefunction must be symmetric or antisymmetric with respect to any possible interchange of two identical particles.

That is:

$$\psi(q_1,q_2,\ldots,q_j,\ldots,q_i,\ldots q_N) = \pm\psi(q_1,q_2,\ldots,q_i,\ldots,q_j,\ldots q_N)$$

All particles in nature are divided into these two categories:

- Particles whose wave functions are ***symmetric*** with respect to interchange (+ sign) are called ***Bosons***.
- Particles whose wavefunctions are ***antisymmetri*** (- sign) are called ***Fermions***.

Furthermore:

- half integral spin particles (electrons, positrons, neutrons, protons) are Fermions
- integral spin particles (photons, mesons) are Bosons

These are relations form the so called spin-statistics theorem, which can be proven by invoking relativity.

Composite particles behave as their net spin. The ^{3}He isotope is a Fermion, whereas ^{4}He is a boson. Due to this difference the two isotopes behave very different at temperatures close to absolute zero, ^{4}He follows Bose statistics whereas ^{3}He follows Fermi statistics.

Since electrons are Fermions, we have another Fundamental Postulate of Quantum Mechanics. *The wavefunctions of a system of electrons must be antisymmetric with respect to interchange of any two of them.*

This is the PAULI EXCLUSION PRINCIPLE.

One cannot write an antisymmetric wave function for an atom that has more than two electrons in one orbital. This has interesting implications, and gives rise to the periodic behavior of the elements. It turns out that Bosons do not have this restriction. If electrons were Bosons, all electrons could go into the same orbital. *Think how this might change chemistry!!*

In hydrogen, since there is only one electron, we don't need to worry about symmetric or antisymmetric behavior with respect to particle interchange. Since the Hamiltonian is independent of spin, the wave function is just a product of spin and spatial parts.

Now let's go back to Helium and consider the effect of the Pauli Exclusion Principle. (We will then go to lithium). Specifically, let us see why it was ok to neglect spin in our previous treatment.

We can write the ground state helium wave functions as

$$\psi = 1s(1)1s(2)$$

where the 1*s* functions could be Hartree-Fock orbitals for example. The parenthesis denotes which particle (one or two).

To take spin into account, we must multiply the spatial function by a spin eigenfunction. The total wave function must satisfy the Pauli Exclusion Principle and be antisymmetric.

We will use notation such as $\alpha(1)\beta(2)$ which means particle 1 has spin up, and particle 2 has spin down. That is the number in parenthesis refers to the particle, not a spin quantum number.

For our helium atom with spatial wavefunction 1*s*(1)1*s*(2) we can have four possible spin functions:

$$\alpha(1)\,\alpha(2) \qquad \beta(1)\,\beta(2) \qquad \alpha(1)\,\beta(2) \qquad \beta(1)\alpha(2)$$

The first two functions are perfectly valid spin functions, since they don't distinguish between identical particles.

However, the last two violate the principle of indistinguishability. They distinguish between the two electrons. If we apply the permutation operator to these last two, we find they are neither symmetric or antisymmetric.

However, if you take normalized linear combinations of these two you can generate a symmetric and an antisymmetric function.

$$\frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2) \pm \beta(1)\alpha(2)\right]$$

So we have 4 normalized two-electron spin functions.

$$\left.\begin{array}{c} \alpha(1)\alpha(2) \\ \beta(1)\beta(2) \\ \frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2)+\beta(1)\alpha(2)\right] \end{array}\right\} \quad \text{symmetric}$$

$$\frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2)-\beta(1)\alpha(2)\right] \quad \text{antisymmetric}$$

We now want to combine the spatial and spin parts.

Since 1*s*(1)1*s*(2) is symmetric with respect to exchange, we must multiply it by an antisymmetric spin function since the overall wave function must be antisymmetric.

We only have one choice. Thus the zeroth-order wavefunction is

$$\psi = 1s(1)1s(2)\frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2)-\beta(1)\alpha(2)\right]$$

So you can see that like hydrogen, the wave function for helium can be factored into a function of the spatial coordinates times a function of spin coordinates.

To a very high degree of approximation the Hamiltonian for a helium atom (or for hydrogen for that matter) is independent of spin coordinates. In this case, the energy will not be affected if we neglect the spin part of the wave function (because it can be factored in this way).

Leaving off the spin part didn't really change the problem. This is not so when you have 3 electrons as in the case of Lithium. To treat the Lithium atom, our normal procedure would be to construct a Hartree-Fock wave function as a product of one-electron functions. Following in the manner in which we dealt with the ground state of Helium, we would write the spatial part of the wave function as a product of 1*s* single electron functions:

$$\psi = 1s(1)1s(2)1s(3)$$

Let us follow this line of reasoning and see where it goes wrong. Consider spin and the requirements imposed by the Pauli Principle. Since the zeroth-order wavefunction $\psi = 1s(1)1s(2)1s(3)$ is symmetric with respect to electron exchange, we need to find an antisymmetric spin function involving 3 electrons. It turns out that it is easy to construct totally *symmetric* spin functions for 3 particles but impossible to construct *antisymmetric* spin functions.

Consider for a moment how to construct an antisymmetric combination of 3 functions *f*, *g*, and *h*.

With 3 electrons, you get 6 permutations of these functions:

$f(1)g(2)h(3)$ $g(1)f(2)h(3)$ $h(1)g(2)f(3)$ $f(1)h(2)g(3)$ $g(1)h(2)f(3)$ $h(1)f(2)g(3)$

We need to combine these to make an antisymmetric function that does not distinguish between electrons. One could show that a general method for constructing such an antisymmetric function for 3 electrons is to use a determinant that includes all the possible functions

$$\psi = \frac{1}{\sqrt{6}}\begin{vmatrix} f(1) & g(1) & h(1) \\ f(2) & g(2) & h(2) \\ f(3) & g(3) & h(3) \end{vmatrix}$$

Note that the rows contain contributions from the same particle and the columns contain contributions from the same function. (We could have done this for helium).

The fact that this determinant will give us the proper antisymmetric combination of these functions can be easily seen from the properties of determinants. Interchanging two electrons amounts to the interchange of any two rows, and we know the interchange of any two rows of a determinant causes the determinant to be multiplied by -1. That is exactly the property we are looking for.

Recognizing this as the most general way to construct an antisymmetric combination of 3 functions (or *n* functions for that matter), we are ready to see the restrictions due to the Pauli Principle. (We actually did this for helium when we formed the linear combination $\alpha\beta-\beta\alpha$)

The functions *f*, *g*, and *h* may each be either α or β. If we let *f*= α , *g*= β, *h*= α

we get the determinant:

$$\psi = \frac{1}{\sqrt{6}}\begin{vmatrix} \alpha(1) & \beta(1) & \alpha(1) \\ \alpha(2) & \beta(2) & \alpha(2) \\ \alpha(3) & \beta(3) & \alpha(3) \end{vmatrix}$$

However, we can recognize that this determinant equals zero, since it has two columns that are the same.

The problem arises because we are trying to get an antisymmetric spin function for three electrons using only two spin functions. The determinant above will always vanish in cases with more than two electrons. This is why we didn't run into problems in out treatment of Helium.

In the case of helium, we were able to take the wavefunctions as independent functions of spatial coordinates and spin coordinates. However, for more than two electrons, this approach fails. Instead, we must consider each zeroth-order function as a combination of spatial and spin variables and then take linear combinations of these functions to get the proper symmetry.

For instance, we could take as our function

$$f(1) = 1s(1)\alpha(1)$$

A function such as this which is a product of a one electron spatial orbital and a one electron spin function is called a *spin-orbital.* Although we had only two different spin functions we can construct many different spin-orbitals by using different spatial parts. We can now construct our determinant to find the proper antisymmetric combination of such functions.

We can see, that if we let our function $g(1) = 1s(1)\alpha(1)$, then the determinant will equal zero. This is where the implications of the Pauli Exclusion Principle become important. The requirement that the total wave function for a system of electrons be antisymmetric is what has led us to using such a determinant in finding the proper wave function.

The properties of the determinant that gives us the proper symmetry behavior (mandated by the Pauli exclusion principle) indicate that *NO TWO ELECTRONS CAN OCCUPY THE SAME SPIN ORBITAL i.e.* have the same quantum numbers. If they do, it will violate the requirement that the total wavefunction be antisymmetric.

The proper determinant for a three electron system is then

$$\psi = \frac{1}{\sqrt{6}}\begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) & 2s(1)\alpha(1) \\ 1s(2)\alpha(2) & 1s(2)\beta(2) & 2s(2)\alpha(2) \\ 1s(3)\alpha(3) & 1s(3)\beta(3) & 2s(3)\alpha(3) \end{vmatrix}$$

Note that if 3rd column were to contain a 1*s* orbital, the determinant would equal zero.

If you work out this determinant, you will see that it is not a simple product of space and spin parts, but is a combination of terms, each of which is a product of space and spin parts. (NOTE: Such determinants of spin-orbital functions are called *Slater Determinants.* These determinants can be written for an *n*-electron atom, in which case the coefficient in front is $\frac{1}{\sqrt{n!}}$)

Two properties of determinants make these wave functions satisfy the Pauli Principle. Interchanging two rows multiplies the determinant by -1 and thus gives a properly antisymmetric wave function. A determinant with two identical rows will equal zero, thus fulfilling the requirement that no electrons occupy the same spin-orbital.

So the Pauli Exclusion principle has required us to put the third electron in a 2*s* orbital, making the ground state configuration:

$1s^2 2s^1$

Note that we could have equally well used the β function in the last column of our determinant. Thus this state is twofold degenerate:

1s	2s		or		1s	2s
↑↓	↑				↑↓	↓

So you can see that the Pauli Exclusion Principle ultimately requires that no two electrons have the same four quantum numbers.

We can trace this requirement back to the fact the electron is a spin 1/2 particle (*i.e*. a Fermion). If it were a Boson, the world would be a different place. If it were a spin 3/2 particle, *the world would be a different place as well.*

I would like to say a few brief words about Hartree-Fock calculations for atoms with more than two electrons. When we discussed the Hartree-Fock method a few lectures ago, we used helium as an example, although I commented on how one would treat a many electron system. Now that we have considered the topic of electron spin, there is just one point I would like to add to the previous discussion.

Recall that the basic Hartree-Fock approach was to start with a wave function that is a product of one electron functions and then calculate the average inter-electronic repulsion by using the one electron functions to calculate a probability density for the individual electrons. This average interaction potential was then used to

form a one electron Hamiltonian which was used to refine the one electron wave functions. This was done iteratively until the functions no longer change.

Our previous treatment neglected spin, and this was ok since the wave functions for a two-electron atom factor into a spatial part and a spin part and because the Hamiltonian is, to a high degree of approximation, independent of spin. Our treatment of spin and the Pauli Principle simply tells us that when we choose our one-electron functions for the Hartree-Fock procedure for atoms of more than two electrons, we need to use Slater determinants of spin-orbitals to insure the proper symmetry and indistinguishability.

Also, when we generalize the Hatree-Fock procedure to more than two electrons, we get a slightly different expression for the energy than we had for Helium. (Our result was a special case of this more general result). The approach is still the same: find the optimized orbitals and then use the variational principle to find the energy.

When we assign electrons in atoms into individual spin orbitals, we are assigning them each a set of quantum numbers, particularly spin and orbital angular momentum quantum numbers. A particular electronic configuration can have a variety of states with different energies depending on how the individual angular momenta couple together. A ***term symbol*** is a designation that indicates the total, orbital, and spin angular momenta for the whole atomic system.

10 Atomic Term Symbols and Coupling of Angular Momentum

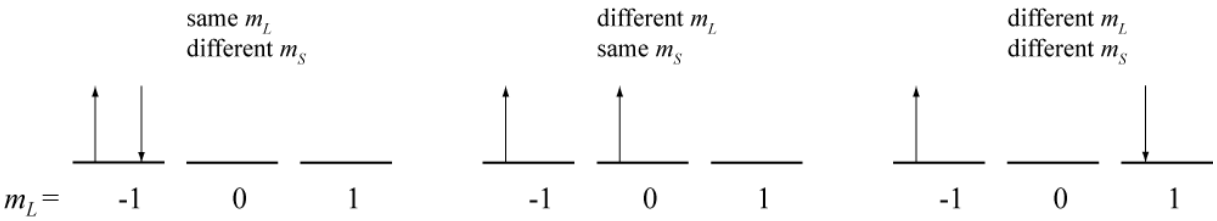
In many electron atoms, each electron has both a spin and an orbital angular momentum. In the zeroth-order picture where we neglect inter-electronic repulsion, the energy of an electron in an atom (*i.e.*, the energy of a particular orbital) depends only on the *n* quantum number of that electron or orbital. In this approximation, a particular electronic configuration (in which we specify only *n* and *l* for each electron) has a specific, well-defined energy.

However, when we specify the electronic configuration of an atom, there are many different values of the projections of both the spin and orbital angular momenta (m_l and m_s) for a given *n* and *l*. (Remember, for hydrogen, the energy is independent of *m*.)

Let's consider the example of the ground state of carbon, which has an electronic configuration of

$$1s^2 2s^2 2p^2$$

The two 2*p* electrons can be in any of the three 2*p* orbitals ($2p_{-1}$, $2p_0$, $2p_{+1}$) and have different m_L and can have either spin up or spin down (different m_S).



When we take into consideration inter-electronic repulsion, the energy of the atoms depends upon these quantum numbers as well *i.e.,* the inter-electronic repulsion lifts the degeneracy.

We therefore need a way to label the states of an atom in a more precise way in order to account for these differences. We need to label a state not only by its configuration, but also in a manner that indicates the projections of the angular momenta of the individual electrons, because each state labeled in this way will have a different net angular momentum and a different energy.

There is no analytical way to differentiate these states, since the electronic Schrödinger equation is not exactly solvable. However, having some qualitative principles will allow us to predict which states will be higher in energy than others. This is done using atomic term symbols.

The approach that we will consider is called ***Russell-Sanders coupling*** or ***L-S coupling*** and is simply a method by which one determines the resultant angular momenta from many electrons.

The Russell-Saunders scheme determines the total orbital angular momentum, **L**, and the total spin angular momentum, **S**, of all the electrons, and then adds these two vectors to get a total electronic angular momentum, **J.**

The result of coupling these angular momenta in this way (*i.e*., the resulting state) is designated by an atomic term symbol.

The term symbol has the form: $^{2S+1}L_J$

where : *L* is the total orbital angular momentum quantum number
S is the total spin quantum number
J is the total angular momentum quantum number

We will see that the total orbital angular momentum quantum number *L* will be an integer ≥ 0.

In a manner analogous to the way we indicate the orbital angular momenta of individual electrons by letters (*s*,*p*,*d*,*f*...), we indicate the total *L* by letters

L =	0	1	2	3	4
	S	*P*	*D*	*F*	*G*

The total *S* will turn out to be integral or half-integral, so the left superscript 2*S*+1 will be an integer.

The quantity 2*S* +1 is called the multiplicity, since if $L \geq S$, there are 2*S* +1 possible values of *J*.

Examples of atomic term symbols (otherwise just called terms) are

^{3}S ^{2}D ^{1}P

These symbols represent states of an atom with the same electronic configuration but with different energy.

These are read as "triplet *S*", "doublet *D*", and "singlet *P*".

At this point it may seem a bit mysterious why we chose to add the angular momenta in this way (that is, why we chose first to find the total *L* and total *S* and then combine them to find total *J*).

Let me digress a moment to talk about good quantum numbers and constants of motion. We know that the square of the magnitude of the total angular momentum operator, $\hat{J}^2$, commutes with the Hamiltonian (if the potential only depends on the distance not the angles) , it is a constant of the motion. This is true no matter how we couple angular momenta. (For a quantity to be constant in time it must commute with $\hat{H}$ so its eigenfunctions have trivial time dependence.)

If we neglect inter-electronic repulsion, the angular momentum operator of each electron commutes with the Hamiltonian and hence is a conserved quantity or a constant of motion. This means that the angular momentum quantum numbers of each electron have meaning since they are time independent.

However, if you include inter-electronic repulsion in a multi-electron system, the orbital angular momentum operators of the individual electrons no longer commute with $\hat{H}$. Therefore the quantum numbers l_i for the individual electrons are no longer good quantum numbers and the angular momenta of the individual electrons need not be conserved (*i.e.* they are no longer constants of the motion).

However, it turns out that for atoms of atomic number less than about 40, the sum of the orbital angular momentum

$$\mathbf{L} = \sum_i l_i$$

will still be a constant of the motion. This says that the individual orbital angular momentum vectors can change in time, but their vector sum remains constant.

The same holds true for spin angular momentum. The total spin angular momentum

$$\mathbf{S} = \sum_i s_i$$

will also be a constant of the motion.

Why does the coupling scheme outlined above work for atoms with $Z \leq 40$? The coupling scheme I have presented in which we first couple the orbital angular momentum and the spin together and then couple these to get the total angular momentum only holds in the case in which we neglect relativistic terms in the Hamiltonian. That is, *L* and *S* are only good quantum numbers when we neglect the relativistic spin-orbit coupling term.

The spin-orbit term in the Hamiltonian has the form

$$\hat{H}_{SO} = \sum_j \xi(r_j)\mathbf{l}_i \cdot \mathbf{s}_i$$

and it enters if we do a relativistic treatment of quantum mechanics. It basically arises from the fact that a moving electron creates a magnetic field which has a magnitude that is proportional to the orbital angular momentum, l_j of the electron. This magnetic field interacts with the spin magnetic moment of the electron which is proportional to the spin, s_j

One can easily show that $\hat{L}^2$ and $\hat{S}^2$ (where these are the operators for the total angular momentum) do not commute with this term in the Hamiltonian.

In the Russell-Sanders or *L-S* coupling scheme outlined above we are neglecting the H_{SO} term and assume

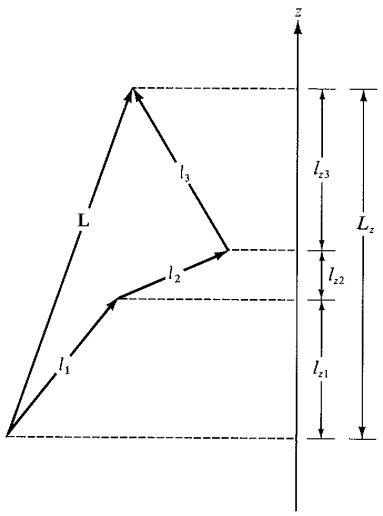
$$\left[\hat{L}^2, \hat{H}\right] = 0 \qquad\qquad \left[\hat{S}^2, \hat{H}\right] = 0$$

This assumption (and hence the coupling scheme) is best when the atomic number is low (<40).

The average velocity of the electron depends on *Z*. Higher atomic number means faster $\langle v \rangle$ and hence, it's more likely that relativistic terms will become important. So in this approximation (which is quite good for low *Z*), we can characterize an atomic state with a total orbital angular momentum quantum number, *L*, where $l(l + 1)\hbar^2$ is the square of the magnitude of **L**.

Because *L* and *S* are constants of the motion, it is meaningful to label states by them.

To get the quantum numbers corresponding to the total orbital or total spin angular momentum, we cannot simply add the individual quantum numbers. The angular momenta are vector quantities, and we have to add them vectorially. However, the projections of the angular momentum vectors on the *z*-axis add like scalar quantities, and from this we can find the quantum numbers for the total angular momentum.



Thus,

$$L_z = \sum_i l_{zi} = \sum_i m_{Li}\hbar = M_L\hbar \qquad \text{where} \qquad M_L = \sum_i m_i$$

and

$$S_z = \sum_i s_{zi} = \sum_i m_{Si}\hbar = M_S\hbar \qquad \text{where} \qquad M_s = \sum_i m_{si}$$

Note: I will use M_L for M to emphasize that it is orbital angular momentum.

The goal is to couple a set of angular momentum vectors by adding the *z*-components to get a total *z*-component (*i.e*. M_L and M_S), and then determine what the corresponding total angular momentum quantum numbers are (*L* and *S*). From these we can determine the Term Symbol. Each term symbol will represent a state of different energy.

We need to be able to derive the *Terms* for a given electronic configuration.

1) First consider closed subshells (full number of electrons)

The total spin will always equal zero for a closed shell. (note that shells are defined by *n*, subshells by *l*)

For each electron with m_s = 1/2 there is another with m_s = -1/2

So

$$M_s = \sum_i m_{si} = 0 \qquad \text{for a filled subshell and } S=0$$

Similarly, the total orbital angular momentum *L* will = 0 for a closed subshell

For every electron in an orbital with quantum number *m*, there will be one with quantum number *-m*, since *m* runs from *l* to *-l.*

$$M_L = \sum_i m_{Li} = 0 \qquad \text{for a filled subshell.}$$

So for a totally closed subshell

$L = 0$ $S = 0$ and we have only a $^1S_0$

If we have a configuration with a closed subshell and an open one, we need only consider the open subshells. (The closed ones don't contribute to *L* and *S*).

2) Consider two electrons in different subshells

One doesn't have to worry about the Pauli principle since they will have different values of *n* or *l* or both. (Remember, no two electrons can have all the same quantum numbers.)

Ignore closed subshells.

Write down all the possible combinations of quantum numbers of the open shell electrons. By summing the *m* and m_s to get M_L and M_S you can then find the possible values of *S* and *L* and the terms.

3) When we have two electrons in the same subshell (same as *n* and *l*), we have some restrictions.

For instance for a $1s^2\ 2s^2\ 2p^2$ carbon ground state, one can make a list of possible values of the *m* quantum numbers, see below. Certain values are not valid, however, because of the Pauli Exclusion Principle.

For instance, one can't have

$m_1 = 1 \quad m_{s1} = 1/2 \qquad\qquad m_2 = 1 \quad m_{s2} = 1/2$

Or you can't have both

$m_1 = 1 \quad m_{s1} = 1/2 \qquad\qquad m_2 = 1 \quad m_{s2} = -1/2$

and

$m_1 = 1 \quad m_{s1} = -1/2 \qquad\qquad m_2 = 1 \quad m_{s2} = 1/2$

If you allowed both of these, it implies that you can distinguish between electrons. By summing the m and m_S to get M_L and M_S you can then find the possible values of S and L and the terms.

There is one last part to these term symbols: the subscript J which is the total electronic angular momentum. The total angular momentum **J** is the vector sum **J** = **L** + **S**. We know this commutes with H so J will be a good quantum number

J can take the values

$$J = L + S, L + S\text{-}1, \ldots |L\text{-}S|$$

You can see this by coupling just two angular momentum vectors.

J is written as a subscript in the term symbol.

For ^3P $L = 1, S = 1$ J can range from $J = 2, 1, 0$

so we get $^3P_0\ ^3P_1\ ^3P_2$ These are called levels.

In the absence of spin-oribit interaction, different levels have the same energy, but $\hat{H}_{SO}$ splits them slightly. Once we have determined the states which correspond to a particular electronic configuration, how do we decide which is lowest in energy?

Hund's Rules

1) The highest in multiplicity is lowest (least inter-electronic repulsion)

2) If there are two states with the same multiplicity, the one with the largest L is the lowest.

3) If the subshell < half filled the state with the lowest J is lowest in energy, if the subshell > half filled state with highest J is lowest in energy.

This works well for ground state configuration, but not always as well for excited configurations.

Friedrich Hund

Example:

Let us have a look at an example, the carbon atom. The 6 electrons of the carbon atom give rise to the following electronic configuration $1s^2 2s^2 2p^2$. We have just seen that the filled subshells do not contribute to the total orbital and spin angular momentum, since they have $L=S=0$. So we only have to consider the two electrons in the $2p$ shell. If we take into account the Pauli Exclusion Principle we can write down the following table for the different values of m and m_S to give M_L and M_S.

	m_1	m_{s1}	m_2	m_{s2}	M	M_S	M_J
1	+1	+½	+1	-½	+2	0	+2
2	+1	+½	0	+½	+1	+1	+2
3	+1	+½	0	-½	+1	0	+1
4	+1	+½	-1	+½	0	+1	+1
5	+1	+½	-1	-½	0	0	0
6	+1	-½	0	+½	+1	0	+1
7	+1	-½	0	-½	+1	-1	0
8	+1	-½	-1	+½	0	0	0
9	+1	-½	-1	-½	0	-1	-1
10	0	+½	0	-½	0	0	0
11	0	+½	-1	+½	-1	+1	0
12	0	+½	-1	-½	-1	0	-1
13	0	-½	-1	+½	-1	0	-1
14	0	-½	-1	-½	-1	-1	-2
15	-1	+½	-1	-½	-2	0	-2

We must now deduce the possible values of L and S from the values of M and M_S in the table above. The largest value is 2 and this value occurs only with M_S is 0. Therefore there must be a state with $L=2$ and $S=0$. This state accounts for the entries 1, 3, 5, 12 and 15 in the table. Remember, an $L=2$ state with $S=0$ has M values of -2, -1, 0, +1 and +2 and $M_S=0$.

If we take these entries out of the table we are left with:

	m_1	m_{s1}	m_2	m_{s2}	M	M_S	M_J
2	+1	+½	0	+½	+1	+1	+2
4	+1	+½	-1	+½	0	+1	+1
6	+1	-½	0	+½	+1	0	+1
7	+1	-½	0	-½	+1	-1	0
8	+1	-½	-1	+½	0	0	0
9	+1	-½	-1	-½	0	-1	-1
10	0	+½	0	-½	0	0	0
11	0	+½	-1	+½	-1	+1	0
13	0	-½	-1	+½	-1	0	-1
14	0	-½	-1	-½	-1	-1	-2

The largest value of M remaining is $M=1$, implying $L=1$. This $L=1$ state has as possible M values, $M=0, \pm 1$. Each of these values occurs with a value of $M_S=0$ or ± 1. So what we have is a state with $L=1$ and $S=1$. If we remove the nine entries in the table corresponding to this state we are only left with entry 10 with $M=0$ and $M_S=0$ implying $L=0$ and $S=0$.

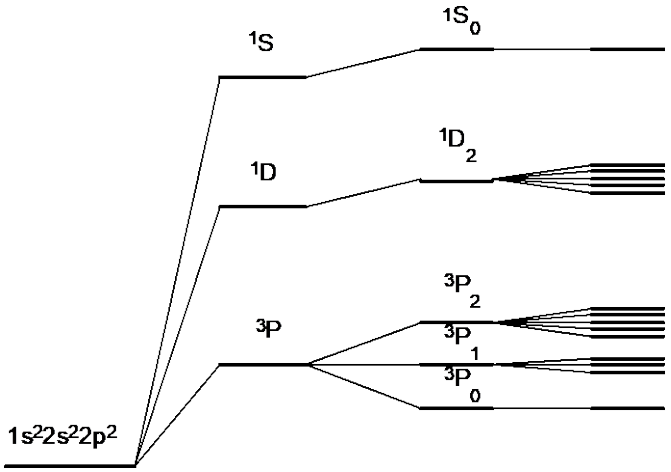
Al that has to be done now is to assign term values to these states.

$L=2,\ S=0 \quad \Rightarrow \quad J=2 \quad \Rightarrow \quad {}^1D_2$

$L=1,\ S=1 \quad \Rightarrow \quad J=0,1,2 \quad \Rightarrow \quad {}^3P_0\ ,\ {}^3P_1\ ,\ {}^3P_2$

$L=0,\ S=0 \quad \Rightarrow \quad J=0 \quad \Rightarrow \quad {}^1S_0$

The ordering of the different states according to Hund's rule are given in the figure below.



To close out and summarize our discussion of atoms, let us step back for a moment and look at the effect of the different terms in the Hamiltonian.

1) To zeroth-order there is only one state for a particular electron configuration.

2) If we consider electron repulsion, it splits the different terms. This says that states with different amounts of orbital and spin angular momentum will have different amounts of electron repulsion.

3) If we add spin orbit interaction, we split different *J* values.

4) If we add Zeeman term (*i.e.*, due to an external magnetic field), split m_j's.

11 Quantum Mechanical Treatment of Simple Molecules

One of the great achievements of quantum mechanics was a description of the stability of the chemical bond. It turns out that this is mainly a quantum mechanical effect; classical mechanics cannot fully account for its stability.

Consider the case of H_2. As you bring in two hydrogen atoms from infinity, their electron clouds begin to overlap, and at some point there is a buildup of electron probability between the nuclei; the electrons feel the attractive forces of both atoms. We want to see physically why this might occur.

First I will discuss the H_2 molecule itself, since it is one of the simplest examples of a chemical bond (H_2^+ is actually stable and is even simpler, but it doesn't display some of the more general principles that I would to demonstrate).

After discussing H_2 I will then talk about one-electron orbitals for molecules in the same way we had one-electron orbitals for atoms. These orbitals are appropriately called **Molecular Orbitals.**

We will only scratch the surface here, since the treatment of the electronic Schrödinger equation for molecules is a field in itself. My goals are (1) to be able to give you enough of a background in the quantum mechanics of simple molecules to be able to understand molecular spectroscopy, and (2) to give you some basis for understanding applications of these concepts to organic and inorganic chemistry.

The place to start a discussion of the quantum mechanical treatment of H_2 is with the **Born-Oppenheimer Approximation**.

The Hamiltonian for the Hydrogen molecule is

$$\hat{H} = -\frac{\hbar^2}{2M}\left(\nabla_A^2 + \nabla_B^2\right) - \frac{\hbar^2}{2m_e}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Ze^2}{4\pi\varepsilon_0}\left(\frac{1}{r_{1A}} + \frac{1}{r_{1B}} + \frac{1}{r_{2A}} + \frac{1}{r_{2B}}\right) + \frac{e^2}{4\pi\varepsilon_0 r_{12}} + \frac{Z^2e^2}{4\pi\varepsilon_0 R}$$

where *A* and *B* refer to the two nuclei and 1 and 2 to the two electrons.

As we discussed earlier in the course, because the nuclei are much more massive than the electrons, ($M/m_e \approx$ 1835 for hydrogen and is higher for any other molecule), we can neglect the nuclear kinetic energy term from this Hamiltonian.

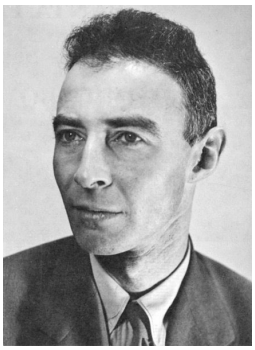
We are saying that on the timescale of electron motion, the nuclei are essentially fixed. One then solves the electronic part of the Schrödinger equation considering the internuclear separation *R* as a parameter (*i.e.* the Schrödinger equation is solved as a function of *R*).

The Hamiltonian for H_2 then becomes (in atomic units):

$$\hat{H} = -\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}} + \frac{1}{R}$$

Our general approach to solving this equation will be to apply the variational method using various types of basis functions for the trial function.


Max Born


Robert Oppenheimer

11.1 Valence-Bond method for H_2

This approach was introduced by Heitler and London in 1927 and gave the first satisfactory explanation of the stability of a chemical bond. Later this method was extended by John Slater and Linus Pauling.

We will ignore spin in this treatment because as we have seen, for any two-electron system, the spin and spatial parts of the wave function are separable.

The valence bond method takes the following approach:

At large internuclear separation, an H_2 molecule looks like two separate hydrogen atoms. The ground state wave function in this case would be

$$\psi_1 = 1s_A(1)1s_B(2)$$

where $1s_A$ denotes a 1*s* hydrogen orbital centered on nucleus *A* and $1s_B$ a 1*s* hydrogen orbital centered on nucleus *B*.

Because the electrons are indistinguishable, an equally good wave function is

$$\psi_2 = 1s_A(2)1s_B(1)$$

Heitler and London's approach was to take a linear combination of these two functions as a variational trial function:

$$\psi = c_1\psi_1 + c_2\psi_2 = c_1 1s_A(1)1s_B(2) + c_2 1s_A(2)1s_B(1)$$

As we showed in our treatment of the variational principle, when we use a trial function that is a linear combination of functions with the expansion coefficients as variable parameters, the problem reduces to solving what we called the secular determinant,

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{12} - ES_{12} & H_{22} - ES_{22} \end{vmatrix} = 0$$

where I have used the fact that H_{21}=H_{12}.


Walter Heitler


Fritz London


John Slater


Linus Pauling

Solving this equation gives us an estimate to the lowest two energies (recall that the number of energies it estimates is determined by the number of terms included in the trial function).

One can also get the coefficients c_1 and c_2 by substituting the eigenvalues, E, back into the set of equations that gave rise to the determinant.

Recall that the S terms are what we call overlap integrals. Let's first look at those along the diagonal.

$$S_{11} = \int\int 1s_A(1)1s_B(2)1s_A(1)1s_B(2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

$$= \int 1s_A(1)1s_A(1)d\mathbf{r}_1 \int 1s_B(2)1s_B(2)d\mathbf{r}_2$$

$$= 1$$

These terms are equal to 1 because the orbitals are normalized. (note that $d\mathbf{r}_1$ and $d\mathbf{r}_2$ represent the volume element for electron 1 and 2, independent of the nucleus.)

The value of the integrals S_{22} will be the same.

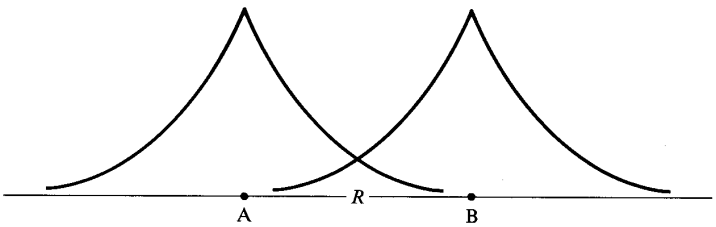
The off-diagonal overlap integrals are somewhat different, however.

$$S_{12} = \int\int 1s_A(1)1s_B(2)1s_A(2)1s_B(2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

$$= \int 1s_A(1)1s_B(1)d\mathbf{r}_1 \int 1s_B(2)1s_A(2)d\mathbf{r}_2$$

This is different from the diagonal term since the 1s orbitals in each integral are centered on different nuclei.

You can see that these two integrals are the same. Thus we can write $S_{12} = S^2$ where S is a single overlap integral. These integrals represent the overlap of the wave function on one nucleus with one on the other.



You can see that the region of overlap of these functions will be a strong function of *R*. Evaluation of the overlap integral as a function of *R* is a fairly long (but not difficult) procedure.

The result is

$$S(R) = e^{-R}\left(1 + R + \frac{R^2}{3}\right)$$

We still need to evaluate the other terms in the determinant.

$$H_{11} = \int\int 1s_A(1)1s_B(2)\hat{H}1s_A(1)1s_B(2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

Recall that

$$\hat{H} = -\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}} + \frac{1}{R}$$

You can see that this Hamiltonian is comprised of two one electron Hamiltonians plus a few additional terms due to attraction of an electron by the opposite nucleus, the inter-electronic repulsion, and the nuclear repulsion.

Since the 1*s* functions are eigenfunctions of the one electron Hamiltonians, we can see that

$$H_{11} = -\frac{1}{2} - \frac{1}{2} + J = -1 + J$$

where

$$J = \int\int 1s_A(1)1s_B(2)\left(-\frac{1}{r_{1B}} - \frac{1}{r_{2A}} + \frac{1}{r_{12}} + \frac{1}{R}\right)1s_A(1)1s_B(2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

Recall that the energy of the H atom in atomic units is -1/2, so this looks like twice the H atom energy plus the integral.

We can break up this integral to see its physical significance.

$$J = -\int \frac{|1s_A(1)|^2}{r_{1B}} d\mathbf{r}_1 - \int \frac{|1s_B(2)|^2}{r_{2A}} d\mathbf{r}_2 + \int\int \frac{|1s_A(1)|^2\,|1s_B(2)|^2}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + \frac{1}{R}$$

The first term is the coulomb interaction of the electron on nucleus A with nucleus *B*. The second term is the interaction of the electron on *B* with nucleus *A*. The third term is the inter-electronic repulsion. The last term is the nuclear repulsion. This integral is therefore called a ***coulomb integral.***

I will not take the time to evaluate this integral here. The result, which will be a function of the distance *R*, is

$$J = e^{-2R}\left(\frac{1}{R} + \frac{5}{8} - \frac{3}{4}R - \frac{R^2}{6}\right)$$

One could easily show that $H_{22}=H_{11}$.

The final term to evaluate is H_{12}

$$H_{12} = \int\int 1s_A(1)1s_B(2)\hat{H}1s_A(2)1s_B(1)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

If we were to substitute the Hamiltonian into the integral and use the fact that the 1*s* functions are eigenfunctions of the 1 electron Hamiltonians one obtains

$$H_{12} = -S^2 + K$$

where $$K = \int\int 1s_A(1)1s_B(2)\left(-\frac{1}{r_{1A}}-\frac{1}{r_{2B}}+\frac{1}{r_{12}}+\frac{1}{R}\right)1s_A(2)1s_B(1)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

K cannot be interpreted as a coulomb integral in the same way as *J*. We can rearrange this integral to get

$$K = \int\int 1s_A(1)1s_B(1)\left(-\frac{1}{r_{1A}}-\frac{1}{r_{2B}}+\frac{1}{r_{12}}+\frac{1}{R}\right)1s_A(2)1s_B(2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$

This integral arises because we are using a trial function that does not distinguish between the two electrons. (If we used only one of the two terms in the trial function, this *K* integral would not have arisen.) Because of this, *K* is called an ***exchange integral,*** and as such it is strictly a quantum mechanical property.

The evaluation of this exchange integral is pretty involved and I will not do it here. I will plot it as a function of *R* in a moment.

We are left with the following secular equation:

$$\begin{vmatrix} -1+J-E & -S^2+K-ES^2 \\ -S^2+K-ES^2 & -1+J-E \end{vmatrix} = 0$$

This gives us a quadratic equation with two roots.

$$E_{\pm} = -1+\frac{J \pm K}{1 \pm S^2}$$

Note however that -1 is the energy of two isolated hydrogen atoms (in atomic units). We can define ΔE as the energy of H_2 relative to that of the isolated atoms. We then have

$$\Delta E_{\pm} = \frac{J \pm K}{1 \pm S^2}$$

If we take E_+ and put it back into the equations that gave us the secular determinant, we can get the constants c_1 and c_2 that go into our expression for the wave function.

We would find that

$$c_1 = c_2 = \frac{1}{\sqrt{2\left(1+S^2\right)}}$$

This gives us

$$\psi_+ = \frac{1}{\sqrt{2\left(1+S^2\right)}}\left(\psi_1+\psi_2\right)$$

Recall that

$$\psi_1 = 1s_A(1)1s_B(2)$$

and

$$\psi_2 = 1s_A(2)1s_B(1)$$

You can see that this function is symmetric with respect to interchange of electrons 1 and 2, and hence its designation as ψ_+.

Similarly, E_- would give us

$$\psi_- = \frac{1}{\sqrt{2\left(1-S^2\right)}}\left(\psi_1 - \psi_2\right)$$

This function is antisymmetric with respect to particle interchange.

We have neglected spin up to this point. Remember we can do this because it is a two-electron system and the wave functions factor into a spatial and spin part. We can look at these functions and see what the spin parts need to be.

Since ψ_+ is symmetric with respect to interchange, the spin part must be antisymmetric.

Remember from our treatment of Helium, there is only one antisymmetric spin function involving α and β, but 3 symmetric functions:

$$\frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2) - \beta(1)\alpha(2)\right] \qquad \text{antisymmetric}$$

$$\left.\begin{array}{c} \alpha(1)\alpha(2) \\ \beta(1)\beta(2) \\ \frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2) + \beta(1)\alpha(2)\right] \end{array}\right\} \qquad \text{symmetric}$$

Thus, to have the correct overall symmetry behavior, ψ_+ must go with the antisymmetric spin function, and ψ_- goes with any of the symmetric spin functions.

Now let us look at the energy eigenvalues $\Delta E_\pm$ as a function of *R*. Recall that

$$\Delta E_\pm = \frac{J \pm K}{1 \pm S^2}$$

and that *J*, *K* and *S* are functions of *R*. If we plot these we see the following

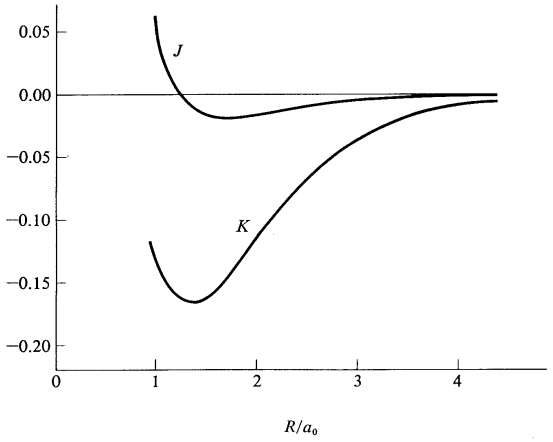
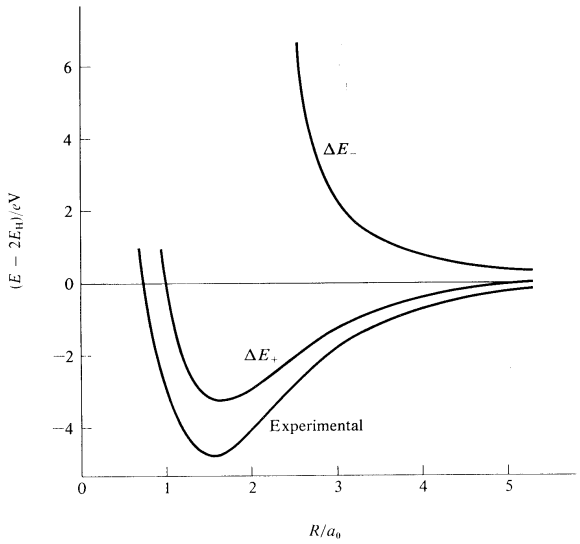


Figure 9-4. The Heitler-London valence-bond Coulomb (J) and exchange (K) integrals as a function of internuclear separation R. All quantities are expressed in atomic units.

As you can see, it is the exchange integral that accounts for most of the stability of the H_2 bond (in the context of the trial function we have chosen). Note that the overlap integral is a monotonically decaying function of *R*. Because the exchange integral is a quantum mechanical quantity, *the existence of the chemical bond is mainly a quantum mechanical effect.*



(The two curves in the figure above represent the electronic energy as a function of *R*. We could take this as a 1-dimensional function and solve the vibrational problem!)

While the valence bond approach doesn't give very good quantitative agreement, it gives us a good physical picture for the stability of the chemical bond.

We will now take a slightly different approach to solving the electronic Schrödinger equation for molecules. We will still use the variational principle, but rather than starting with wave functions that represent two hydrogen atoms at infinite distance (*i.e.* two atomic orbitals), we will construct one-electron molecular orbitals and then put electrons into these one-electron functions.

We will derive the molecular orbitals themselves by considering the H_2^+ molecule. It is a one-electron diatomic molecule, and will play the same role in molecular wave functions as H did for atomic wave functions. This way of describing molecular bonding was pioneered by Robert Mulliken.

Robert Mulliken

11.2 Molecular Orbital treatment of H_2^+

By using an approximate treatment of H_2^+, we develop the framework in which we can treat many electron molecules. We want to find functions to serve as *molecular orbitals* in the same way the H atom functions serve as atomic orbitals. We will begin by using the variational principle on H_2^+.

The Hamiltonian for H_2^+ is

$$\hat{H} = -\frac{1}{2}\nabla^2 - \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{R}$$

We will try a linear trial function of the form

$$\psi = c_1 1s_A + c_2 1s_B$$

This is called a LCAO-MO (linear combination of atomic orbitals-molecular orbital) where $1s_A$ and $1s_B$ are 1*s* H atom wave functions centered on nucleus *A* or *B*.

Remember from our discussion of the variational principle that the optimal values of c_1 and c_2 can be found from a solution of the secular equation.

$$\begin{vmatrix} H_{AA} - E & H_{AB} - ES \\ H_{BA} - ES & H_{BB} - E \end{vmatrix} = 0$$

where

$$H_{AA} = \int 1s_A \hat{H} 1s_A \, d\tau$$

$$H_{BB} = \int 1s_B \hat{H} 1s_B \, d\tau$$

$$S = \int 1s_A 1s_B \, d\tau$$

We can see that $H_{AA} = H_{BB}$ since the Hamiltonian is symmetrical with respect to interchange of r_a and r_b.

Also, $H_{AB} = H_{BA}$ by symmetry

$$H_{AB} = \int 1s_A \hat{H} 1s_B \, d\tau$$

Solving the secular equation (analogous to the way we did with the valence-bond method) yields two roots:

$$E_{\pm} = \frac{H_{AA} \pm H_{AB}}{1 \pm S}$$

These two roots are upper bounds for the energies of the ground and first excited state of H_2^+.
(We get the first 2 states because we included 2 terms in our linear trial function)

We can get the coefficients c_1 and c_2 in our linear trial function by substituting the roots E_+ and E_- back into our system of linear equations (which we never wrote explicitly).

We get

$$\psi_{\pm} = \frac{1}{\sqrt{2 \pm 2S}} \left(1s_A \pm 1s_B \right)$$

What we have found here is basically a one electron analog of the valence bond wave function that we used earlier, however there is an important difference between the valence bond and molecular orbital approaches.

Let's digress for a moment to compare the wave function for H_2 using both the VB and MO approaches.

Recall the valence bond wavefunction for H_2 is given by

$$\psi_{VB} = 1s_A(1)1s_B(2) + 1s_A(2)1s_B(1)$$

Consider what the MO for H_2 might look like. If we put each of the two electrons in a one electron molecular orbital, the wave function would be a product of those one electron functions:

$$\psi_{MO} = \left(1s_A(1) + 1s_B(1)\right)\left(1s_A(2) + 1s_B(2)\right)$$

$$= 1s_A(1)1s_B(2) + 1s_B(1)1s_A(2) + 1s_A(1)1s_A(2) + 1s_B(1)1s_B(2)$$

The first two terms here are just the terms in the valence-bond wave function. The second two terms correspond to electron configurations in which both electrons are on one atom.

Using electron dot formulas this would look like:

$H_A{:}\,H_B$ and $H_A{:}H_B$

or

$H_A^- H_B^+$ and $H_A^+ H_B^-$

These last two terms represent ionic structures.

So $\psi_{MO} = \psi_{VB} + \psi_{ionic}$

Using a trial function like this gives a better estimate than the *VB* wavefunction alone. *This suggests that the true wave function has some ionic character.*

Getting back to the solution of our problem we find for the energy:

$$E_{\pm} = \frac{H_{AA} \pm H_{AB}}{1 \pm S}$$

If we were to look a little more closely at the integrals H_{AA} and H_{AB}, we would see that we could break them up into coulomb and exchange integrals in a similar manner to the valence-bond approach.

The result is

$$\Delta E_{\pm} = \frac{J' \pm K'}{1 \pm S}$$

where J' is given by

$$J' = \int 1s_A \left(-\frac{1}{r_B} + \frac{1}{R}\right) 1s_A \, d\mathbf{r}$$

and K' by

$$K' = \int 1s_A \left(-\frac{1}{r_A} + \frac{1}{R}\right) 1s_B \, d\mathbf{r}$$

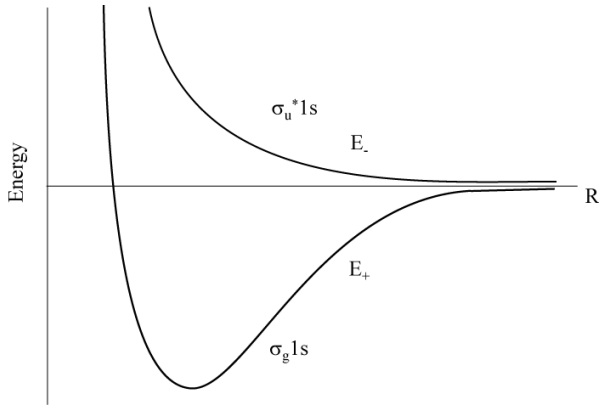
You can see these are basically the one-electron equivalents for the coulomb and exchange integrals we had earlier.

Note that the integrals are over the electron positions at a fixed value of *R*.

Thus $E_{\pm}$ is a function of *R*.

One can evaluate these integrals fairly easily, although we will not do it.

We can solve for different R and plot these as a function of *R*.



A brief digression about the notation of these states:

Each state of H_2^+ has a definite value of *m*. It turns out that $\left[\hat{L}_z, \hat{H}\right] = 0$. This will happen when the system has cylindrical symmetry. The electronic energy depends only on |*m*| since l_z^2 appears in the Hamiltonian. We did not write out the Hamiltonian explicitly to see this.

The absolute value of *m* is called λ: $\lambda = |m|$

One designates the states of H_2^+ by the value of λ in a similar way to the *s,p,d,f* notation for H atom states.

λ =	0	1	2	3	4
	σ	π	δ	φ	γ

The designation *s,p,d,f* for H atom states refers to the value of *l*. The designation $\sigma,\pi,\delta,\varphi$ for molecules indicates the value of $\lambda = |m|$.

Recall that *m* tells us how many nodes the wave function will have in the φ coordinate. We will later see that λ will also tell us something about the symmetry of the wave function.

We will also classify these states according to their properties upon inversion at the origin

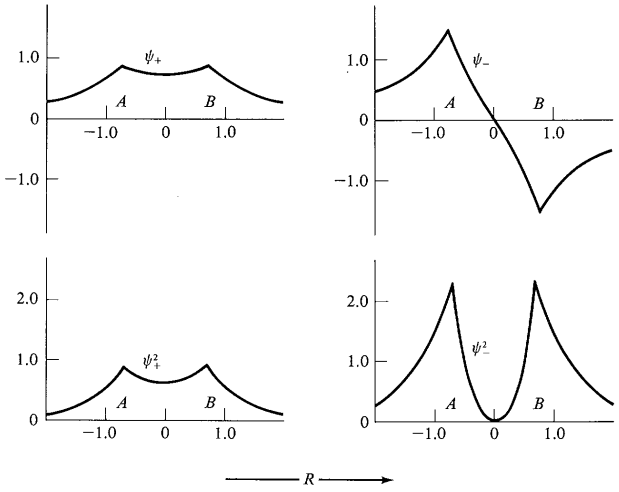
Even $\Rightarrow$ *g* for *gerade* Odd $\Rightarrow$ *u* for *ungerade*

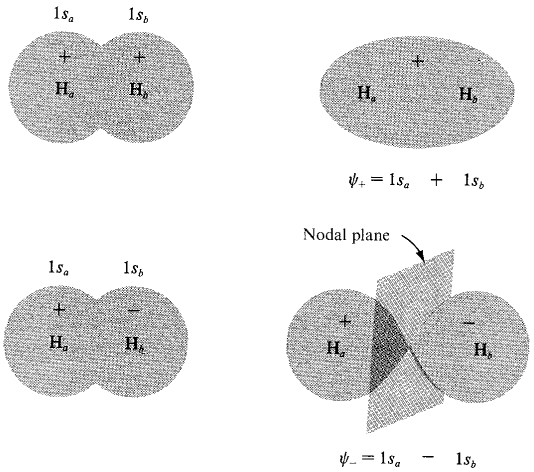
Finally, the H_2^+ states are classified by the state of the H atom to which the molecule correlates at large *R*.

Thus, the lowest state of H_2^+ would be designated σ_g 1*s*.

Why we labeled the bonding σ_g and the antibonding σ_u will become clear in a moment.

If we look at the wavefunctions corresponding to these two potential curves, we can get a little better physical understanding of why one is bonding and the other is antibonding.





If we look at $\psi_+ = 1s_A + 1s_B$

we see there is a buildup of electronic charge between the two nuclei, helping to cause mutual attraction. This arises from constructive interference between the two wave functions.

For $\psi_- = 1s_A - 1s_B$

we see that the probability goes to zero in between the two nuclei, *i.e.* the electron density is zero here, which gives rise to an antibonding orbital. (It will always be the case that antibonding orbital will have a nodal plane midway between the two nuclei.) Consequently the state is labeled $\sigma_u^* 1s$

To determine whether an orbital is bonding or antibonding, one thus must look at the reflection of the wave function through a plane midway between the atoms. If the wavefunction changes sign, it must have a node.

⇒ This means there will be destructive interference and depleted electron density and will be antibonding

⇒ If the wavefunction does not change sign, it will be a bonding orbital.

From this simple LCAO-MO picture, we can get some reasonable insight into the nature of the chemical bond.

11.3 Higher MO's of H_2^+

We now only have orbitals for the 1st two states. This was because we included only two terms in our linear variational function.

We could have written

$$\psi = c_1\psi_{1s}(A) + c_2\psi_{2s}(A) + c_3\psi_{2p_0}(A) + c_4\psi_{1s}(B) + c_5\psi_{2s}(B) + c_6\psi_{2p_0}(B)$$

Because of the symmetry of H, we would get a similar result to what we obtained previously--the coefficients of the *B* orbitals will be ±1 times those of the *A* orbitals.

$$\psi = \left[c_1\psi_{1s}(A) + c_2\psi_{2s}(A) + c_3\psi_{2p_0}(A)\right] \pm \left[c_1\psi_{1s}(B) + c_2\psi_{2s}(B) + c_3\psi_{2p_0}(B)\right]$$

Consider for a moment the two electronic states that will dissociate to a 1s hydrogen atom. For this state, we would expect c_1 to be greater than c_2 or c_3. *Certainly at large R this is true.*

As a first approximation we can take

$$\psi = c_1\left[\psi_{1s}(A) \pm \psi_{1s}(B)\right]$$

which is what we had done before.

We don't know for sure that this is the lowest state, but it is a good guess.

The same arguments hold for the two states that dissociate to a 2*s* H-atom

$$\psi = c_2\left[\psi_{2s}(A) \pm \psi_{2s}(B)\right]$$

These functions are approximations to what we would get if we solved the secular equation of variational theory.

Another way to look at it, however, is from the point of view of perturbation theory. Taking the separated atoms as the unperturbed problem, these two wavefunctions are the correct zero-order wavefunctions. In general, molecular states will correlate with each state of the separated atoms, and rough approximations to their wave functions will be given by $f_A + f_B$ and $f_A - f_B$ where f_A and f_B are hydrogen like wave functions.

Thus

$$\psi = c_2\left[\psi_{2s}(A) \pm \psi_{2s}(B)\right]$$

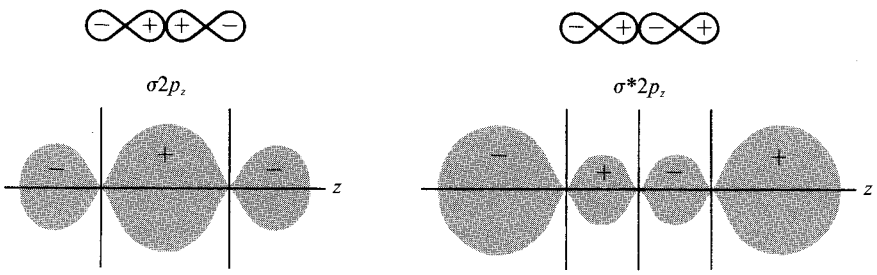
will give σ_g 2*s* and σ_u *2*s* molecular orbitals.

Think about what this zero-order picture means. In the case of atoms, the use of H atom orbitals as zeroth-order wave functions for many electron atoms says that to zeroth-order we are neglecting the inter-electronic repulsion. Once we have the zeroth-order function, we can use Perturbation Theory to improve our energy and wavefunction.

We are now doing the same thing for homonuclear diatomic molecules. Using H_2^+ *molecular orbitals*, we can put electrons in these orbitals in a manner consistent with the Pauli principle. This zeroth-order wavefunction neglects inter-electronic repulsion.

If you then use perturbation theory, the first-order correction to the wave function will mix in contributions from other functions in our basis set of atomic orbitals. However, only the functions closest in energy mix very strongly (remember the energy denominator in the first order perturbation theory correction to the wave function).
Let's look qualitatively at the few next higher MO's.

If we take the combinations of the $2p_0$ ($2p_z$) H atom state, we get the following:



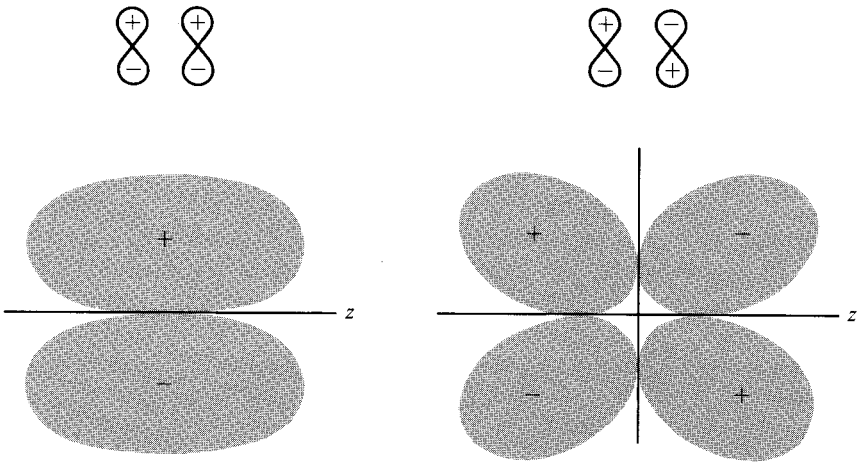
$\psi_{2p0}(A) - \psi_{2p0}(B)$ $\psi_{2p0}(A) + \psi_{2p0}(B)$

These orbitals are designated σ since we are taking linear combinations of atomic orbitals with $m = 0$. *These wave functions have cylindrical symmetry about the z-axis.*

Orbitals with $m \neq 0$ are a little more difficult to visualize because the $2p_1$ and $2p_{-1}$ orbitals are complex.

We can use $2p_x$ and $2p_y$, however these are not eigenfunctions of $\hat{L}_z$. Remember $2p_x$ and $2p_y$ are linear combinations of $2p_1$ and $2p_{-1}$. Thus, they won't have cylindrical symmetry about the *z*-axis

When we draw the linear combinations $\psi = \psi_{2p_x}(A) \pm \psi_{2p_x}(B)$ we get things that look like:



$\pi_u 2p_x$ $\pi_g^* 2p_x$

These are not cylindrically symmetrical about *z*. They have an extra nodal surface. They are not eigenfunctions of $\hat{L}_z$. This is what one often sees as π orbitals. We can do this because any linear combination of eigenfunctions of H with the same energy is still an eigenfunction of $\hat{H}$. But it will not necessarily be an eigenfunction of $\hat{L}_z$.

We can now use these orbitals to discuss many-electron homonuclear diatomic molecules. If we ignore inter-electronic repulsion, we can take the zeroth-order wave function of homonuclear diatomics to be a Slater determinant of one electron H_2^+ spin orbitals, using the LCAO-MO's for spatial parts.

The approximate relative ordering of these orbitals is

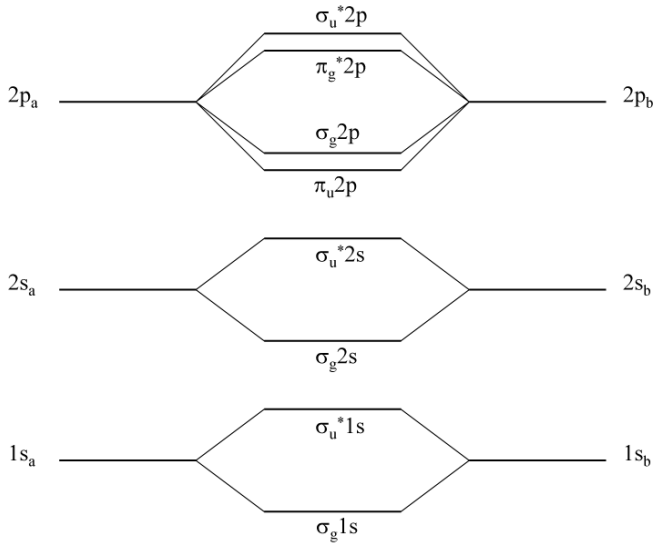
$$\sigma_g 1s < \sigma_u^* 1s < \sigma_g 2s < \sigma_u^* 2s < \pi_u 2p_{+1}, \pi_u 2p_{-1} < \sigma_g 2p < \pi_g^* 2p_{+1}, \pi_g^* 2p_{-1} < \sigma_u^* 2p$$

Note that we have used the $2p_{+1}$ and $2p_{-1}$ atomic orbitals here to form our MO's rather than the $2p_x$ and $2p_y$. Note also that this order can be slightly different for different molecules.

Each bonding orbital fills before it's corresponding antibonding orbital (indicated by the *).

The energies of these orbitals indicated by the order shown above are determined using the variational principle. However for two closely spaced levels, the order may be reversed. The energies of such molecular orbitals can be verified experimentally using photoelectron spectroscopy.

Another way to think of the levels is in a diagram like this:



11.4 Molecular orbitals for multi-electron diatomic molecules

We can now begin to get some rough idea of molecular bonding in multi-electron diatomic molecules by simply placing electrons in these orbitals.

We saw that H_2^+ has one electron in a bonding orbital, $\sigma_g 1s$.

For H_2, one puts 2 electrons in $\sigma_g 1s$ with opposite spins. This gives $(\sigma_g 1s)^2$
The two bonding electrons give a single bond.

Consider $He_2 = (\sigma_g 1s)^2(\sigma_u^* 1s)^2$ 2 bonding electrons 2 antibonding. $\Rightarrow$ No net bond.

Experimentally, He_2 shows no significant minimum in the potential energy curve.

In a more formal sense we could write:

 # bonds = ½ (# electrons in bonding orbitals - # electrons in anti-bonding orbitals)

Let us pursue this concept a little further. If you promote one of the antibonding electrons in helium to a higher state one has.

$$(\sigma_g 1s)^2(\sigma_u^* 1s)^1(\sigma_g 2s)^1$$

This has 3 bonding and 1 anti-bonding electrons. It will therefore be chemically bound. This is correct. He_2 is called an *excimer* in that it is bound in the upper electronic state, but not in the lowest (ground) electronic state.

For Li_2 we have

$$(\sigma_g 1s)^2(\sigma_u{}^*1s)^2(\sigma_g 2s)^2$$

Two net bonding electrons lead to a single bond. Experimentally Li_2 is a stable molecule.

Sometimes its configuration is written

$$KK(\sigma_g 2s)^2$$

where the *K* denotes filled *K* lithium atoms.

We did not talk about this notation when we did HF-SCF. *K*, *L*, *M* represent the quantum numbers n = 1, 2, 3. So the n. The n= 1 atomic orbitals constitute the *K* shell, n=2 the *L* shell, etc. . .

Let's jump to Nitrogen, N_2. The configuration is

$$KK(\sigma_g 2s)^2\,(\sigma_u{}^*\,2s)^2\,(\pi_u\,2p)^4\,(\sigma_g\,2p)^2$$

6 bonding electrons imply a triple bond, which is what is observed.

For O_2 $\quad KK(\sigma_g 2s)^2\,(\sigma_u{}^*\,2s)^2\,(\sigma_g\,2p)^2\,(\pi_u\,2p)^4(\pi_g{}^*\,2p)^2$

$\uparrow \quad \uparrow$

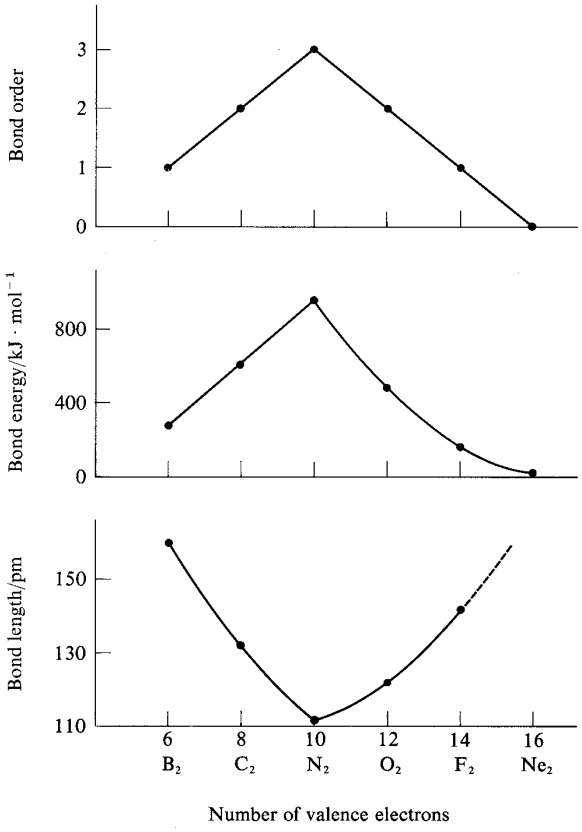
experimental evidence suggests these are reversed here.

In O_2 we have unfilled π orbitals. Hund's rule says that the triplet state has lower energy. The triplet state is the one in which the electrons are unpaired. Experimentally O_2 is paramagnetic because of the two unpaired electrons! One can use the magnetism of air to measure oxygen levels.

One usually uses the term *bond-order* to indicate the number of pairs of electrons are involved in the bond. A single bond has a bond order of 1. A double bond has a bond order of 2. etc.

As the table and figures below demonstrate, one can roughly correlate the bond order of a diatomic molecule with its bond length. The higher the bond order, the shorter (and stronger) the bond.

Species	Ground-state electron configuration	Bond order	Bond length/Å	Bond energy kJ mol^{-1}
H_2^+	$(\sigma 1s)^1$	1/2	1.06	255
H_2	$(\sigma 1s)^2$	1	0.74	431
He_2^+	$(\sigma 1s)^2(\sigma^* 1s)^1$	1/2	1.08	251
He_2	$(\sigma 1s)^1(\sigma^* 1s)^2$	0	-	-
Li_2	$KK(\sigma 2s)^2$	1	2.67	105
Be_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2$	0	-	-
B_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^2$	1	1.59	289
C_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^4$	2	1.24	599
N_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^4(\sigma 2p_z)^2$	3	1.10	942
O_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^4(\sigma 2p_z)^2(\pi^* 2p)^2$	2	1.21	494
F_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^4(\sigma 2p_z)^2(\pi^* 2p)^4$	1	1.41	154
Ne_2	$KK(\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^4(\sigma 2p_z)^2(\pi^* 2p)^4(\sigma^* 2p_z)^2$	0	-	-



You can clearly see that this simple molecular orbital picture allows us to make good qualitative predictions about the properties of molecules.

To be more quantitative, one can use the Hartree-Fock method for multi-electron diatomic molecules in an analogous way to its application to multi-electron atoms. (I won't discuss this in detail, but I would like to mention the terminology that is used.)

In using the HF method for molecules, we keep the concept of molecular orbitals that can be filled with two electrons each (consistent with the Pauli principle), however these orbitals are constructed from flexible functions that allow us to vary the parameters in such a way that we can reach the Hartree-Fock limit.

In our simple approach up to now, which is often called the LCAO-MO method (linear combination of atomic orbitals-molecular orbital), we took each molecular orbital to be a linear combination of an atomic orbital on each nucleus in such a way that the atomic orbitals maintain their identity.

As we started adding atomic orbitals to get higher molecular orbitals, I wrote things like

$$\psi = \left[c_1\psi_{1s}(A) + c_2\psi_{2s}(A) + c_3\psi_{2p_0}(A) + \ldots\right] \pm \left[c_1\psi_{1s}(B) + c_2\psi_{2s}(B) + c_3\psi_{2p_0}(B) + \ldots\right]$$

I indicated that to a good approximation, one coefficient would be dominant for each state of the separated atoms. To a first approximation we neglected the others.

Thus, I indicated that the first two states could be approximated as

$$\psi_{1,2} = c_1\left[\psi_{1s}(A) \pm \psi_{1s}(B)\right]$$

and hence we called this a $\sigma_g 1s$ molecular orbital.
The same arguments hold for the two states that dissociate to a 2*s* H-atom

$$\psi_{3,4} = c_2\left[\psi_{2s}(A) \pm \psi_{2s}(B)\right]$$

We called this a $\sigma_g 2s$ orbital.

However, if we want a more accurate estimate of the energy by using the HF method, we have to include not only the coefficients we neglected, but also make the atomic functions themselves flexible enough so that we can reach the HF limit. In doing this, the molecular orbitals lose their identity (in terms of the atomic orbitals) since the coefficients of many atomic orbitals can become significant. In this case, one often uses a different manner to label the molecular orbitals.

Correspondence between various notations for molecular orbitals

Simple LCAO-MO		HFSCF-LCAO-MO
$\sigma 1s$	$\sigma_g 1s$	$1\sigma_g$
$\sigma^* 1s$	$\sigma_u 1s$	$1\sigma_u$
$\sigma 2s$	$\sigma_g 2s$	$2\sigma_g$
$\sigma^* 2s$	$\sigma_u 2s$	$2\sigma_u$
$\pi 2p_x$	$\pi_u 2p_x$	$1\pi_u$
$\pi 2p_y$	$\pi_u 2p_y$	$1\pi_u$
$\sigma 2p_z$	$\sigma_g 2p_z$	$3\sigma_g$
$\pi^* 2p_x$	$\pi_g 2p_x$	$1\pi_g$
$\pi^* 2p_y$	$\pi_g 2p_y$	$1\pi_g$
$\sigma^* 2p_z$	$\sigma_u 2p_z$	$3\sigma_u$

11.5 Molecular term symbols for diatomic molecules

In a manner similar to that for atoms, the electronic states of diatomic molecules are designated by term symbols. In the case of atoms, I indicated that although the individual spin and orbital angular momentum quantum numbers of the electrons are no longer good quantum numbers (*i.e.*, the individual *l* and s are not conserved), the sum *L* and the sum *S* of all the electrons are still a good quantum number (if we ignore spin orbit coupling).

In the case of diatomic molecules, even if there is only one electron (*i.e.*, H_2^+), the *l* quantum number of the individual electrons is not a good quantum number. However, because of the cylindrical symmetry, m_l remains a good quantum number (for H_2^+) and $l_z = m_l \hbar$ is a conserved quantity.

Once we go to multi-electron diatomic molecules, l_z of the individual electrons is no longer conserved and thus m_l is no longer a good quantum number. However the sum, M_L of all the electrons is still a good quantum number. Also, the total *S* for the electrons is a good quantum number. Thus we can label the states of a diatomic molecule by M_L and *S*.

One constructs a molecular term symbol in an analogous way to the atomic term symbols:

$^{2S+1}|M_L|$ or since $\Lambda = |M_L|$ we can write $^{2S+1}\Lambda$

Where $M_L = m_{l1} + m_{l2} + \ldots$ and $M_s = m_{s1} + m_{s2} + \ldots$

and *S* is determined from the values of M_s.

Note that the difference from the atomic case is that the main symbol represents a scalar quantity, $|M_L|$ rather than a vector quantity *L*.

The various values of $|M_L|$ are associated with capital Greek letters according to

$\|M_L\|$	Letter
0	Σ
1	Π
2	Δ
3	Φ

(Note that these letters correspond to the *S*, *P*, *D*, and *F* in the atomic case.)

Examples of molecular term symbols are $^1\Sigma$, $^3\Pi$, and $^2\Delta$.

The determination of the molecular terms symbols from molecular-orbital electron configurations is simpler than the atomic case since M_L is a scalar quantity. However, it is important to remember that *S* is still a vector quantity. I illustrate this procedure below with 3 examples.

Consider the case of H_2, which has an electron configuration of $(\sigma 1s)^2$.

The occupied σ orbitals have $m_l = 0$.

Thus $M_L = 0 + 0 = 0$

The two electrons must have opposite spins in order to satisfy the Pauli principle, thus

$$M_S = +\frac{1}{2} - \frac{1}{2} = 0$$

Because there is only one value of M_S, then *S* must equal 0.

The terms symbol will therefore be: $^1\Sigma$

Now consider the case of He_2^+. The ground state electronic configuration is $(\sigma 1s)^2(\sigma^* 1s)^1$.
One can construct a table of the possible values of m_l and m_s.

m_{l1}	m_{s1}	m_{l2}	m_{s2}	m_{l3}	m_{s3}	M_L	M_S
0	+ ½	0	- ½	0	+ ½	0	+ ½
0	+ ½	0	- ½	0	- ½	0	- ½

The fact that $M_L = 0$ says that we have a Σ state.

The $M_S = \pm 1/2$ corresponds to the two projections of $S = 1/2$. This means $2S+1 = 2$ and we have a doublet state.

So the term symbol for He_2^+ is $^2\Sigma$.

These first two examples are quite simple since there is only one possible value of *S*.

The next example, B_2, is a little more complicated and illustrates the general scheme that one should use.
The electron configuration for B_2 is

$$(\sigma 1s)^2(\sigma^* 1s)^2\ (\sigma 2s)^2(\sigma^* 2s)^2(\pi 2p)^2$$

The first 4 molecular orbitals in B_2 have M_L=0 and M_S=0, and thus we need to consider only the last two electrons. (Recall that in the atomic case, we also ignored filled subshells.)

Each of these last two electrons is in a π orbital and can have $m_l = \pm 1$ and $m_s = \pm 1/2$. One can construct a table of the possible values, keeping in mind the indistinguishability of the electrons and the Pauli principle.

	m_{l1}	m_{s1}	m_{l2}	m_{s2}	M_L	M_S
1	+1	+ ½	+1	- ½	2	0
2	+1	+ ½	-1	+ ½	0	1
3	+1	+ ½	-1	- ½	0	0
4	+1	- ½	-1	+ ½	0	0
5	+1	- ½	-1	- ½	0	-1
6	-1	+ ½	-1	- ½	-2	0

Entries 1 and 6 in the table correspond to $|M_L| = 2$ and $M_S = 0$. This gives rise to a $^1\Delta$ (singlet delta) state.

Entries 2, 3, and 5 correspond to $M_L=0$ and $S=1$, and so we have a $^3\Sigma$ (triplet sigma) state.

Entry number 4 has to $M_L=0$ and $M_S=0$, an thus it corresponds to a $^1\Sigma$ state.

Thus we have 3 possible molecular states for this electronic configuration of B_2: $^1\Delta$, $^3\Sigma$, and $^1\Sigma$.

Hund's rules apply to molecular electronic states as well as to atomic states. Hund's rule says that the state with the largest spin multiplicity will be the ground state. Thus, we predict that the ground state of B_2 is a $^3\Sigma$ state.

(For the Σ states of homonuclear diatomics, there is also a right superscript of + or - that indicates the symmetry of the wave function with respect to reflection in a plane containing the internuclear axis. We will not concern ourselves with this for the moment.)

11.6 MO treatment of Heteronuclear Diatomics

The treatment of heteronuclear diatomic molecules is similar to that of homonuclear diatomics.

In the case where the two atoms in a diatomic have atomic number that differ only slightly, such as in CO, we could consider the molecule being formed from the isoelectronic molecule N_2 by a gradual transfer of nuclear charge from one nucleus to the other. During this hypothetical transfer, the original N_2 MO's would slowly vary to give finally the CO MO's. We therefore expect the CO molecular orbitals to bear some resemblance to those of N_2.

In a case such as this, the symbols used for the MO's are similar to those for homonuclear diatomics.

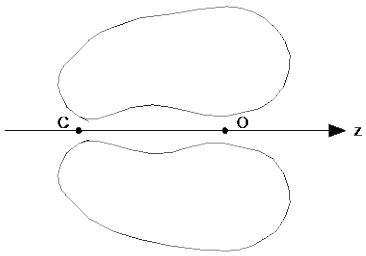
However, in the heteronuclear case, the electronic Hamiltonian lacks the symmetry of the homonuclear diatomic case, and hence the *g*, *u* property of the molecular orbitals disappears.

The correlation between the N_2 and CO subshell designations is

N_2	$1\sigma_g$	$1\sigma_u$	$2\sigma_g$	$2\sigma_u$	$1\pi_u$	$3\sigma_g$	$1\pi_g$	$3\sigma_u$
CO	1σ	2σ	3σ	4σ	1π	5σ	2π	6σ

The MO's of the same symmetry are numbered in order of increasing energy. Because of the absence of the g, u property, the numbers of corresponding homonuclear and heteronuclear MO's differ.

Shown below is a a sketch of a contour of the CO 1π MO taken from an SCF calculation. Note its resemblance to the π MO that I had drawn earlier for homonuclear diatomics.



A Hartree-Fock SCF calculation for the expansion coefficients of this molecular orbital in terms of Slater atomic orbitals gives

$$1\pi = 0.469\ (2p\pi_C) + 0.771\ (2p\pi_O)$$

This can be compared to the corresponding bonding orbital of N_2 which is given by

$$1\pi_u = 0.624\ (2p\pi_a + 2p\pi_b)$$

You can see both from the picture of the orbital as well as in the SCF calculation that the amount of atomic orbital centered on the carbon and the oxygen are not the same. This makes sense since the symmetry is now broken and the charge is different on the two nuclei. (Note, however, that the wave functions are not normalized).

11.7 More on Valence-Bond Theory

I would like to say a bit more about valence-bond theory. Although this approach is not accurate enough to be used for computations, it provides some important qualitative notions about chemical bonding. More specificially, valence-bond theory guides our intuitive ideas regarding Lewis formulas, resonance formulas, etc. In our treatment of H_2, we constructed a wave function that had one electron on each atom.

If we let

$$\psi_1 = \begin{vmatrix} 1s_A\alpha(1) & 1s_B\beta(1) \\ 1s_A\alpha(2) & 1s_B\beta(2) \end{vmatrix}$$

and

$$\psi_2 = \begin{vmatrix} 1s_A\beta(1) & 1s_B\alpha(1) \\ 1s_A\beta(2) & 1s_B\alpha(2) \end{vmatrix}$$

each of which represent H_2 with one electron on each atom, then we can use

$$\psi = c_1\psi_1 + c_2\psi_2$$

as a trial function and minimize the energy with respect to c_1 and c_2.

The ground state wave function is given by

$$\psi = \frac{1}{\sqrt{2\left(1+S^2\right)}}\left(\psi_1 - \psi_2\right)$$

(Note that the problem is formulated differently than our original treatment, and the definition of ψ_1 and ψ_2 are different, but if you multiply out these determinants and compare the result to the one we had before, you will see that it is the same).

Consider now the molecule LiH. One of the (unnormalized) Slater determinants in the valence bond wave function of LiH is of the form

$$\psi_1 = \begin{vmatrix} \psi_{1sLi}\alpha(1) & \psi_{1sLi}\beta(1) & \psi_{2sLi}\alpha(1) & \psi_{1sH}\beta(1) \\ \psi_{1sLi}\alpha(2) & \psi_{1sLi}\beta(2) & \psi_{2sLi}\alpha(2) & \psi_{1sH}\beta(2) \\ \psi_{1sLi}\alpha(3) & \psi_{1sLi}\beta(3) & \psi_{2sLi}\alpha(3) & \psi_{1sH}\beta(3) \\ \psi_{1sLi}\alpha(4) & \psi_{1sLi}\beta(4) & \psi_{2sLi}\alpha(4) & \psi_{1sH}\beta(4) \end{vmatrix}$$

We can abbreviate these Slater determinants by only listing the elements on the diagonal. This would give the notation:

$$\psi_1 = \left|\psi_{1sLi}\alpha(1) \quad \psi_{1sLi}\beta(2) \quad \psi_{2sLi}\alpha(3) \quad \psi_{1sH}\beta(4)\right|$$

The other contribution to the valence-bond wave function that is distinct (not just changing labels on the electrons) is

$$\psi_2 = \left|\psi_{1sLi}\alpha(1) \quad \psi_{1sLi}\beta(2) \quad \psi_{2sLi}\beta(3) \quad \psi_{1sH}\alpha(4)\right|$$

Thus we can write the valence-bond wave function as

$$\psi_{cov} = c_1\psi_1 + c_2\psi_2$$

where I have used the subscript ψ_{cov} to indicate that we have chosen configurations that correspond to covalent bonding where the electrons are equally shared between the bonded atoms.

A variational calculation would give $c_1 = -c_2$, since there should be no preference for which electron has spin up or down. Thus we have

$$\psi_{cov} = \left|\psi_{1sLi}\alpha(1) \quad \psi_{1sLi}\beta(2) \quad \psi_{2sLi}\alpha(3) \quad \psi_{1sH}\beta(4)\right| - \left|\psi_{1sLi}\alpha(1) \quad \psi_{1sLi}\beta(2) \quad \psi_{2sLi}\beta(3) \quad \psi_{1sH}\alpha(4)\right|$$

If we were to use this function to calculate the bond length and bond energy, we would get values of 3.01 au and -215.98 au respectively.

This should be compared to the experimental values of 3.02 au and -219.71 au.

One reason for the poor agreement in the energy is that we have not allowed for any ionic character in our valence-bond wave function.

We can include the ionic structure Li^+H^- into our valence-bond treatment by using

$$\psi_{ionic} = \left|\psi_{1sLi}\alpha(1) \quad \psi_{1sLi}\beta(2) \quad \psi_{1sH}\alpha(3) \quad \psi_{1sH}\beta(4)\right|$$

Note that this wave function describes Li^+H^- in that there are two electrons on Li^+ and two electrons on H^-.

We can now take a linear combination of ψ_{cov} and ψ_{ionic} and write

$$\Psi_{VB} = c_{cov}\Psi_{cov} + c_{ionic}\Psi_{ionic}$$

The inclusion of the ionic term improves the energy calculation to a value of -217.0 au.

Because the two 1*s* electrons in the lithium atom do not play a great role in the formation of the bond in LiH, it is a convenient, common approximation to ignore inner-core electrons in the valence-bond wave function and to consider only the bonding, or valence, electrons. In this approximation we would have

$$\psi_{cov} = \left|\psi_{2sLi}\alpha(1) \quad \psi_{1sH}\beta(2)\right| - \left|\psi_{2sLi}\beta(1) \quad \psi_{1sH}\alpha(2)\right|$$

and

$$\psi_{ionic} = \left|\psi_{1sH}\alpha(1) \quad \psi_{1sH}\beta(2)\right|$$

The neglect of these inner-core electrons is not so important for LiH, since this is a small enough molecule that it can be solved to a high degree of accuracy. However as we will see shortly, this approach becomes very useful when discussing larger systems.

The square of the expansion coefficients, c_{cov}^2 and c_{ionic}^2, give some indication of the ionic and covalent character in the molecule. However, one must be careful with this interpretation. The precise values of these coefficients depend on the form of the atomic orbitals that are used.

The idea of introducing ionic terms into valence-bond wave functions nicely illustrates the concept of resonance that you learned in first year chemistry and organic chemistry. Quantum mechanically, we see that if we can write two or more sensible Lewis structural formulas for a molecule, then the wave function for that molecule is a linear combination of these structures and the "true" picture is some intermediate structure. The variational principle, which gives the numerical values of the coefficients in the linear combination, provides an indication of the relative importance of various possible Lewis formulas.

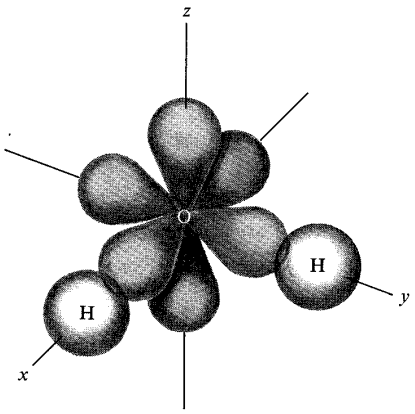
Thus, the qualitative ideas about resonance structures have a quantitative basis in quantum chemistry. (This is true about many of the qualitative ideas presented in organic chemistry.)

The extension of the valence-bond method to non-linear polyatomic molecules is straightforward in principle.

Consider the molecule H_2O. The electron configuration of the oxygen atom is

$$1s^2\ 2s^2\ 2p_x^1\ 2p_y^1\ 2p_z^2$$

suggesting that the unpaired $2p_x^1$ and $2p_y^1$ electrons are available for bonding with the hydrogen atoms.



From this simple valence-bond picture considering only covalent terms, one would expect the bond angle in H_2O to be 90°, since the $2p_x^1$ and $2p_y^1$ orbitals lie along the *x* and *y* axes and we would expect the hydrogen 1*s* orbitals to maximize the overlap with the oxygen 2*p* orbitals.

The prediction of a 90° H-O-H bond angle is in poor agreement with the measured bond angle of 104°.

However, if one introduces ionic terms in the wave function, the hydrogen atoms will develop some positive character and repel each other, giving a bond angle closer to 104°.

For the corresponding molecules H_2S, H_2Se, and H_2Te, the electronegativities of the sulfur, selenium, and tellurium atoms are such that the ionic terms contribute progressively less. In this case, the predicted angle becomes progressively closer to the prediction of 90°.

If we go on to the case of NH_3, the valence-bond approach predicts that the H-N-H bond angles are 90° as compared to the experimental value of 107°. Once again, the situation can be improved by introducing ionic terms, however, this approach clearly fails us when we get to methane, CH_4.

Clearly, the carbon electron configuration of $1s^2\ 2s^2\ 2p_x^1\ 2p_y^1$ does not explain (using a valence-bond picture) the well-known tetrahedral bonding in methane and other saturated hydrocarbons.

An approach that better describes the directionality of chemical bonding involves a consideration of hybrid orbitals , a subject that you will have likely discussed both in introductory chemistry and in organic chemistry. We will treat this subject from a quantum mechanical point of view.

11.8 Hybrid Orbitals

Like valence bond theory, the concept of hybrid orbitals is not particularly accurate in a computational sense, but it provides an important qualitative picture from which we can make predictions about the geometries of simple molecules. I will briefly introduce some of the mathematical background behind hybrid orbitals.

11.8.1 sp hybrid orbitals

Consider first the molecule beryllium hydride, BeH_2. The two Be-H bonds in this molecule are equivalent and the H-Be-H bond angle is 180°.

The ground state electron configuration of Be is: $1s^2\ 2s^2$

To represent the two equivalent Be-H bonds in BeH_2 that make an angle of 180° with respect to one another, we will take a linear combination of the beryllium 2s orbital and one of the beryllium 2*p* orbitals (*i.e.*, the $2p_z$). We can do this for each of the 2*s* electrons since the orbital holds two electrons. Thus, we have

$$\xi = a_1\, 2s_{Be} + b_1\, 2p_{zBe}$$

$$\xi' = a_2\, 2s_{Be} + b_2\, 2p_{zBe}$$

Linear combinations of orbitals on the same atom are called ***hybrid orbitals.***

(Recall that if we neglect inter-electronic repulsion, the s and p orbitals on an atom would have the same energy, so in this limit, these hybrid orbitals would still be eigenfunctions of the atomic Hamiltonian. In reality, the *s* and *p* do not have the same energy, however this approach still provides a useful qualitative picture of molecular bonding.)

The two bonds in BeH_2 are described by the following bond orbitals:

$$\phi = c_1 1s_A + c_2 \xi$$

$$\phi' = c_1' 1s_B + c_2' \xi'$$

where the $1s_A$ and $1s_B$ are the 1*s* orbitals of the two hydrogen atoms.

So the bond orbitals are linear combinations of an atomic orbital on the hydrogen atom and a hybrid orbital on the beryllium.

We will now determine the form of the two hybrid orbitals ξ and ξ' so that ϕ and ϕ' describe two equivalent bond orbitals that are directed 180° from one another.

We can approximate the 2*s* and $2p_z$ orbitals of the beryllium atom by the Slater orbitals

$$\psi_{2s} = \sqrt{\frac{1}{4\pi}} R(r)$$
$$\psi_{2p_z} = \sqrt{\frac{3}{4\pi}} \cos\theta \, R(r)$$

where for simplicity I have used the same functional form for the radial part of the wave functions. (Recall that the angular part is simply given by the spherical harmonics). The constants in front insure that the functions are normalized.

If we substitute these functions into the expression for ξ we get

$$\xi = \frac{R(r)}{\sqrt{4\pi}} \left(a_1 + \sqrt{3} b_1 \cos\theta \right)$$

If a_1 and b_1 are both positive, then ξ is directed along the positive *z*-axis, and we can choose this to be the case since we are free to orient the molecule in space as we wish.

The other hybrid orbital, ξ' is then given by

$$\xi' = \frac{R(r)}{\sqrt{4\pi}} \left(a_2 + \sqrt{3} b_2 \cos\theta \right)$$

Recall that the original *s* and *p* functions are orthogonal, and when we take a linear combination of these orbitals to get another set of two orbitals, ξ and ξ', it is convenient to require that ξ and ξ' be orthogonal.

In this case we have that:

$$\int_0^\infty \int_0^\pi \int_0^{2\pi} \xi(r,\theta) \xi'(r,\theta) r^2 \sin\theta \, dr \, d\theta \, d\varphi = 0$$

If we have chosen our radial function to be normalized, then

$$\int_0^\infty R^2(r) r^2 dr = 1$$

Combining this with the expressions for ξ and ξ' gives us for the first integral

$$\int_0^\pi \left(a_1 + \sqrt{3} b_1 \cos\theta \right) \left(a_2 + \sqrt{3} b_2 \cos\theta \right) \sin\theta d\theta = 0$$

The evaluation of this integral is straightforward and gives

$$a_1 a_2 + b_1 b_2 = 0$$

Because we have taken a_1 and b_1 to be positive, then a_2 and b_2 must have opposite signs, and this causes ξ' to be directed along the negative *z*-axis, 180° from ξ.
Because the two Be-H bonds in BeH_2 are equivalent, we require that the two hybrid orbitals have the same shape.

Thus we require that $a_1 = a_2$ and that $b_1 = -b_2$.

Furthermore, because $a_1 a_2 + b_1 b_2 = 0$

we find that $a = \pm b$

Finally, if the two equivalent orbitals are normalized, then

$$\xi = \frac{1}{\sqrt{2}}(2s + 2p_z)$$
$$\xi' = \frac{1}{\sqrt{2}}(2s - 2p_z)$$

Because these hybrid orbitals are made up of a 2*s* orbital and one 2*p* orbital, they are called *sp* hybrid orbitals.

The BeH_2 molecule is formed by overlapping a hydrogen 1*s* orbital with each of the sp hybrid orbitals.
The electron configuration of BeH_2 in this bond-orbital description is therefore

$$K(\phi)^2(\phi')^2$$

where ϕ and ϕ' are the bond orbitals described earlier. In this picture, a chemical bond is described as two electrons of opposite spin occupying a bond orbital.

11.8.2 sp^2 hybrid orbitals

Consider now the case of BH_3. The three B-H bonds in BH_3 are equivalent and lie in a plane, directed 120° from each other.

To describe the three equivalent bonds in BH_3, we must construct three hybrid orbitals on the boron atom. We will construct each of these orbitals as a linear combination of one *s* and two *p* orbitals

$$\xi_1 = a_1 2s + b_1 2p_z + c_1 2p_x$$

$$\xi_2 = a_2 2s + b_2 2p_z + c_2 2p_x$$

$$\xi_3 = a_3 2s + b_3 2p_z + c_3 2p_x$$

where the 2*s* and $2p_z$ orbitals are the same we used in the previous example, and the $2p_x$ is a Slater orbital given by:

$$\psi_{2p_x} = \sqrt{\frac{3}{4\pi}} \sin\theta \cos\varphi R(r)$$

Because these hybrid orbitals are constructed from one 2*s* and two 2*p* orbitals, they are called sp^2 hybrid orbitals.

Because we can choose the overall orientation of the molecule with respect to the axis system, we can let one hybrid orbital lie along the *z*-axis. In this case, the contribution of the p_x orbital will be zero, and

$$\xi_1 = a_2 2s + b_1 2p_z$$

In analogy to the case of BeH_2, because the *s* orbital is spherically symmetric and the three hybrid orbitals are equivalent, we shall take equal contribution of the 2*s* orbital to each hybrid orbital.

This means that $a_1 = a_2 = a_3$

Furthermore, because there is one 2*s* orbital to be distributed among the three hybrids, it must be true that

$$a_1^2 + a_2^2 + a_3^2 = 1$$

(Note: This expresses the conservation of the 2*s* orbital. Whenever we take three orthogonal functions and take linear combinations to make three others, there will always be both a normalization condition and what I call a conservation condition, meaning that the total amount of the original orbitals must be conserved.)

From these two conditions we find that

$$a_1 = a_2 = a_3 = \frac{1}{\sqrt{3}}$$

If we now substitute a_1 into ξ_1 we have

$$\xi_1 = \frac{1}{\sqrt{3}} 2s + b_1 2p_z$$

Because ξ_1 is normalized, we have that

$$\frac{1}{3} + b_1^2 = 1$$

$$b_1 = \sqrt{\frac{2}{3}}$$

Thus

$$\xi_1 = \frac{1}{\sqrt{3}} 2s + \sqrt{\frac{2}{3}} 2p_z$$

The second hybrid orbital is

$$\xi_2 = \frac{1}{\sqrt{3}} 2s + b_2 2p_z + c_2 2p_x$$

The requirement that ξ_1 and ξ_2 be orthogonal leads to the condition that

$$a_1 a_2 + b_1 b_2 + c_1 c_2 = 0$$

Note that this is the same condition that we had in the previous example. If we have a set of orthogonal functions and take linear combinations of them to get a new set of orthogonal functions, this condition will always arise. It is simply the condition of two orthogonal vectors.

With $c_1 = 0$, this condition leads to

$$\frac{1}{3}+\sqrt{\frac{2}{3}}b_2=0$$

$$b_2=-\frac{1}{\sqrt{6}}$$

Thus

$$\xi_2=\frac{1}{\sqrt{3}}2s-\frac{1}{\sqrt{6}}2p_z+c_2 2p_x$$

By requiring ξ_2 to be normalized we have

$$\frac{1}{3}+\frac{1}{6}+c_2^2=1$$

$$c_2=\frac{1}{\sqrt{2}}$$

We therefore have that

$$\xi_2=\frac{1}{\sqrt{3}}2s-\frac{1}{\sqrt{6}}2p_z+\frac{1}{\sqrt{2}}2p_x$$

Once we have ξ_1 and ξ_2, we can find many expressions that relate the remaining coefficients to those that are not yet determined.

For example, we have

$$b_1^2+b_2^2+b_3^2=1$$

$$c_1^2+c_2^2+c_3^2=1$$

together with the orthogonality conditions

$$a_1a_3+b_1b_3+c_1c_3=0$$

$$a_2a_3+b_2b_3+c_2c_3=0$$

and the normalization condition

$$a_3^2+b_3^2+c_3^2=1$$

At this point, we have an over determined system, and we need not use all of these relations. I will use the orthogonality conditions.

From the first of the two orthogonality conditions, we have

$$a_1 a_3 + b_1 b_3 + c_1 c_3 = 0$$

$$\frac{1}{3} + \sqrt{\frac{2}{3}} b_3 + 0 = 0$$

$$b_3 = -\frac{1}{\sqrt{6}}$$

From the second orthogonality condition

$$a_2 a_3 + b_2 b_3 + c_2 c_3 = 0$$

$$\frac{1}{3} + \frac{1}{6} + \frac{1}{\sqrt{2}} c_3 = 0$$

$$c_3 = -\frac{1}{\sqrt{2}}$$

Thus

$$\xi_3 = \frac{1}{\sqrt{3}} 2s - \frac{1}{\sqrt{6}} 2p_z - \frac{1}{\sqrt{2}} 2p_x$$

I will leave it as an exercise for you to demonstrate that these three hybrid orbitals lie 120° apart from each other.

11.8.3 sp^3 hybrid orbitals

From what we have done up to this point, the development of the four sp^3 hybrid orbitals should directly follow. By choosing the first to lie along the *z*-axis and using the orthogonality and normalization relations, you can easily find the coefficients of the following orbitals.

$$\xi_1 = a_1 2s + b_1 2p_x + c_1 2p_y + d_1 2p_z$$

$$\xi_2 = a_2 2s + b_2 2p_x + c_2 2p_y + d_2 2p_z$$

$$\xi_3 = a_3 2s + b_3 2p_x + c_3 2p_y + d_3 2p_z$$

$$\xi_4 = a_4 2s + b_4 2p_x + c_4 2p_y + d_4 2p_z$$

You can then show that they have the shape of a tetrahedron. I will leave this as an optional exercise.

11.9 π-electron approximation and Hückel molecular orbital theory

I would like to take our discussion of simple molecules one step further to relate the quantum mechanical treatment that we have developed to some of the concepts that you will encounter in organic chemistry. I will not give a very sophisticated treatment but rather introduce some of the basic notions that will help you to better understand what you do in organic chemistry.

Our discussion of hybrid orbitals leads us to consider the case of unsaturated hydrocarbons. Consider, for example, the molecule ethylene, C_2H_4. It is planar and has all of its bond angles equal to 120°. One can describe the structure of the molecule by assuming that all the carbon atoms form sp^2 hybrid orbitals as shown in the figure below:

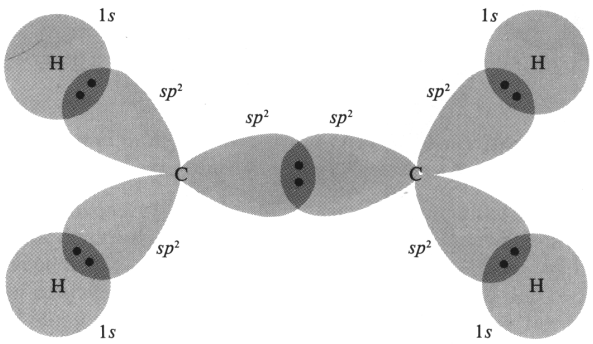


Figure 9-28. The σ-orbital framework of the ethylene molecule.

The CH bonds result from the overlap of the 1*s* hydrogen orbital with an sp^2 carbon hybrid orbital. Part of the CC bond results from the overlap of the two sp^2 carbon hybrid orbitals.

Each of the bonds shown in the figure above are σ bonds. This is because the carbon sp^2 hybrid orbitals are planar with no nodes about their respective bond axis and the hydrogen 1*s* orbitals have no nodes since l=0 and m=0. (Note that when we are considering the symmetry of these bonds, in the present case a σ bond, we are considering the local symmetry of each bond and not of the overall molecule)

The representation in the figure above is therefore called the σ bond framework of the ethylene molecule. If we let this framework be in the *x-y* plane, the carbon atomic orbitals used to construct the hybrid orbitals would have been the $2p_x$ and the $2p_y$ orbitals. This leaves the two carbon $2p_z$ orbitals available for further bonding, and their overlap can contribute to the C-C bond, as shown below:

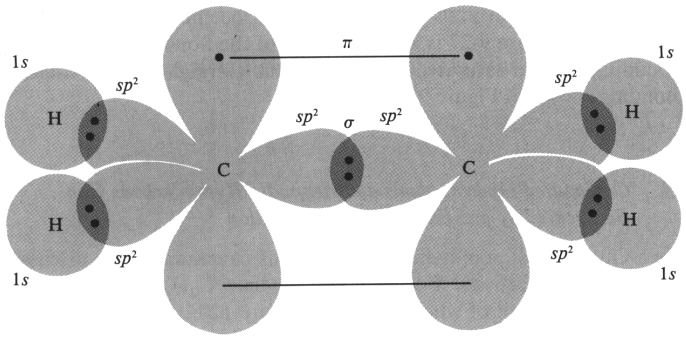


Figure 9-29. A schematic representation of the π bond in the ethylene molecule.

The charge distribution along the C-C bond due to the overlap of the $2p_z$ orbitals produces a π bond.

We are developing here a σ–π description of unsaturated hydrocarbons. It turns out to be a fairly good approximation to treat the π electrons as moving in a fixed, effective, electrostatic potential due to the electrons in the σ framework. This approximation is called the ***π-electron approximation.*** It can be developed from the Schrödinger equation, but we will simply accept it as being physically intuitive.

In 1930, Eric Hückel developed a treatment of conjugated and aromatic molecules that has found wide success in organic chemistry. This theory, which is referred to as ***Hückel molecular orbital theory*** , is based upon the π-electron approximation. The σ electrons are described as localized hybrid bond orbitals and the π electrons are described by molecular wave functions that extend over each of the atoms that contribute a π electron. Thus, the π electrons are delocalized.

I will illustrate this technique by considering some specific examples.

11.9.1 Ethylene

Let us first consider the simple case of ethylene.

The σ framework of ethylene is shown on the previous pages. Each carbon atom contributes a $2p_z$ orbital to the delocalized π orbital, and as in the case of the molecular orbital treatment of H_2 we write

$$\psi_\pi = c_1\chi_1 + c_2\chi_2$$

where χ_1 and χ_2 are the carbon $2p_z$ orbitals. It is important to realize that the Hamiltonian operator in this theory involves the effective potential due to the electrons in the σ framework of the molecule and so itself is an effective Hamiltonian. A principal advantage of Hückel theory is that it is not necessary to ever specify this effective Hamiltonian.

The secular determinantal equation associated with the molecular orbital above is

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{12} - ES_{12} & H_{22} - ES_{22} \end{vmatrix} = 0$$

where the H_{ij} are integrals involving the effective Hamiltonian. Because the carbon atoms in ethylene are equivalent, $H_{11} = H_{22}$.

These diagonal elements of the secular determinant, called *Coulomb integrals*, are customarily denoted by α.

The off-diagonal H's in the secular determinant are called *resonance integrals* or *exchange integrals* and are customarily denoted by β. Note that β is a two-center integral because it involves the atomic orbitals from two different carbon atoms. Although it is not necessary, one often neglects the overlap integrals in Hückel theory and so the S_{ij} are given by

$S_{ij} = 0$ if $i \neq j$
$= 1$ if $i = j$

Thus, the Hückel secular determinantal equation describing the ethylene molecule is

$$\begin{vmatrix} \alpha - E & \beta \\ \beta & \alpha - E \end{vmatrix} = 0$$

The two roots of this secular determinant are:

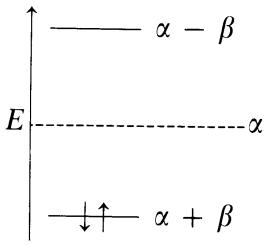
$$E = \alpha \pm \beta.$$

In principle, to evaluate α and β we would have to know the effective Hamiltonian operator. However, we do not have to do this in Hückel theory because α and β are assigned empirical values.

Because α is essentially the energy of an electron in an isolated carbon $2p_z$ orbital, we can use it to set our zero of energy. The quantity β has been determined from a consideration of a variety of data and can be assigned a value of approximately -75 kJ mol^{-1}.

There are two π electrons in ethylene. In the ground state, both electrons occupy the lowest energy orbital. Because β is negative, the lowest energy is $E = \alpha + \beta$.

An energy-level diagram showing the ground state of ethylene is shown below:



The π electronic energy of ethylene is simply the sum of the energies of the two π electrons and is given by $E = 2\alpha + 2\beta$. But because α is used to specify the zero of energy, the π electronic energy of ethylene is simply 2β.

Because α is essentially the energy of an isolated *p* orbital, the two energies, $E = \alpha \pm \beta$ must correspond to bonding and antibonding orbitals.

Let us determine the wave functions for the Hückel molecular orbitals.

Recall that the secular determinantal equation originates from the pair of linear algebraic equations for the expansion coefficients c_1 and c_2

$$(H_{11} - ES_{11})c_1 + (H_{12} - ES_{12})c_2 = 0$$

$$(H_{12} - ES_{12})c_1 + (H_{22} - ES_{22})c_2 = 0$$

where c_1 and c_2 came from our trial wave function in which we approximated the orbitals as linear combinations of π orbitals on the two atoms.

Upon using the Hückel approximations for the H and S, we have

$$(\alpha - E)c_1 + \beta c_2 = 0$$

$$\beta c_1 + (\alpha - E)c_2 = 0$$

To find the *c*'s associated with each value of *E*, we substitute one value of *E* into either of the two linear algebraic equations above.

For example, for the value $E = \alpha + \beta$, either equation yields $c_1 = c_2$, so that

$$\psi_1 = c_1(\chi_1 + \chi_2)$$

The value of c_1 is found by requiring that ψ_1 be normalized. Because we are using $S_{12} = 0$, we find that $c_1 = \frac{1}{\sqrt{2}}$.

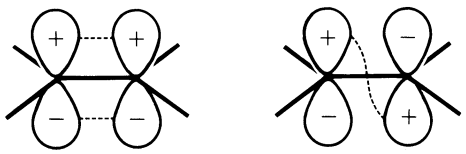
Thus,

$$\psi_1 = \frac{1}{\sqrt{2}}(\chi_1 + \chi_2)$$

It is not difficult to show that the root $E = \alpha - \beta$ yields

$$\psi_2 = \frac{1}{\sqrt{2}}(\chi_1 - \chi_2)$$

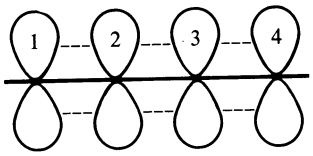
The figure below shows the two molecular orbitals schematically:



Note that it would not be difficult to extend this treatment to the case where we do not assume that the overlap integrals, S_{ij} to be zero for i≠j.

11.9.2 Butadiene

The case of butadiene is more interesting than that of ethylene. Although butadiene exists in cis or trans configuration, we will picture this molecule as simply a linear sequence of four carbon atoms, each of which contributes a $2p_z$ orbital to a π-electron orbital.



Because we have a linear combination of four atomic orbitals, we are going to have a 4 x 4 secular determinant, four different energies, and four different π-molecular orbitals.
We can introduce the notation

$$\psi_i = \sum_{n=1}^{4} c_{in}\chi_n$$

where c_{in} is the coefficient of the atomic orbital of the n^{th} atom in the i^{th} molecular orbital.

The secular determinantal equation for the butadiene molecule is

$$\begin{vmatrix} H_{11}-ES_{11} & H_{12}-ES_{12} & H_{13}-ES_{13} & H_{14}-ES_{14} \\ H_{12}-ES_{12} & H_{22}-ES_{22} & H_{23}-ES_{23} & H_{24}-ES_{24} \\ H_{13}-ES_{13} & H_{23}-ES_{23} & H_{33}-ES_{33} & H_{34}-ES_{34} \\ H_{14}-ES_{14} & H_{24}-ES_{24} & H_{34}-ES_{34} & H_{44}-ES_{44} \end{vmatrix} = 0$$

Because we are taking the four carbon atoms in the butadiene molecule to be equivalent, all the H_{ii} in this determinant are equal, and as in the case of ethylene, we denote them by α.

The H_{ij} , on the other hand, are two-center integrals. They involve the $2p_z$ orbital centered on carbons *i* and *j*. In the simplest version of Hückel theory, one sets $H_{ij} = \beta$ if the i and j carbon atoms are adjacent and $H_{ij} = 0$ if they are not adjacent. The justification of this is that the overlap of the $2p_z$ orbitals from two carbon atoms decreases with their separation. Following this argument, one might set $S_{ij} = S$ for adjacent carbon atoms and $S_{ij} = 0$ for nonadjacent carbon atoms. In the simplest version of Hückel theory, one goes even one step further and sets $S = 0$.

Under these approximations and assumptions, the Hückel theory secular determinantal equation for butadiene becomes

$$\begin{vmatrix} \alpha - E & \beta & 0 & 0 \\ \beta & \alpha - E & \beta & 0 \\ 0 & \beta & \alpha - E & \beta \\ 0 & 0 & \beta & \alpha - E \end{vmatrix} = 0$$

If we factor β from each column and let $x = (\alpha - E)/\beta$, then the determinant above becomes

$$\begin{vmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} = 0$$

If we expand this determinant, then the secular equation is

$$x^4 - 3x^2 + 1 = 0$$

We can solve this equation for x^2 to obtain

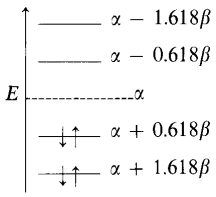
$$x^2 = \frac{3 \pm \sqrt{5}}{2}$$

We therefore find the four roots to be

$$x = \pm 1.61804$$

$$x = \pm 0.61804$$

Recalling that $x = (\alpha - E)/\beta$ and that β is a negative quantity, we can construct a Hückel theory energy-level diagram for butadiene. There are four π electrons in butadiene. In the ground state, these four π electrons occupy the two orbitals of lowest energy as shown below



The total π electronic energy of butadiene is

$$E_\pi = 2(\alpha + 1.618\beta) + 2(\alpha + 0.618\beta)$$

$$= 4\alpha + 4.472\beta$$

It is interesting to compare the energy given in this equation to the energy of the localized structure in which the two double bonds are localized between carbon atoms 1 and 2 and carbons atoms 3 and 4 in butadiene.

In the simple Hückel theory, this localized structure is equivalent to two isolated ethylene molecules. We have shown above that $E_\pi = 2\alpha + 2\beta$ for ethylene, and so we can define a delocalization energy by

$$E_D = E_\pi(butadiene) - 2E_\pi(ethylene)$$

$$= 0.472\beta$$

If β is given the value -75 kJ mol^{-1}, then we see that the delocalization energy in butadiene is about -35 kJ mol^{-1}. This is the energy by which butadiene is stabilized relative to two isolated double bonds, or in other words, the stability that butadiene derives because its π electrons are delocalized over the entire length of the molecule instead of being localized to the two end bonds.

Associated with each of the four molecular orbital energies of butadiene is a wave function given by the expansion coefficients.

Recall that the molecular orbital wave functions are given by

$$\psi_i = \sum_{n=1}^{4} c_{in}\chi_n$$

where the c_{in} are determined by the set of linear algebraic equations that lead to the secular determinantal equation. The algebra is a little bit longer than in the case of ethylene although straightforward. The resulting wave functions are

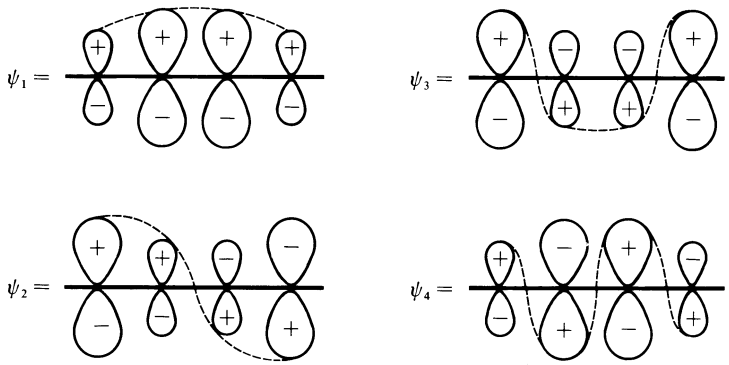
$$\psi_1 = 0.3717\chi_1 + 0.6015\chi_2 + 0.6015\chi_3 + 0.3717\chi_4$$

$$\psi_2 = 0.6015\chi_1 + 0.3717\chi_2 - 0.3717\chi_3 - 0.6015\chi_4$$

$$\psi_3 = 0.6015\chi_1 - 0.3717\chi_2 - 0.3717\chi_3 + 0.6015\chi_4$$

$$\psi_4 = 0.3717\chi_1 - 0.6015\chi_2 + 0.6015\chi_3 - 0.3717\chi_4$$

These wave functions are presented schematically below:



Notice that the energy increases as the number of nodes increases.

Because we have set $S_{ij} = \delta_{ij}$, we have in effect assumed that the χ_n are orthonormal. Using this fact, one can see that

$$\sum_{n=1}^{4} c_{in}^2 = 1$$

This allows us to interpret c_{in}^2 as the fractional π electronic charge on the n^{th} carbon atom due to an electron in the i^{th} molecular orbital. Thus, the total π electronic charge on the n^{th} carbon atom is

$$q_n = \sum_i n_i c_{in}^2$$

Where n_i is the number of π electrons in the i^{th} molecular orbital. For butadiene, you will find that all the *q*'s are 1, indicating that the *p* electrons are uniformly distributed over the molecule. You can see this schematically by putting 2 electrons in each of the first two molecular orbitals and adding the electron probabilities.

One can also use Hückel theory to define the π bond order. We can interpret the product $c_{ir}c_{is}$ as the π electron charge in the ith molecular orbital between the adjacent carbon atoms *r* and *s*. The π bond order can be defined as:

$$P_{rs}^{\pi} = \sum_i n_i c_{ir} c_{is}$$

Where n_i is the number of π electrons in the i^{th} molecular orbital.

For butadiene, one gets

$$\begin{aligned} P_{12}^{\pi} &= 2c_{11}c_{12} + 2c_{21}c_{22} + 0c_{31}c_{32} + 0c_{41}c_{42} \\ &= 2(0.3717)(0.6015) + 2(0.6015)(0.3717) \\ &= 0.8942 \end{aligned}$$

$$\begin{aligned} P_{23}^{\pi} &= 2c_{12}c_{13} + 2c_{22}c_{23} \\ &= 2(0.6015)(0.6015) + 2(0.3717)(-0.3717) \\ &= 0.4473 \end{aligned}$$

You can see by symmetry that

$$P_{12}^{\pi} = P_{34}^{\pi}$$

Remember that this is the π bond order. To find the total bond order, we must remember that there is a σ bond between each carbon atom. So we can write:

$$P_{rs} = 1 + P_{rs}^{\pi}$$

For butadiene, this leads to

$$P_{12} = P_{34} = 1.894$$

$$P_{23} = 1.447$$

These values are in good agreement with the relative reactivity of these bonds.

12 Electronic Structure Calculations