An emergent reality, part 2: Quantum Mechanics

Randall S. Dumont

Aug. 29, 2017

Contents

Chapter 1

Before quantum mechanics

Imagine a time when it was finally accepted that the world consisted of atoms - each atom with a positively charged nucleus surrounded by negatively charged particles. These particles are subject to electromagnetic forces. Maxwell's equations of electromagnetism are known. It is also known that these equations explain light as an electromagnetic wave. It is 1911, and the Rutherford model of the atom has just been established.

There are *particles* and *waves*, all subject to *equations of motion.* The particles are thought to be described by *classical mechanics* - *Newton's laws* - equivalently, *Lagrange equations* or *Hamilton's equations.* The latter are subsequent reformulations of Newton's laws.

The forces (electromagnetic and gravity), that are responsible for most of what happens on the surface of this planet, were known at this snapshot in human history. In principle, if the exact state of a system is known at a specific time, then the entire future of the system can be determined - at least according to classical mechanics.

1.1 Classical mechanics

Newton's second law is expressed by

$$\text{Force} = \text{mass} \times \text{acceleration}.$$

The first law is really just the special case of no net force. Newton's third law is just the extension of the second law to a pair of particles (bodies).

For a system of N particles, all with the same mass m, the position of all particles is specificied by a $3N$ dimensional vector, $\mathbf{x}$. The components of $\mathbf{x}$ are the x, y and z spatial coordinates for each particle. Newton's second law takes the form,

$$\mathbf{F} = m\frac{d^2}{dt^2}\mathbf{x}. \tag{1.1}$$

This is a *second order* differential equation in time. To obtain a *unique solution* - for all time - requiries specifying initial position, $\mathbf{x}(0)$, and initial velocity, $\mathbf{v}(0)$;

$$\mathbf{v}(t) = \frac{d}{dt}\mathbf{x}(t).$$

According to the classical mechanics, we can uniquely *predict the future* - given the initial state of the system. The classical state of the system consists of specific positions and *momenta* (velocities), $\mathbf{x}$ and $\mathbf{p} = m\mathbf{v}$.

1.1.1 Hamilton's formulation of classical mechanics

Momentum, as opposed to velocity, is used in Hamilton's formulation of classical mechanics. This formulation begins with the construction of the system Hamiltonian, $H(\mathbf{x}, \mathbf{p})$, as a function of position and momentum - the Hamiltonian is the energy function. The Hamiltonian consists of *kinetic* plus *potential* energy;

$$H(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2m} + V(\mathbf{x}), \tag{1.2}$$

where

$$p^2 = \mathbf{p}^T\mathbf{p}$$

is the square magnitude of the momentum vector, and $V(\mathbf{x})$ is the potential energy which depends only on the positions of the particles - the *configuration.*

The differential equatons in Hamilton's formulation are first order. Hamilton's equations are given by

$$\frac{d}{dt}\mathbf{x}(t) = \frac{\partial H}{\partial \mathbf{p}} \tag{1.3}$$

and

$$\frac{d}{dt}\mathbf{p}(t) = -\frac{\partial H}{\partial \mathbf{x}}. \tag{1.4}$$

For the above Hamiltonian, these equation take the form,

$$\frac{d}{dt}\mathbf{x} = \frac{\mathbf{p}}{m}$$

and

$$\frac{d}{dt}\mathbf{p} = -\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{F}(\mathbf{x}).$$

The first of these equations is just the definition of momentum, while the second equation is Newton's second law - where the force is gven by the gradient of the potential energy function. Hamilton's formulation is equivalent to Newton's formulation. Its advantage arises from the equations being first order, rather than second order.

A unique solution to the first order differential equation in time,

$$\frac{d}{dt}\begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{p}/m \\ \mathbf{F}(\mathbf{x}) \end{pmatrix},$$

results from a given initial state of the system, $\begin{pmatrix} \mathbf{x}(0) \\ \mathbf{p}(0) \end{pmatrix}$. This system has $3N$ *degrees of freedom*. Each degree of freedom has a position and momentum - the components of $\mathbf{x}$ and $\mathbf{p}$.

1.2 The trouble with classical mechanics

Classical mechanics works well for macroscopic bodies - e.g., baseballs, cars, planes, machine components. But it does not work for systems of elementary particles. Classical mechanics was already in trouble before the atomic nature of matter was universally accepted.

1.2.1 The ultraviolet catastrophe

The statistical theory of systems of large numbers of particles, *classical statistical mechanics*, was known in 1911. In particular, the *equipartition theorem* was known. It states that, for a system in thermal equilibrium, each unbound degree of freedom has energy, $\frac{1}{2}k_{\mathrm{B}}T$, while each bound degree of freedom has energy, $k_{\mathrm{B}}T$. When this theorem is applied to the electromagnetic waves inside a box in thermal equilibrium, there is a startling result.

In the classical theory of light, light is an electromagnetic wave. Specifically, light consists of waves in electric and magnetic fields at right angles. The components of the electric (and magnetic) field(s) satisfy the wave equation,

$$\frac{d^2}{dx^2}\mathcal{E}(x,y,z,t)+\frac{d^2}{dy^2}\mathcal{E}(x,y,z,t)+\frac{d^2}{dx^2}\mathcal{E}(x,y,z,t)=\frac{1}{u^2}\frac{d^2}{dt^2}\mathcal{E}(x\,y,z,t)\,. \quad (1.5)$$

The following sine wave satisfies this equation:

$$\mathcal{E}(x,y,z,t)=\mathcal{E}_0\sin\left(k_x x+k_y y+k_z z-\omega t\right), \quad (1.6)$$

where

$$k_x^2+k_y^2+k_z^2=k^2=\frac{\omega^2}{u^2}. \quad (1.7)$$

Here, u is the velocity of the wave,

$$u=\frac{c}{n},$$

where n is the refractive index and $c = 2.99792458 \times 10^8$ m s^{-1} is the speed of light in a vacuum. The wave vector,

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

is inversely proportion to the wavelength, λ. $\hat{\mathbf{k}}$ is a unit vector giving the direction of wave propagation.

$$\omega=2\pi\nu$$

is the angular frequency.

For light constrained to a finite volume (consider a cube), there is a *discrete* set of possible values of k_x, k_y and k_z, corresponding to distinct standing waves of light. Each possible set of the values determines a degree of freedom, or *mode*, of light. The number of modes consistent with frequency, $\nu = \omega/2\pi$, is proportional to ν^2. This results because the wave vector components satisfy Eq. 1.7 which is the equation for the surface of a sphere of radius, ω/u. The number of modes with frequency, ν, is proportional to the area of that sphere which increases in proportion to ν^2.

Consider light in thermal equilibrium with the walls of the box, at temperature T. This light is called *blackbody radiation.* According to the equipartition theorem, the amount of energy stored in the light is infinite. In particular, the amount of energy in blackbody radiation at frequency, ν, increases as ν^2. This is called the *ultraviolet catastrophe.*

The problem was solved by Planck who modified the equipartition theorem, replacing the fixed energy, $k_\mathrm{B}T$, by a function - equal to $k_\mathrm{B}T$ for small ν, and decreasing exponentially with ν when ν is large. The resulting blackbody spectrum - i.e., the distribution of energy over frequency - matches experimental observations of blackbody radiation (e.g., the light from the sun, or any other star). This modification required introducing Planck's constant, $h = 6.62606957 \times 10^{-34}$. It was Einstein who interpreted this modification as accounting for the *quantization* of light energy. Specifically, light consists of particles of light, photons, with energy equal to $h\nu$. Light has wave and particle properties.

The *Planck distribution* or *blackbody spectral emissivity* - the power emitted, at frequency ν, by an object at temperature T, per unit frequency and unit surface area (of the object) - is given by

$$\begin{aligned} \mathsf{E}(\nu) &= \frac{2\pi h\nu^3}{c^2}\frac{1}{e^{h\nu/k_\mathrm{B}T}-1} \\ &= \frac{2\pi h\nu^3}{c^2}\frac{e^{-h\nu/k_\mathrm{B}T}}{1-e^{-h\nu/k_\mathrm{B}T}}. \end{aligned} \tag{1.8}$$

Blackbody *emissivity* - the total power emitted (over all frequencies) per unit surface area - is the integral of $\mathsf{E}(\nu)$ over all frequencies.

$$\begin{aligned}\int_0^\infty \mathsf{E}(\nu)\, d\nu &= \frac{2\pi h}{c^2}\int_0^\infty \frac{\nu^3 e^{-h\nu/k_{\mathrm{B}}T}}{1-e^{-h\nu/k_{\mathrm{B}}T}}\, d\nu \qquad (1.9)\\ &= \frac{2\pi h}{c^2}\left(\frac{k_{\mathrm{B}}T}{h}\right)^4\int_0^\infty \frac{x^3e^{-x}}{1-e^{-x}}\, dx\\ &= \frac{2\pi k_{\mathrm{B}}^4}{h^3c^2}T^4\left[\int_0^\infty x^3e^{-x}\, dx + \int_0^\infty x^3 e^{-2x}\, dx + \ldots\right]\\ &= \frac{2\pi k_{\mathrm{B}}^4}{h^3c^2}T^4\left[\int_0^\infty x^3e^{-x}\, dx + \frac{1}{2^4}\int_0^\infty x^3 e^{-x}\, dx + \ldots\right]\\ &= \frac{2\pi k_{\mathrm{B}}^4}{h^3c^2}T^4\left[6+\frac{6}{2^4}+\frac{6}{3^4}+\ldots\right]\\ &= \frac{2\pi^5 k_{\mathrm{B}}^4}{15h^3c^2}T^4 = \sigma T^4\end{aligned}$$

The T^4 dependence of emissivity is the *Stefan-Boltzmann law.* $\sigma = 5.67\times10^{-8}$ W m^{-2} K^{-4} is called the *Stefan-Boltzmann constant.* The frequency of peak spectral emissivity is given by setting

$$\frac{d\mathsf{E}(\nu)}{d\nu}=0$$

or

$$\begin{aligned}0 &= \frac{d}{dx}\frac{x^3}{e^x-1}\\ &= \frac{3x^2\left(e^x-1\right)-x^3e^x}{\left(e^x-1\right)^2},\end{aligned}$$

where $x = h\nu/k_{\mathrm{B}}T$. Since the denominator and x must be positive, this expression simplifies to

$$0 = 3\left(e^x-1\right)-xe^x,$$

or

$$x = 3\left(1-e^{-x}\right).$$

This is a nonlinear equation which has solutions at $x=0$ ($\rho(\nu)$ is a minimum at zero frequency) and $x = 2.822$. The peak frequency is given by

$$\begin{aligned}\nu &= 2.822 k_{\mathrm{B}}T/h \qquad (1.10)\\ &= 5.88\times10^{10}\ \text{Hz K}^{-1}.\end{aligned}$$

When expressed in terms of wavelength, this result is known as *Wien's law.*

Figure 1.1 shows Planck distributions for T from 200 to 800 K. For this temperature range, blackbody emissivity varies from 90 W m^{-2} to 23 kW m^{-2}.

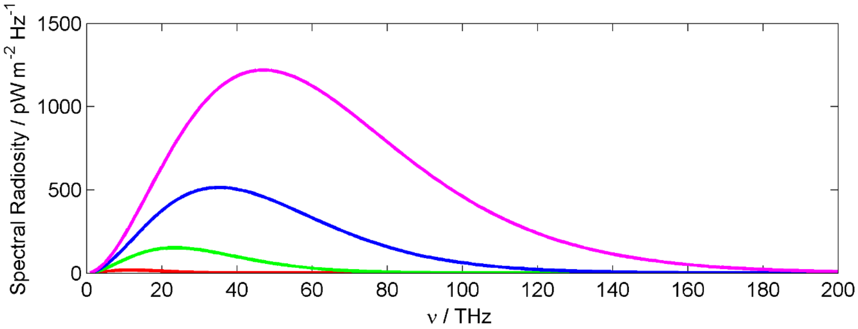


Figure 1.1: Spectral emissivity of blackbodies at different temperatures. The spectral radiosity - power emitted per unit area of the blackbody, per unit frequency of light emitted - in units of pW $\mathrm{m}^{-2}\mathrm{Hz}^{-1}$ is shown for $T = 200$, 400, 600 and 800 K. The area under the curve increases with increasing temperature as - T^4 - while the peak shifts to higher frequency.

1.2.2 The stability of atoms, and atomic spectra

An atom consists of a positively charged nucleus, carrying most of the mass of the atom, surrounded by enough electrons to balance the charge. According to classical mechanics, electrons orbit the nucleus much as planets, asteroids and comets orbit the sun. However, because the electrons are charged and they are accelerated (centripetal acceleration), they should emit electromagnetic waves (i.e., light). As the electrons emit light, they lose energy. Electrons should, according to classical mechanics, eventually fall into the nucleus. The *classical atom* is *unstable*. A ground state hydrogen atom would collapse in 16 ps [J.D. Olsen and K.T. McDonald, *Classical Lifetime of a Bohr Atom*, http://www.physics.princeton.edu/~mcdonald/examples/orbitdecay.pdf].

Another problem with the classical atom, as if instability were not enough, is that the energy of the atom varies continuously and any change in energy is possible. Since $h\nu = \Delta E$, the classical atom can absorb or emit any frequency of light. The classical atom has a *continuous spectrum*. However, a real atom has a *discrete spectrum* - i.e., there is a discrete set of frequencies of light that can be absorbed or emitted.

Bohr tried to solve this problem for the hydrogen atom by positing that the angular momentum of the orbiting electron must be a multiple of $\hbar = h/2\pi$. De Broglie later interpreted this model as requiring that the allowed orbits correspond to standing electron waves, where the electrons have wavelength given by

$$\lambda_{\mathrm{deB}} = \frac{h}{p}, \tag{1.11}$$

where p is the electron momentum.

The Bohr model of the hydrogen atom reproduces the frequencies actually emitted or absorbed by hydrogen atoms - specifically, the Lyman, Balmer, Paschen, Backett, ... spectral series. The Schrödinger equation predicts the same series for hydrogen, and determines every other kind of spectrum of any atom or molecule. The key feature of electrons, according to *quantum mechanics*, is that they have both wave and particle properties. Like light, matter is both particle and wave.

1.2.3 Wave property of matter particles

In 1927, Davisson and Germer observed *diffraction* of a *low energy* electron beam reflected from a crystal of nickel. The pattern was determined by the spacing of nickel atoms at the surface of the crystal, and the de Broglie wavelength of the electrons in the beam. A simplified *low energy electron diffraction* (LEED) setup is shown in Figure 1.2.

Figure 1.3 shows an observed LEED diffraction pattern for a Si(100) crystal. The structure of the surface layers of the crystal is manifest in the pattern of peaks.

Electrons *diffract* in the same manner as X-rays. Electron hit the screen at diffraction peaks. Diffraction is a *wave property.* For specular reflection (angle of reflection equals angle of incidence) - when the *Bragg condition* is met (it is usually written in terms of the sine of the complementary angle);

$$n\lambda = 2a\cos\theta. \tag{1.12}$$

Here, a is the atomic spacing, θ is the angle of incidence (relative to normal), λ is the wavelength of the electrons in the beam (all the same, as the electrons all have the same momentum), and $n = 1, 2, 3, \ldots.$

In general, for a crystal with three lattice vectors (these are the three sides of a unit cell of the crystal), **a**, **b** and **c**, the von Laue equations determine the diffraction peaks for any angle of reflection - i.e., not necessarily specular:

$$\begin{aligned} \frac{n_a\lambda}{a} &= \cos\theta_{a,\mathrm{i}} + \cos\theta_{a,\mathrm{r}} \qquad (1.13) \\ \frac{n_b\lambda}{b} &= \cos\theta_{b,\mathrm{i}} + \cos\theta_{b,\mathrm{r}} \\ \frac{n_c\lambda}{c} &= \cos\theta_{c,\mathrm{i}} + \cos\theta_{c,\mathrm{r}}, \end{aligned}$$

where n_a, n_b and n_c are independent positive integers: $1, 2, 3, \ldots$ The angles correspond to the angles depicted in Fig. 1.3, relative to the crystal planes indicated - a, b or c.

Diffraction has been observed for particles heavier than electrons - e.g., He,[I. Estermann, O. Stern: Z. Phys. 61, 95 (1930)] and even large molecules such as Buckminsterfullerene [B. Brezger, L. Hackermüller, S. Uttenthaler, J. Petschinka, M. Arndt, A. Zeilinger: Phys. Rev. Lett. 88, 100404 (2002)].

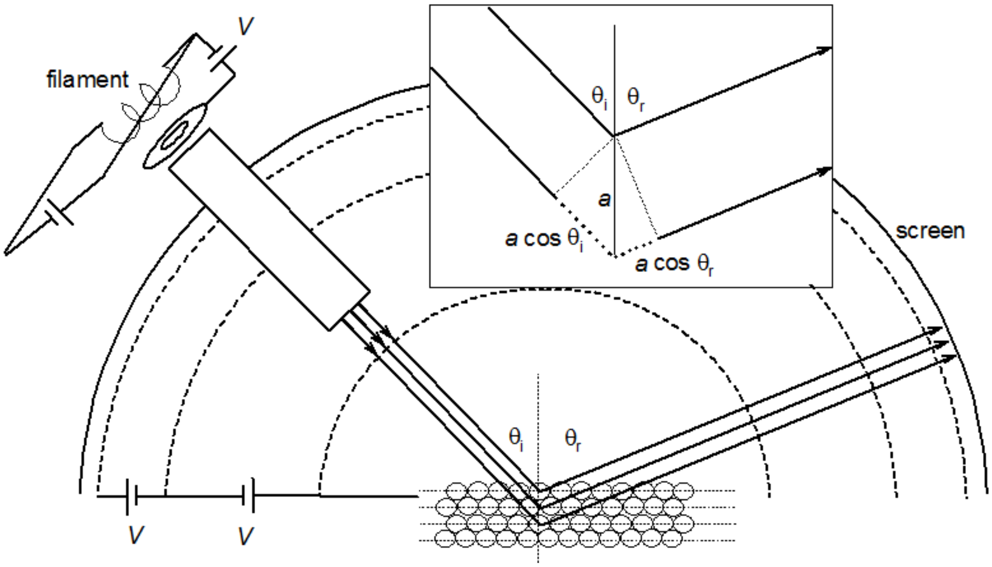


Figure 1.2: A low energy electron diffraction (LEED) experimental setup. Current flows through the thin filament, making it hot - the filament has high resistance (it is like an incandescent light-bulb filament). Electrons emitted by the hot filament are accelerated through a potential difference, $V = 20$ to 200 V, and collimated into a fixed velocity electron beam velocity. Electrons impact a crystal with angle of incidence, θ_{i}, and reflects at angles θ_{r} (only one reflection is angle shown here). Grids between sample and screen filter out electrons with less energy than the incoming electrons (they have energy, eV). The inner grid shields the sample from the electric field induced by the filter. Phosphors in the screen emit visible light when they are strck by electrons. The inset shows the extra distance traveled by electrons that reflect off of the second layer of atoms - in comparison to the first layer.

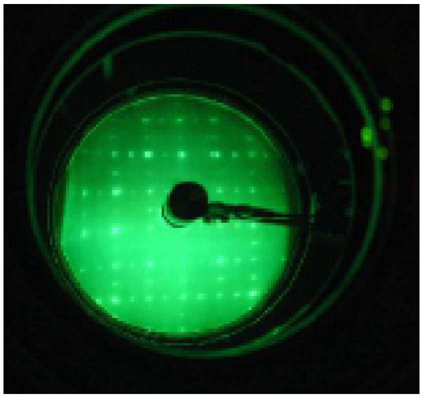


Figure 1.3: The LEED pattern of diffraction peaks for a crystal of silicon cut with the 100 surface exposed. The 100 designation gives the components of a vector normal to the crystal surface, within the standard crytallographic unit cell. Si(100) consists of the front face of the unit cell - only y and z values vary across the surface. Diffraction peaks appear at angles satifying the von Laue equations. From [http://en.wikipedia.org/wiki/Low-energy_electron_diffraction# /media/File:Si100Reconstructed.png. Retrieved Aug. 28, 2014.]

Example 1.1. Electron diffraction peaks. Consider a beam of electrons, accelerated through a potential of 100. V, incident on a crystal of nickel with $\theta_{a,\mathrm{i}} = 60°$. Nickel is face centered cubic. As such, it has just one lattice parameter; $a = b = c$. $a = 0.352$ nm for nickel.

(a) What is the de Broglie wavelength of electrons in the beam?

(b) Determine the angles of reflection of the two diffraction peaks closest to specular reflection, $\theta_{b,\mathrm{r}} = \theta_{b,\mathrm{i}} = 60°$.

(c) How many diffraction peaks will be observed? Note that only reflected beams - i.e., $-\frac{\pi}{2} < \theta_{b,\mathrm{r}} < \frac{\pi}{2}$ - can be seen.

(d) If the electron gun voltage were 200.0 V, rather than 100. V, how many diffraction peaks would be seen?

Solution 1.1.

(a) The kinetic energy of an electron accelerated through 100. V is

$$\begin{aligned} E &= eV = 100.\ \mathrm{eV} \\ &= \frac{mu^2}{2} = \frac{p^2}{2m}, \end{aligned}$$

since $p = mu$, where u is the velocity of the electron. Therefore, the electron momentum is

$$\begin{aligned} p &= \sqrt{2mE} = \sqrt{2 \times 9.109 \times 10^{-31}\ \text{kg} \times 1.602 \times 10^{-19}\ \text{C} \times 100.\ \text{V}} \\ &= 5.40_2\ \times 10^{-24}\ \text{kg m s}^{-1}. \end{aligned}$$

The electron velocity is $5.930\,8 \times 10^6$ m s^{-1}. Since this is well below the speed of light, the use of the above non-relativistic kinetic energy formula is acceptable.

The de Broglie wavelength is

$$\begin{aligned} \lambda &= \frac{h}{p} = \frac{6.626 \times 10^{-34}\ \text{J s}}{5.40_2\ \times 10^{-24}\ \text{kg m s}^{-1}} \\ &= 1.22_7 \times 10^{-10}\ \text{m} \\ &= 0.122_7\ \text{nm}. \end{aligned}$$

(b) Since $\cos\theta_{a,\text{i}} = \frac{1}{2}$, and the von Laue equation (there is only one for nickel) takes the form,$3 \times 0.348\,58 - \frac{1}{2} = 0.545\,74$

$$\begin{aligned} \cos\theta_{a,\text{r}} &= \frac{n_a\lambda}{a} - \frac{1}{2} \\ &= n_a\frac{0.122_7}{0.352} - \frac{1}{2} \\ &= n_a \times 0.348_{\,6} - \frac{1}{2}. \end{aligned}$$

For specular reflection, $\cos\theta_{a,\text{r}} = \frac{1}{2}$. This corresponds to $n_a = 1/0.348_6 = 2.87$. Since this is not an integer, there is no diffraction peak. The closest diffraction peaks occur at $n_a = 2$ and 3;

$$\cos\theta_{a,\text{r}} = \begin{cases} 0.197_2 & n_a = 2 \\ 0.545_{\,7} & n_a = 3 \end{cases}$$

The associated angles are $\cos^{-1}(0.197_2) = 78.6°$ and $\cos^{-1}(0.545_7) = 56.9°$.

(c) For angles in the observed hemisphere, the electron beam is reflected and $\cos\theta_{a,\text{r}} > 0$. For $n_a = 1$, the von Laue equation produces a negative $\cos\theta_{a,\text{r}}$ (specifically, $-0.151\,4$) corresponding to an angle of reflection of $\cos^{-1}(-0.151\,4) = 98.7°$. The $n_a = 1$ actually corresponds to transmission. However, unless the sample is extremely thin, this transmission beam cannot be seen. *Electron transmission microscopy* is based on observing transmitted electrons through ultra-thin samples. Here, the $n_a = 1$ peak cannot be seen. The peaks begin with $n_a = 2$, then 3. Since $\cos\theta_{a,\text{r}}$ cannot excede 1, then sequence ends with $n_a = 4$. Three diffraction peaks can be seen here.

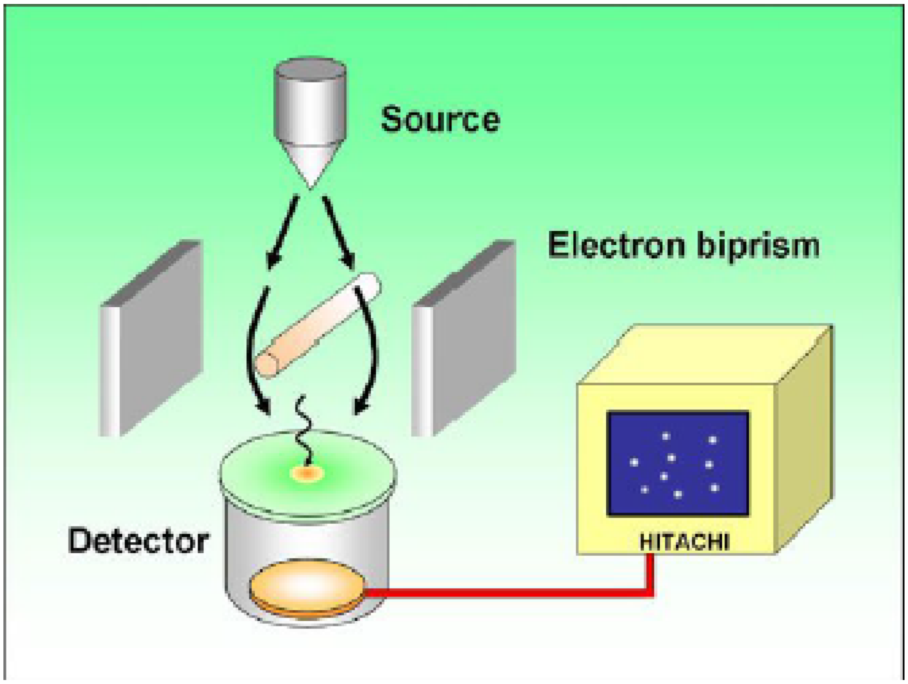


Figure 1.4: The double slit experiment. From [http://www.hitachi.com/rd/research/em/doubleslit.html. Retrieved Aug. 24, 2014.]

(d) In this case, the electron energy is twice as big. This makes the momentum larger by $2^{1/2}$, and the de Broglie wavelength smaller - it changes by the factor $2^{-1/2}$. In this case,

$$\begin{aligned}\cos\theta_{a,\mathrm{r}} &= \frac{n_a\lambda}{a} - \frac{1}{2} \\ &= n_a\frac{2^{-1/2}\times 0.122_7}{0.352} - \frac{1}{2} \\ &= n_a \times 0.246\,_5 - \frac{1}{2}.\end{aligned}$$

The observed diffraction peaks begin with $n_a = 3$ and end with $n_a = 6$. The upper limit is found by setting $\cos\theta_{a,\mathrm{r}} = 1$ in the von Laue equation, then solving for n_a. The result is $n_a = 6.09$. $n_a = 6$ is the largest peak index below this value. Therefore, four diffraction peaks are observed.

In a *double slit experiment*, shown in Fig. 1.4, there are two pathways to a target available to the wave. The two portions of the wave that take different pathways interfere when they arrive at the target together. When the difference in total pathlengths is a multiple of wavelength of the particles, the pathways interfere constructively and particles are observed. When the difference in total pathlengths is a half-integral multiple of the wavelength, the pathways destructively interfere and particles are not observed.

When an electron is incident on a target (detector), it is seen at a specific location. Each electron arrives at a different location. The interference effect is manifest in the distribution of these locations on the detector. The interference effect is seen even when the beam intensity is so low that there is only one electron (or none) between the source and target (detector) at any given time. In this case, the interference pattern builds up and becomes evident only after very many electrons have been observed.

Chapter 2

The emergence of quantum mechanics

Light is both wave and particle, and the components of matter are both particle and wave. In 1926, *Erwin Schrödinger* introduced the *wave equation* for matter waves. In the previous year, 1925, *Werner Heisenberg* introduced *matrix mechanics* - equivalent to Schrödinger's *quantum mechanics*.

2.1 The Schrödinger equation

Schrödinger noticed that the *Hamilton-Jacobi equation* of classical mechanics has the same form as the *eikonal equation* of *geometric optics*. The eikonal equation of geometric optics arises from Maxwell's equations (the wave equations for light) when the wavelength is small. Treating the Hamilton-Jacobi equation as an eikonal equation, Schrödinger worked out the associated wave equation - the *time independent Schrödinger equation*.

2.1.1 Geometric optics, Fermat's principle and the eikonal equation

The wavelengths of visible light - 400-700 nm - are small on the macroscopic scale. As such, for many purposes (not including diffraction and interference), visible light can be viewed as consisting of rays whose paths satisfy the *eikonal equation*,

$$\left\| \frac{\partial \varphi\left(\mathbf{x}\right)}{\partial \mathbf{x}} \right\| = k\left(\mathbf{x}\right) = \frac{\omega}{c} n\left(\mathbf{x}\right).$$

Here, $\mathbf{x}$ is a 3 dimensional vector - position in physical space; $\partial/\partial\mathbf{x}$ is the gradient. This equation follows from the wave equation for light (Eq. 1.5), where $\varphi\left(\mathbf{x}\right)$ is the phase of the wave - if the oscillations in $\varphi\left(\mathbf{x}\right)$ are much faster than the variation in the amplitude of the wave.

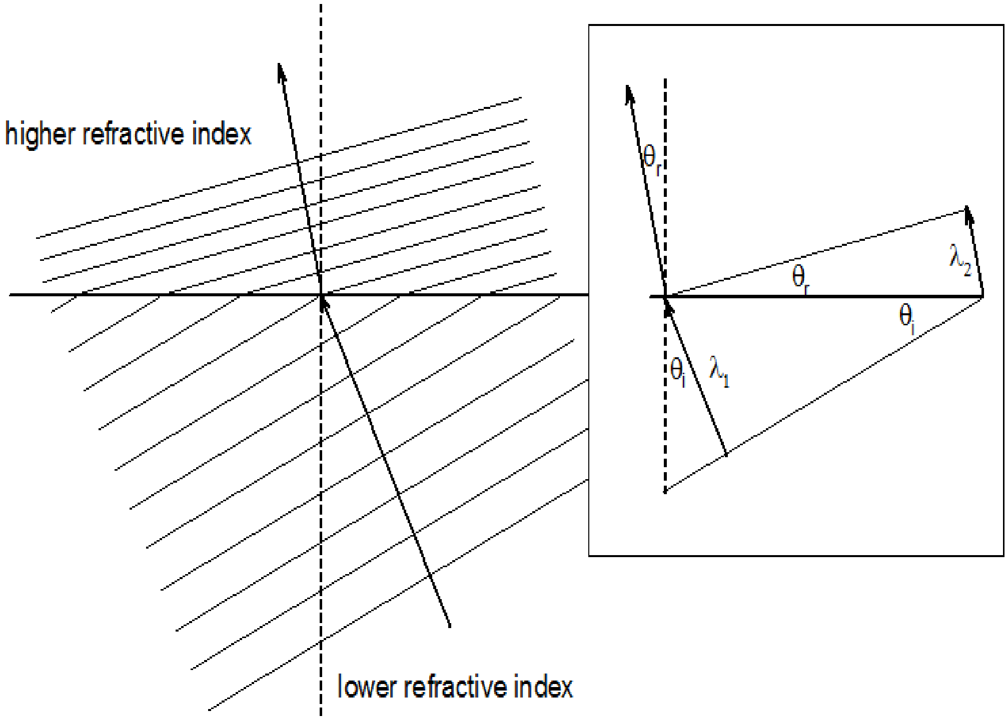


Figure 2.1: The wavefronts - lines of constant $\varphi(\mathbf{x})$, where $\varphi(\mathbf{x})$ satisfies the eikonal equation - across an interface between high and low refractive index media. The inset shows how the wavefronts satisfies Snell's law.

Solutions, $\varphi(\mathbf{x})$, to the above equation determine rays along which light travels - the rays are always perpendicular to the level surfaces of $\varphi(\mathbf{x})$. The level surfaces of $\varphi(\mathbf{x})$ can be viewed as wavefront surfaces; $\varphi(\mathbf{x})$ is the phase of the wave. All directions of propagation are possible. The eikonal equation simply fixed the magnitude of the gradient of the phase.

For homogeneous materials, the eikonal equation produces straight line rays. When refractive index changes discontinuously across a boundary, rays travel in accord with *Snell's law* - see Fig. 2.1. The associated solution to the eikonal equation consists of parallel wavefronts, with different wavelengths in the two media, connecting continuously across the boundary. This requires bending of the normal to the wavefronts - the rays - in accord with Snell's law. The rays bend at the interface so as to increase distance travelled in the material with lower refractive index - i.e., faster speed. In general, light rays satisfy *Fermat's principle of least time*: light rays travel along the path that takes the least time. To see how Fermat's principle gives rise to Snell's law rays, we consider the *lifeguard's dilema.* [Feynman's Lectures on Physics]

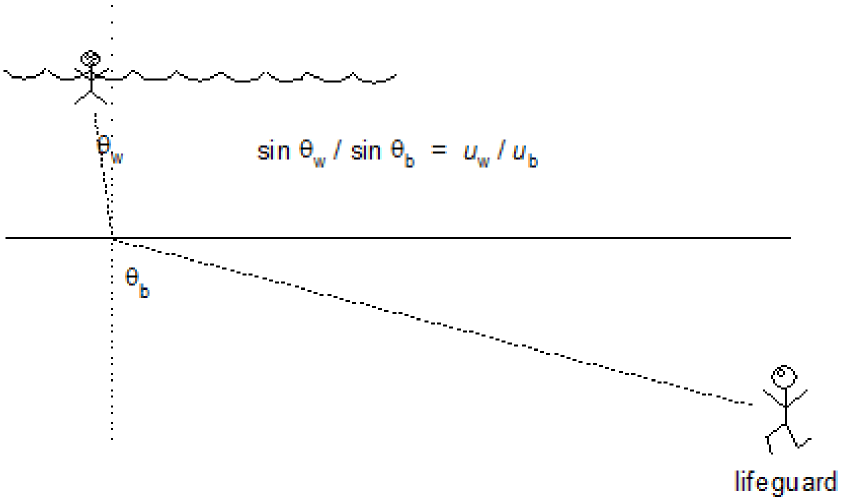


Figure 2.2: The lifeguard's dilema. The optimal path satisfies Snell's law.

The lifeguard's dilema

When a swimmer is in trouble, a lifeguard must choose the angle of approach to the shore that will minimize the time to get to the swimmer. The optimal path is in accord with Snell's law. This means keeping the swimming portion of the path almost as short as possible. The speed in water, u_{w}, is likely (excluding, e.g., rocky beaches) much smaller than the speed on the beach, u_{b}. In general, the solution to the lifeguard's dilema is the path that satisfies Snell's law.

Proof of Snells's law from Fermat's principle:

To prove Snell's law from Fermat's principle, we express the time to get from A to B - see Fig. 2.2 - in terms of a single variable, x_{w}, and then minimize with respect to x_{w}, the x component of the distance travelled in the slow medium (water in the above example).

The time to get from A to B is just

$$\begin{aligned} t &= \frac{d_{\mathrm{b}}}{u_{\mathrm{b}}} + \frac{d_{\mathrm{w}}}{u_{\mathrm{w}}} \\ &= \frac{\left((x_{\mathrm{t}} - x)^2 + y_{\mathrm{w}}^2\right)^{1/2}}{u_{\mathrm{b}}} + \frac{\left(x^2 + y_{\mathrm{w}}^2\right)^{1/2}}{u_{\mathrm{w}}}. \end{aligned}$$

The minimum of this function, with respect to $x = x_{\mathrm{w}}$, is given by setting dt/dx

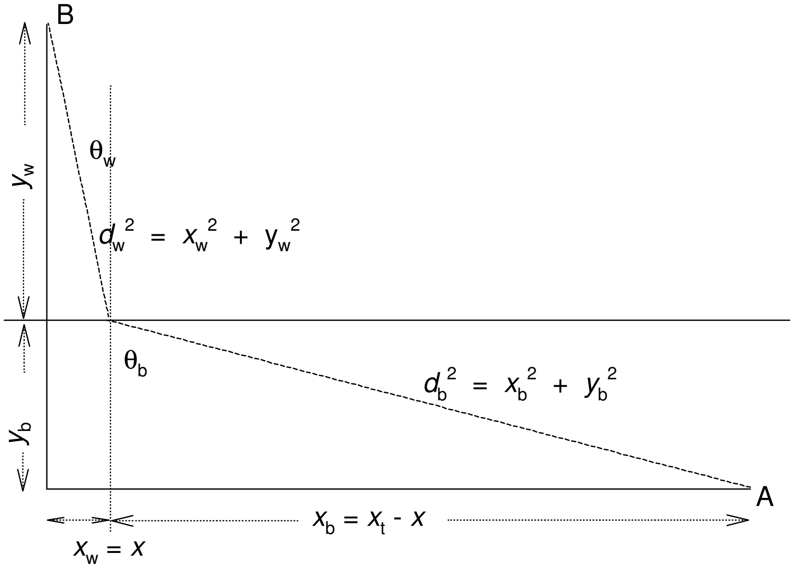


Figure 2.3: Proof of Snell's law.

to zero, and solving for x. Specifically,

$$\begin{aligned}\frac{dt}{dx} &= -\frac{(x_\mathrm{t}-x)\left((x_\mathrm{t}-x)^2+y_\mathrm{w}^2\right)^{-1/2}}{u_\mathrm{b}} + \frac{x\left(x^2+y_\mathrm{w}^2\right)^{-1/2}}{u_\mathrm{w}} \\ &= -\frac{x_\mathrm{b}d_\mathrm{b}^{-1}}{u_\mathrm{b}} + \frac{x_\mathrm{w}d_\mathrm{w}^{-1}}{u_\mathrm{w}} \\ &= -\frac{\sin\theta_\mathrm{b}}{u_\mathrm{b}} + \frac{\sin\theta_\mathrm{w}}{u_\mathrm{w}} \\ &= 0,\end{aligned}$$

if

$$\frac{\sin\theta_\mathrm{b}}{u_\mathrm{b}} = \frac{\sin\theta_\mathrm{w}}{u_\mathrm{w}}. \qquad \text{Snell's law} \tag{2.1}$$

That this condition determines a minimum time is verified as follows. (1) When x exceeds the Snell's law value, the positive term (the second term) contributing to dt/dx is larger in magnitude than the negative term (the first term). In this case, the derivative is positive and the time increases with further displacement from the Snell's law value. (2) When x is smaller than the Snell's law value, the first term dominates and the derivative is negative. In this case, time again increases with further (negative) displacement of x.

2.1.2 The Hamilton-Jacobi equation

The Hamilton-Jacobi equation arises when one attempts to construct new coordinates in phase space (the space of positions *and* momenta) such that some coordinates increase in proportion to time, while other coordinates remain fixed. The desired *conserved* coordinates are expressed as the partial derivative of a certain function determined by solving the Hamilton-Jacobi equation. Here, we consider only the time independent Hamilton-Jacobi equation (there is also a time dependent Hamilton-Jacobi equation). The time independent Hamilton-Jacobi equation is normally written as

$$\frac{1}{2m}\left(\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right)^2 + V(\mathbf{x}) = E, \tag{2.2}$$

where $\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}$ is the gradient of the function, $W(\mathbf{x})$, and

$$\left(\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right)^2 = \left(\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right)^T \left(\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right) = \left\|\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right\|^2.$$

Any solution, $W(\mathbf{x})$, is called *Hamilton's characteristic function.*

If we re-write the Hamilton-Jacobi equation in the form,

$$\left\|\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\right\| = p(\mathbf{x}) = [2m(E - V(\mathbf{x}))]^{1/2},$$

it looks like the eikonal equation for the phase of some wave. If we suppose that the wave in question has the *de Broglie* wavelength of a particle,

$$\lambda(\mathbf{x}) = \frac{h}{p(\mathbf{x})},$$

then the eikonal equation takes the form,

$$\begin{aligned}\left\|\frac{\partial \varphi(\mathbf{x})}{\partial \mathbf{x}}\right\| &= k(\mathbf{x}) = \frac{2\pi}{\lambda(\mathbf{x})} \\ &= \frac{p(\mathbf{x})}{\hbar}, \qquad \text{where } \hbar = \frac{h}{2\pi}\end{aligned}$$

Thus, we identify $W(\mathbf{x})/\hbar$ as the phase of a wave,

$$\psi(\mathbf{x}) = A(\mathbf{x}) \exp\left(\frac{iW(\mathbf{x})}{\hbar}\right), \tag{2.3}$$

where $A(\mathbf{x})$ is a slowly varying wave amplitude.

In the Hamilton-Jacobi equation, momentum appears as the gradient of Hamilton's characteristic function,

$$\mathbf{p} = \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}.$$

In the limit of slowly varying $A(\mathbf{x})$ (compared with $W(\mathbf{x})/\hbar$) we can neglect the gradient of $A(\mathbf{x})$ and write

$$\begin{aligned}\frac{\partial\psi(\mathbf{x})}{\partial\mathbf{x}} &= \frac{i}{\hbar}\frac{\partial W(\mathbf{x})}{\partial\mathbf{x}}A(\mathbf{x})\exp\left(\frac{iW(\mathbf{x})}{\hbar}\right)\\ &= \frac{i}{\hbar}\mathbf{p}\psi(\mathbf{x}),\end{aligned}$$

or

$$\mathbf{p}\psi(\mathbf{x}) = -i\hbar\frac{\partial\psi(\mathbf{x})}{\partial\mathbf{x}}. \tag{2.4}$$

Here, momentum is identified with the linear operator,

$$\hat{\mathbf{p}} = -\,i\hbar\frac{\partial}{\partial\mathbf{x}} \tag{2.5}$$

and the above equation,

$$\hat{\mathbf{p}}\psi(\mathbf{x}) = \mathbf{p}\psi(\mathbf{x}),$$

valid in the limit of constant amplitude, $A(\mathbf{x}) = A$, is an eigenvalue equation.

Taking the second derivative of the wavefunction (again, neglecting the gradient of $A(\mathbf{x})$) gives

$$\begin{aligned}\left(\frac{\partial}{\partial\mathbf{x}}\right)^2\psi(\mathbf{x}) &= -\frac{1}{\hbar^2}\left(\frac{\partial W(\mathbf{x})}{\partial\mathbf{x}}\right)^2\psi(\mathbf{x})\\ &= -\frac{2m}{\hbar^2}(E - V(\mathbf{x}))\psi(\mathbf{x}), \qquad \text{from the Hamilton-Jacobi equation}\end{aligned}$$

or

$$\begin{aligned}\frac{-\hbar^2}{2m}\left(\frac{\partial}{\partial\mathbf{x}}\right)^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) &= E\psi(\mathbf{x}) \qquad (2.6)\\ \frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) &= E\psi(\mathbf{x}).\end{aligned}$$

This is the *time independent Schrôdinger equation (TISE)* - the eigenvalue equation associated with the linear operator,

$$\hat{H} = \frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{x}).$$

The eigenvalue(s) of this Hamiltonian operator are the energies of the system. With the TISE, we can predict the Rydberg formula for the the energy levels of hydrogen. Extending the equation to more electrons (and nuclei), we can (as far as we can tell) predict the properties of all materials and (in principle) explain almost all natural phenomena.

The idea here is that this is the fundamental equation, and that neglecting the gradient of $A(\mathbf{x})$ is the approximation that leads to the Hamilton-Jacobi equation from the TISE - it is the Hamilton-Jacobi equation that is approximate, not the TISE.

There is also the time dependent Hamilton-Jacobi equation,

$$\frac{1}{2m}\left(\frac{\partial S(\mathbf{x},t)}{\partial \mathbf{x}}\right)^2 + V(\mathbf{x}) = -\frac{\partial S(\mathbf{x},t)}{\partial t}.$$

A solution, $S(\mathbf{x},t)$, of this equation is called Hamilton's principle function. Associating it with a time dependent wave,

$$\Psi(\mathbf{x},t) = A(\mathbf{x},t)\exp\left(\frac{iS(\mathbf{x},t)}{\hbar}\right),$$

and neglecting both space and time derivatives of $A(\mathbf{x},t)$) gives the *time dependent Schrôdinger equation (TDSE)*,

$$\frac{-\hbar^2}{2m}\nabla^2\Psi(\mathbf{x},t) + V(\mathbf{x})\Psi(\mathbf{x},t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t). \tag{2.7}$$

Here, $\Psi(\mathbf{x},t)$ is the wavefunction of a particle (or particles) that evolves in time. This TDSE gives the equation of motion for the particle(s) in quantum mechanics - i.e., its solution gives the future state of the particle(s) for all time.

Quantum mechanics (QM) is generally presented as a series of postulates.

2.2 The postulates of quantum mechanics

1. The *state* of a system of N particles (spinless particles - we add spin later) is represented by a unique *complex-valued wavefunction*, $\psi(\mathbf{x})$, such that

   $$\|\psi\|^2 = \langle\psi|\,\psi\rangle = \int_{\mathbb{R}^{3N}} |\psi(\mathbf{x})|^2\, d\mathbf{x} = .1. \qquad \text{normalization condition} \tag{2.8}$$

   where $\mathbb{R}^{3N}$ is the set of all positions of the N in three dimensional physical space. In general, the wavefunction belongs to a *Hilbert space* - a vector space with an inner product (and some more technical properties). Here,

   $$\langle\psi_1|\,\psi_2\rangle = \int_{\mathbb{R}^{3N}} \psi_1^*(\mathbf{x})\,\psi_2(\mathbf{x})\; d\mathbf{x} \tag{2.9}$$

   is the inner product of two wavefunctions (two states), ψ_1 and ψ_2.

   (a) The state of a system is *everything* we can know about the system.

   (b) The wavefunction (uniquely representing a state), $\psi(\mathbf{x})$, is a *smooth*, *single-valued*, *bounded* function of $\mathbf{x}$. Here, $\psi(\mathbf{x})$ is smooth if it, and its first *derivative(s)*, are *continuous* functions of $\mathbf{x}$.

   (c) Since $|\psi(\mathbf{x})|^2$ integrates to a finite value, it must decay to zero outside a finite volume of space.

2. Each *observable*, $\mathcal{A}$, of the system is *represented* by a unique *Hermitian linear operator*, $\hat{A}$, such that the *eigenvalues* of $\hat{A}$, $\{a_j\}_{j=1,2,\ldots}$, are the *possible outcomes* of a *measurement* of the observable - for all possible system states. An observable is any property of the system you can measure - e.g., the energy or momentum of the system, or atom(s) within the system.

   (a) A *linear operator* satisfies the condition,

   $$\hat{A}\left(c_1\psi_1 + c_2\psi_2\right) = c_1\hat{A}\psi_1 + c_2\hat{A}\psi_2, \tag{2.10}$$

   for any states, ψ_1 and ψ_2, and any real coefficients, c_1 and c_2.

   (b) A *Hermitian operator* satisfies the condition,

   $$\left\langle \psi_1 \right| \hat{A}\psi_2 \Big\rangle = \left\langle \psi_2 \right| \hat{A}\psi_1 \Big\rangle^*; \qquad \text{matrix representations are Hermitian} \tag{2.11}$$

   i.e.,

   $$\int_{\mathbb{R}^{3N}} \psi_1^*(\mathbf{x})\, \hat{A}\psi_2(\mathbf{x})\, d\mathbf{x} = \left(\int_{\mathbb{R}^{3N}} \psi_2^*(\mathbf{x})\, \hat{A}\psi_1(\mathbf{x})\, d\mathbf{x}\right)^*.$$

   Note that

   $$\begin{aligned} \left\langle \psi_2 \right| \hat{A}\psi_1 \Big\rangle^* &= \int_{\mathbb{R}^{3N}} \psi_2(\mathbf{x}) \left(\hat{A}\psi_1(\mathbf{x})\right)^* d\mathbf{x} \\ &= \int_{\mathbb{R}^{3N}} \left(\hat{A}\psi_1(\mathbf{x})\right)^* \psi_2(\mathbf{x})\, d\mathbf{x} \\ &= \left\langle \hat{A}\psi_1 \right| \psi_2 \Big\rangle \end{aligned}$$

   provides another equivalent expression of the Hermitian property,

   $$\left\langle \psi_1 \right| \hat{A}\psi_2 \Big\rangle = \left\langle \hat{A}\psi_1 \right| \psi_2 \Big\rangle \tag{2.12}$$

   (c) This postulate provides the source of *quantization.* For bound systems (e.g., a stable atom or molecule), the eigenvalues of Hermitian linear operators generally come in discrete sets - they can be labeled by a *quantum number*, $n = 1, 2, \ldots$.

3. The *expectation value* of an observable ($\mathcal{A}$), associated with a given state of the system ($\psi(\mathbf{x})$), is given by

   $$\left\langle \hat{A} \right\rangle = \left\langle \psi \right| \hat{A}\psi \Big\rangle = \int_{\mathbb{R}^{3N}} \psi^*(\mathbf{x})\, \hat{A}\psi(\mathbf{x})\, d\mathbf{x}. \tag{2.13}$$

   The expectation value is the *average* of many measurements of the observable - all starting with the system in state, $\psi(\mathbf{x})$.

(a) This is where probability enters into quantum mechanics. Quantum mechanics only gives probabilities of outcomes.. In fact, the postulate requires only that quantum mechanics provides the average value of many measurements. Probabilities of outcomes can be derived using this postulate - see below.

(b) Suppose the state of the system, $\psi(\mathbf{x}) = \psi_j(\mathbf{x})$, the normalized eigenfunction of $\hat{A}$ associated with eigenvalue, a_j; i.e.,

$$\hat{A}\psi_j(\mathbf{x}) = a_j\psi_j(\mathbf{x}) . \tag{2.14}$$

In this case,

$$\begin{aligned}\left\langle \hat{A} \right\rangle_j &= \int_{\mathbb{R}^{3N}} \psi_j^*(\mathbf{x})\,\hat{A}\psi_j(\mathbf{x})\; d\mathbf{x} \\ &= \int_{\mathbb{R}^{3N}} \psi_j^*(\mathbf{x})\,a_j\psi_j(\mathbf{x})\; d\mathbf{x} \\ &= a_j \int_{\mathbb{R}^{3N}} \left|\psi_j(\mathbf{x})\right|^2 d\mathbf{x} \\ &= a_j \qquad \psi_j \text{ is normalized}\end{aligned}$$

The expectation value of $\mathcal{A}$ for an eigenfunction of $\hat{A}$ is the associated eigenvalue. In fact there is a stronger result:

(c) Corollary of this postulate: The *probability* that a_j will be observed upon measurement of observable $\mathcal{A}$, when the system is in state ψ, equals

$$\begin{aligned}\rho_j &= \left|\left\langle \psi_j \right| \psi \rangle\right|^2 \\ &= \langle \psi | \left.\psi_j\right\rangle \left\langle \psi_j \right| \psi\rangle .\end{aligned} \tag{2.15}$$

This result follows from applying the postulate to the observable which simply asks the question, "Is the outcome of the measurement of $\mathcal{A}$ equal to a_j?" The values of the observable are 1 for yes and 0 for no. The probability, ρ_j, is the expectation value of this observable which can be written in the abstract form,

$$R_j = \left|\psi_j\right\rangle \left\langle \psi_j\right| .$$

It is the projector onto the subspace spanned by ψ_j.

i. We know that, if $\psi = \psi_j$, then the outcome of a measurement of $\mathcal{A}$ is certain to be a_j. The above formula gives this result. In this case,

$$\begin{aligned}\rho_j &= \left|\left\langle \psi_j \right| \psi_j \rangle\right|^2 \\ &= 1. \qquad \psi_j \text{ is normalized}\end{aligned}$$

The *eigenfunctions* of $\hat{A}$ thus represent states with *well-defined value* of the observable, $\mathcal{A}$, and that value is the associated *eigenvalue.*

ii. Consider the observable equal to one when $\mathbf{x}$ is in some subset of configuration space, S, and zero otherwise; i.e.,

$$\hat{A} = \theta_S(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in S \\ 0 & \mathbf{x} \notin S \end{cases}$$

In this case, the expectation of $\hat{A}$ equals the probability that $\mathbf{x} \in S$,

$$\begin{aligned} \rho_S &= \int_{\mathbb{R}^{3N}} \psi^*(\mathbf{x})\,\theta_S(\mathbf{x})\,\psi(\mathbf{x})\,d\mathbf{x} \\ &= \int_S |\psi(\mathbf{x})|^2\,d\mathbf{x} \end{aligned}$$

If the set S is is a small neighborhood (such that $\psi(\mathbf{x}) \cong$ constant on S) of a point, $\mathbf{x}_0$, then

$$\rho_S \cong |\psi(\mathbf{x}_0)|^2 \int_S d\mathbf{x}.$$

Here, we see that $|\psi(\mathbf{x})|^2$ is the *probability density* for observing configuration $\mathbf{x}$. Think of a single particle. In this case, $|\psi(\mathbf{x})|^2$ is the probability per unit volume for finding the particle at $\mathbf{x}$.

Before considering the next postulate, we consider a subpostulate of postulate 2 above.

2'. Operators representing observables:

(a) The operator that corresponds to position is the multiplication operator, $\mathbf{x}$. Here, it is best to consider first the case of a single particle in one dimension. In this case, $\mathbf{x} = x$ is a scalar and

$$\hat{x}\psi(x) = x\psi(x); \tag{2.16}$$

i.e., simply mulitply the function $\psi(x)$ by the function x. For one particle in three dimensions,

$$\begin{aligned} \hat{x}\psi(x,y,z) &= x\psi(x,y,z), \\ \hat{y}\psi(x,y,z) &= y\psi(x,y,z), \\ \hat{z}\psi(x,y,z) &= z\psi(x,y,z). \end{aligned}$$

i. Any observable which is a function of position is also represented by a multiplication operator - simply multiply by the function of position. The potential energy operator, $\hat{V}$, is a multiplication operator:

$$\hat{V}\psi(\mathbf{x}) = V(\mathbf{x})\,\psi(\mathbf{x}). \tag{2.17}$$

The operator $\hat{A} = \theta_S(\mathbf{x})$, considered above, is also a multiplication operator.

(b) The momentum operator is given by

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

i. An operator which is a function of momentum is given by the function with $-i\hbar\partial/\partial\mathbf{x}$ inserted as the argument. The kinetic energy operator is

$$\begin{aligned} \hat{T}\psi(\mathbf{x}) &= \frac{\hat{\mathbf{p}}^2}{2m}\psi(\mathbf{x}) \qquad (2.18) \\ &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial\mathbf{x}^2}\psi(\mathbf{x}). \end{aligned}$$

(c) The Hamiltonian operator:

$$\begin{aligned} \hat{H}\psi(\mathbf{x}) &= \left(\frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}\right)\psi(\mathbf{x}) \qquad (2.19) \\ &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial\mathbf{x}^2}\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) \end{aligned}$$

The *time independent Schrödinger equation* (TISE) is just the *eigenvalue equation* for the *energy* (Hamiltonian) operator,

$$\hat{H}\psi(\mathbf{x}) = E\psi(\mathbf{x}). \qquad (2.20)$$

The eigenvalues are the allowed energies of the system. Energy is quantized - in agreement with spectroscopic observations. The special significance of the energy operator in QM arises because of the next postulate.

4. The time evolution of the state of a system is given by the solutions to the *time dependent Schrödinger equation* (TDSE),

$$\hat{H}\Psi(\mathbf{x},t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t). \qquad (2.21)$$

Here, we use an upper case psi, $\Psi(\mathbf{x},t)$, to denote the time dependent state of the system. Given that the system is in the state, $\psi(\mathbf{x})$, at time = 0, the solution to the TDSE consistent with this intial condition provides the state of the system at all later times - it is a function of $\mathbf{x}$ and t.

(a) Solution to TDSE: Let

$$\Psi(\mathbf{x},t)\big|_{t=0} = \psi(\mathbf{x}).$$

For a very short time, δt,

$$\Psi(\mathbf{x},\delta t) \cong \Psi(\mathbf{x},0) + \left.\frac{\partial}{\partial t}\Psi(\mathbf{x},t)\right|_{t=0}\delta t \qquad \text{McLaurin series truncated at linear in } \delta t$$

Since

$$\left.\frac{\partial}{\partial t}\Psi\left(\mathbf{x},t\right)\right|_{t=0} = -\frac{i}{\hbar}\hat{H}\Psi\left(\mathbf{x},0\right) = -\frac{i}{\hbar}\hat{H}\psi\left(\mathbf{x}\right),$$

$$\begin{aligned}\Psi\left(\mathbf{x},\delta t\right) &\cong \psi\left(\mathbf{x}\right) - \frac{i}{\hbar}\hat{H}\psi\left(\mathbf{x}\right)\ \delta t. \\ &= \left(\hat{1} - \frac{i}{\hbar}\hat{H}\ \delta t\right)\psi\left(\mathbf{x}\right).\end{aligned}$$

Also,

$$\begin{aligned}\Psi\left(\mathbf{x},2\delta t\right) &\cong \Psi\left(\mathbf{x},\delta t\right) - \frac{i}{\hbar}\hat{H}\Psi\left(\mathbf{x},\delta t\right)\ \delta t \\ &= \left(\hat{1} - \frac{i}{\hbar}\hat{H}\ \delta t\right)^2\psi\left(\mathbf{x}\right).\end{aligned}$$

After N such steps, $t = N\ \delta t$ and

$$\begin{aligned}\Psi\left(\mathbf{x},t\right) &= \Psi\left(\mathbf{x},N\ \delta t\right) = \left(\hat{1} - \frac{1}{N}\frac{i}{\hbar}\hat{H}\ t\right)^N\psi\left(\mathbf{x}\right) \qquad (2.22) \\ &\rightarrow \exp\left(-\frac{i}{\hbar}\hat{H}\ t\right)\psi\left(\mathbf{x}\right), \qquad N\rightarrow\infty \text{ while } \delta t = \frac{t}{N}\rightarrow 0\end{aligned}$$

where

$$\exp\left(\hat{A}\right) = \hat{1} + \hat{A} + \frac{1}{2}\hat{A}^2 + \frac{1}{3!}\hat{A}^3 + \ldots$$

is the exponential function of operator $\hat{A}$.

(b) Suppose $\left.\Psi\left(\mathbf{x},t\right)\right|_{t=0} = \psi_n\left(\mathbf{x}\right)$ where $\psi_n\left(\mathbf{x}\right)$ is the n th energy eigenstate - i.e., the solution to the TISE,

$$\hat{H}\psi_n\left(\mathbf{x}\right) = E_n\psi_n\left(\mathbf{x}\right).$$

In this case,

$$\begin{aligned}\Psi\left(\mathbf{x},t\right) &= \exp\left(-\frac{i}{\hbar}\hat{H}\ t\right)\psi_n\left(\mathbf{x}\right) \\ &= \left(\hat{1} - \frac{i}{\hbar}\hat{H}\ t - \frac{1}{2\hbar^2}\hat{H}^2\ t^2 + \frac{i}{3!\ \hbar^3}\hat{H}^3\ t^3 + \ldots\right)\psi_n\left(\mathbf{x}\right) \\ &= \left(\hat{1} - \frac{i}{\hbar}E_n\ t - \frac{1}{2\hbar^2}E_n^2\ t^2 + \frac{i}{3!\ \hbar^3}E_n^3\ t^3 + \ldots\right)\psi_n\left(\mathbf{x}\right) \\ &= \exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n\left(\mathbf{x}\right)\end{aligned}$$

Here, we see that solutions to the TISE - i.e., states with well-defined energy - change in time only via an overall complex phase factor,

$$\exp\left(-\frac{i}{\hbar}E_n\ t\right).$$

i. Note that

$$\begin{aligned}\hat{H}\exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n(\mathbf{x}) &= \exp\left(-\frac{i}{\hbar}E_n\ t\right)\hat{H}\psi_n(\mathbf{x}) \\ &= E_n\exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n(\mathbf{x});\end{aligned}$$

i.e., the time evolved energy eigenstate is also an eigenstate - associated with the same energy eigenvalue.

ii. Also, the probability density in position space is *independent of time*, when the system state is an energy eigenstate.

$$\begin{aligned}|\Psi(\mathbf{x},t)|^2 &= \left|\exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n(\mathbf{x})\right|^2 \\ &= \left|\exp\left(-\frac{i}{\hbar}E_n\ t\right)\right|^2|\psi_n(\mathbf{x})|^2 \\ &= \exp\left(\frac{i}{\hbar}E_n\ t-\frac{i}{\hbar}E_n\ t\right)|\psi_n(\mathbf{x})|^2 \\ &= |\psi_n(\mathbf{x})|^2=|\Psi(\mathbf{x},0)|^2\end{aligned}$$

iii. Moreover, the probability of measuring *any* observable, $\mathcal{A}$, to have value, a_j, does not change in time, when the system state is an energy eigenstate.

$$\begin{aligned}\rho_j(t) &= \left|\langle\psi_j|\ \Psi(\mathbf{x},t)\rangle\right|^2 \\ &= \left|\langle\psi_j|\ \exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n\right\rangle\Bigg|^2 \\ &= \left|\exp\left(-\frac{i}{\hbar}E_n\ t\right)\langle\psi_j|\ \psi_n\rangle\right|^2 \\ &= \left|\langle\psi_j|\ \psi_n\rangle\right|^2 \qquad \text{see point ii above}\end{aligned}$$

(c) Energy eigenstates are also called *stationary states* - for the above reasons.

(d) The TDSE determines the time dependence of any state, $\Psi(x,t)$, given the state at any time (usually $t=0$). Simply expand the initial state, $\psi(x)$, in terms of the energy eigenstates,

$$\psi(x)=\sum_n c_n\psi_n(x).$$

The state is given at any time, t, by

$$\begin{aligned}\Psi\left(x,t\right) &= \exp\left(-\frac{i}{\hbar}\hat{H}\ t\right)\psi\left(x\right) \\ &= \exp\left(-\frac{i}{\hbar}\hat{H}\ t\right)\sum_n c_n\psi_n\left(x\right) \\ &= \sum_n c_n\exp\left(-\frac{i}{\hbar}\hat{H}\ t\right)\psi_n\left(x\right) && \text{the propagator is linear} \\ &= \sum_n c_n\exp\left(-\frac{i}{\hbar}E_n\ t\right)\psi_n\left(x\right). && \text{see point b above}\end{aligned} \tag{2.23}$$

The coefficients, c_n, are given by

$$c_n = \left\langle\psi_n\right|\left.\psi\right\rangle.$$

This follows because the energy eigenstates, ψ_n, are *orthogonal* - see the second fundamental theorem of quantum mechanics below.

(e) If the Hamiltonian is time independent, the TDSE is *separable* in space and time. On the left side of the equation, the operator $\hat{H}$ affects only on the spatial dependence of the wavefunction. On the right, $i\hbar\,\partial/\partial t$ affects only on the time dependence of the wavefunction. This allows the equation to be solved as the product of a time dependent phase factor, $\exp\left(-\frac{i}{\hbar}E_n\ t\right)$, and the spatial function, $\psi_n\left(x\right)$. The phase factor is an eigenfunction of $i\hbar\,\partial/\partial t$, while the spatial function is an eigenfunction of the time independent Hamiltonian - it is a solution to the TISE.

(f) When considering the interaction of the system with light, we introduce a time dependent term into the Hamiltonian which accounts for the coupling of the system to the electromagnetic field. The time dependence arises because the electric field is oscillatory in time for light which is resonant with a system transition. The interaction term is

$$\hat{\mu}_e\mathcal{E}\left(t\right),$$

where $\hat{\mu}_e$ is the dipole operator of the system, and $\mathcal{E}\left(t\right)$ is the time dependent electric field. Because the resulting Hamiltonian is time dependent, the TDSE is not separable in this case. However, this TDSE can be solved using *time dependent perturbation theory* (see Sec. XX) wherein the interaction is treated as a small term which induces transitions between the eigenstates of the time independent Hamiltonian (i.e., without the interaction term).

Example 2.1 For a particle in one dimension, the state of the system is represented by a function, $\psi\left(x\right)$, of the one variable, x. The normalization condition for the wavefunction is

$$\left\langle\psi\right|\left.\psi\right\rangle = \int_{-\infty}^{\infty}\left|\psi\left(x\right)\right|^2\,dx = 1.$$

Any wavefunction with finite $\langle\psi|\,\psi\rangle$ can be normalized (scaled by an appropriate factor) to satisfy the normalization condition. Which of the following functions can be normalized to represent a state of the particle? For those that do not respresent a physical state, which conditions on admissable wavefunctions are violated?

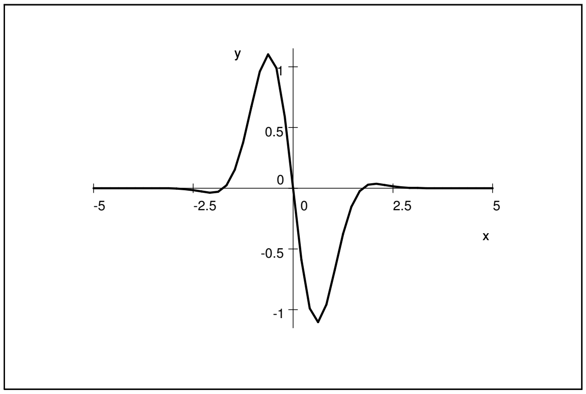
(a)
$$\psi\left(x\right)=\exp\left(-\left|x\right|\right)$$

(b)
$$\psi\left(x\right)=\cos\left(x\right)\exp\left(-x^2\right)$$

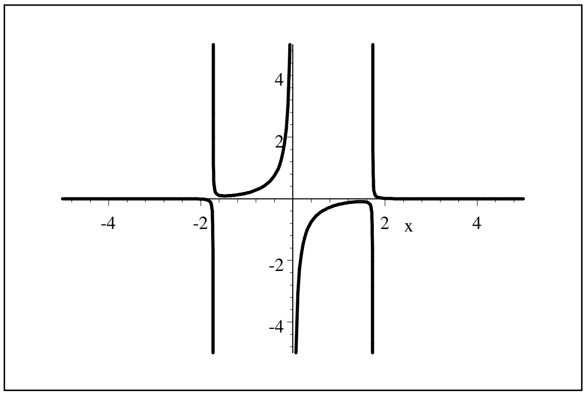
(c)
$$\psi\left(x\right)=\pm\exp\left(x^2\right)$$

(d)
$$\psi\left(x\right)=\left\{\begin{array}{ll}\frac{1}{2}\sin^2\left(x\right)\exp\left(-x^2\right), & x<0\\ \left(1-\cos\left(x\right)\right)\exp\left(-x^2\right), & x<0\end{array}\right.$$

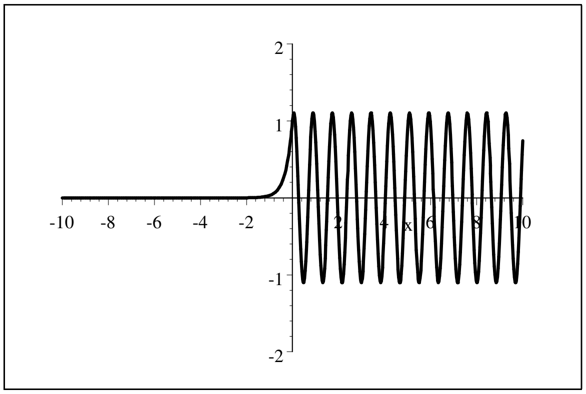
(e)



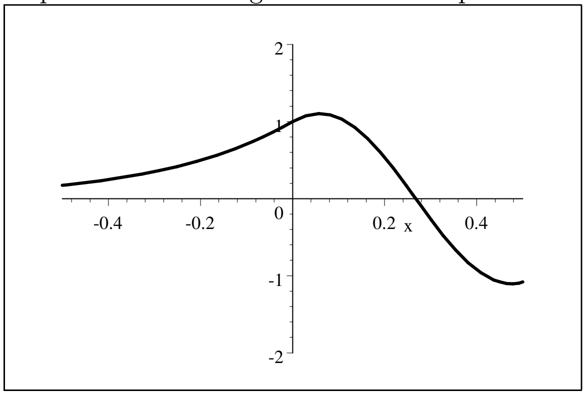
(f)



(g)

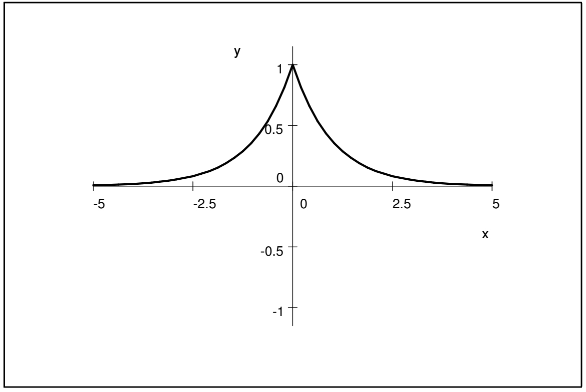


Below is a blow-up of the central region of the above plot.



Solution 2.1

(a) This function has a discontinuous first derivative at $x = 0$ - see figure.



To see this analytically, write

$$\psi(x) = \begin{cases} \exp(-x), & x > 0 \\ \exp(x), & x < 0 \end{cases}$$

and find

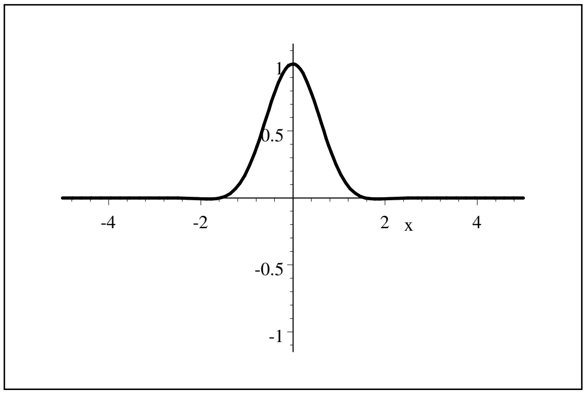
$$\psi'(x) = \begin{cases} -\exp(-x), & x > 0 \\ \exp(x), & x < 0 \end{cases} .$$

Thus,

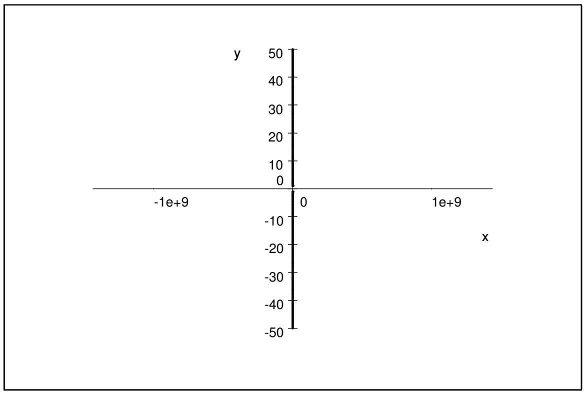
$$\begin{aligned} \lim_{x \to 0^-} \psi'(x) &= \left. \exp(x) \right|_{x=0} = 1 \\ &\neq \lim_{x \to 0^+} \psi'(x) = \left. -\exp(-x) \right|_{x=0} = -1. \end{aligned}$$

This function is not smooth. It thus cannot represent the state of a physical system - it is *not admissable*. However, this function arises as the solution to an idealized potential model system with infinite potential energy in a infinitesimal interval about $x = 0$. It can be viewed as the limit of a sequence of admissable wavefunctions, associated with a increasing sequence of potentials, in a decreasing interval about $x = 0$.

(b) Trigonomic and exponential functions are smooth, and products of smooth functions are smooth. Therefore, this function is smooth. It also decays rapidly to zero as $x \to \pm\infty$ - see figure. It is an *admissable* wavefunction - i.e., if normalized, it represents the state of a physical system.



(c) This function is multivalued - it is ± a single valued function. It is also unbounded as $x \to \pm\infty$. It is *not* an admissable wavefunction.



(d) In this case, we note that

$$\lim_{x\to 0^-} \psi(x) = (1-\cos(x))\exp\left(-x^2\right)\Big|_{x=0} = 0$$

and

$$\lim_{x\to 0^+} \psi(x) = \frac{1}{2}\sin^2(x)\exp\left(-x^2\right)\Bigg|_{x=0} = 0.$$

This function is continuous at $x = 0$, and everywhere else. Now consider the first derivative,

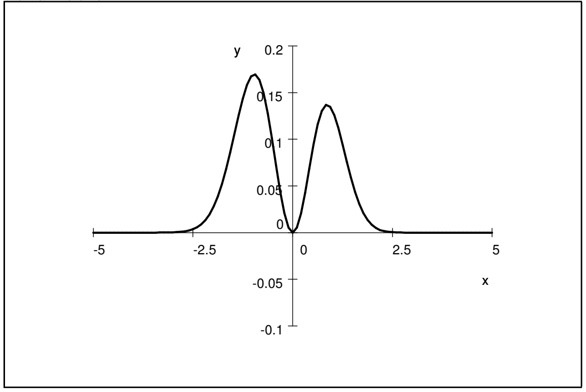
$$\psi'(x) = \begin{cases} \left(\sin(x)\cos(x) - x\sin^2(x)\right)\exp\left(-x^2\right), & x < 0 \\ (\sin(x) - 2x(1-\cos(x)))\exp\left(-x^2\right), & x < 0 \end{cases}.$$

$$\lim_{x\to 0^-} \psi(x) = \left(\sin(x)\cos(x) - x\sin^2(x)\right)\exp\left(-x^2\right)\Big|_{x=0} = 0$$

and

$$\lim_{x\to 0^+} \psi(x) = (\sin(x) - 2x(1-\cos(x)))\exp\left(-x^2\right)\Big|_{x=0} = 0.$$

The first derivative of the function is also continuous everywhere. In addition, the function decays rapidly to zero, as $x \to \pm\infty$. It is an *admissable* wavefunction.



(e) Admissable.

(f) Inadmissable. This function has poles where it blows up (down) to ±∞ on either side of the pole. The function is unbounded near the poles.

(g) Inadmissable. This function and its derivative are continuous - see blow-up for the interval about $x = 0$, where two functions are pieced together. However, it does not decay to zero as $x \to \infty$ (as far as can be ascertained from the figure). This function does arise in the treatment of a model system of a unbound particle - see Sec. XX below. However, it describes a *beam* of particles - rather than just a single particle.

Example 2.2 Which of the following operators, $\hat{A}$, represents an observable of a particle in one dimension?

(a)

$$\hat{A}\psi(x) = \psi^2(x)$$

(b)

$$\hat{A}\psi(x) = \sin(x)\,\psi(x)$$

(c)

$$\hat{A}\psi(x) = \psi'(x)$$

(d)

$$\hat{A}\psi(x) = \psi''(x) + 3\psi(x)$$

Solution 2.2

(a) This operator is nonlinear. $\hat{A}c\psi = c^2\hat{A}\psi \neq c\hat{A}\psi$. It does *not* represent an observable.

(b) This operator is linear. It is a multiplication operator, a function of the position operator, x. It is Hermitian. To see this, consider the condition for *Hermiticity*,

$$\langle\psi|\,\hat{A}\varphi\Big\rangle = \langle\varphi|\,\hat{A}\psi\Big\rangle^*.$$

Since $\hat{A}$ is a multiplication operator, the order the functions appear in the inner product does not matter. Since $\sin(x)$ is a real function of x, the condition for Hermicity,

$$\int_{-\infty}^{\infty}\psi^*(x)\sin(x)\,\varphi(x)\;dx = \left(\int_{-\infty}^{\infty}\varphi^*(x)\sin(x)\,\psi(x)\right)^*\,dx$$

is satisfied. This operator does represent an *observable* of the particle.

(c) This operator is linear. However, it is not Hermitian. Here, the left side of the above condition for Hermiticity takes the form,

$$\begin{aligned}
\left\langle\psi\right|\left.\hat{A}\varphi\right\rangle &= \int_{-\infty}^{\infty}\psi^{*}(x)\,\frac{d}{dx}\varphi(x)\;dx \\
&= \left[\psi^{*}(x)\,\varphi(x)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\left(\frac{d}{dx}\psi^{*}(x)\right)\varphi(x)\;dx \qquad \text{integration by parts} \\
&= -\int_{-\infty}^{\infty}\varphi(x)\,\frac{d}{dx}\psi^{*}(x)\;dx \qquad \psi(x)\ \text{ and }\ \varphi(x)\rightarrow 0,\ \text{as}\ x\rightarrow\pm\infty \\
&= -\left\langle\varphi\right|\left.\hat{A}\psi\right\rangle^{*}.
\end{aligned}$$

Thus, this operator does *not* represent an observable. The Hermitian conjugate of d/dx is $-d/dx$. However, a Hermitian operator - representing an observable - is obtained if the derivative operator is multiplied by i (or $-i$). For example,

$$\begin{aligned}
\left\langle\psi\right|\left.\left(-i\frac{d}{dx}\right)\varphi\right\rangle &= -i\int_{-\infty}^{\infty}\psi^{*}(x)\,\frac{d}{dx}\varphi(x)\;dx \\
&= i\int_{-\infty}^{\infty}\varphi(x)\,\frac{d}{dx}\psi^{*}(x)\;dx \\
&= \left\langle\varphi\right|\left.\left(-i\frac{d}{dx}\right)\psi\right\rangle^{*}.
\end{aligned}$$

Multiplying this operator by $\hbar$ gives the momentum operator.

(d) This is the sum of two operators, d^2/dx^2 and the multiplication operator, 3, which multiplies $\psi(x)$ by 3. Both of these are linear and Hermitian, producing a linear and Hermitian sum. The second derivative operator is proportional to the kinetic energy operator. We show that the second derivative operator is Hermitian as follows:

$$\begin{aligned}
\left\langle\psi\right|\left.\hat{A}\varphi\right\rangle &= \int_{-\infty}^{\infty}\psi^{*}(x)\,\frac{d^2}{dx^2}\varphi(x)\;dx \\
&= \left[\psi^{*}(x)\,\frac{d}{dx}\varphi(x)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\left(\frac{d}{dx}\psi^{*}(x)\right)\frac{d}{dx}\varphi(x)\;dx \qquad \text{integration by parts} \\
&= -\int_{-\infty}^{\infty}\left(\frac{d}{dx}\psi^{*}(x)\right)\frac{d}{dx}\varphi(x)\;dx \qquad \psi(x)\ \text{ and }\ \frac{d}{dx}\varphi(x)\rightarrow 0,\ \text{as}\ x\rightarrow\pm\infty \\
&= \left[\left(\frac{d}{dx}\psi^{*}(x)\right)\varphi(x)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\left(\frac{d^2}{dx^2}\psi^{*}(x)\right)\varphi(x)\;dx \qquad \text{integration by parts} \\
&= -\int_{-\infty}^{\infty}\varphi(x)\,\frac{d^2}{dx^2}\psi^{*}(x)\;dx \qquad \frac{d}{dx}\psi(x)\ \text{ and }\ \varphi(x)\rightarrow 0,\ \text{as}\ x\rightarrow\pm\infty \\
&= \left\langle\varphi\right|\left.\hat{A}\psi\right\rangle^{*}.
\end{aligned}$$

Example 2.3 Match the list of operators with wavefunctions that are eigenstates of the associated observable. What is the associated eigenvalue?

(a) Operators:

i.
$$\hat{A}\psi(x) = -i\psi'(x)$$

ii.
$$\hat{B}\psi(x) = -\psi''(x)$$

iii.
$$\hat{C}\psi(x) = -\psi''(x) + 4x^2\psi(x)$$

iv.
$$\hat{D}\psi(x) = -\psi''(x) - \frac{2}{x}\psi(x)$$

(b) Wavefunctions:

i.
$$\psi(x) = x\exp(-x)$$

ii.
$$\psi(x) = \exp(ipx)$$

p is a real parameter. $\hbar = h/2\pi$ is reduced Planck's constant.

iii.
$$\psi(x) = \exp\left(-x^2\right)$$

iv.
$$\psi(x) = \sin(px)$$

Solution 2.3

(a) i.

$$\begin{aligned}\hat{A}\exp(ipx) &= -i(\exp(ipx))' \\ &= -i(ip)\exp(ipx) \\ &= p\exp(ipx).\end{aligned}$$

$\exp(ipx)$ is an eigenfunction of operator $\hat{A}$. The associated eigenvalue is p. Any real p produces a distinct eigenstate of this operator (it is the momentum operator in units such that $\hbar = 1$ - e.g., atomic units).

ii.

$$\begin{aligned}\hat{B}\sin(px) &= -(\sin(px))'' \\ &= p^2\sin(px).\end{aligned}$$

$\sin(px)$ is an eigenfunction of $\hat{B}$. The associated eigenvalue is p^2.

iii.

$$\begin{aligned}\hat{C}\exp\left(-x^2\right) &= -\left(\exp\left(-x^2\right)\right)'' + 4x^2\exp\left(-x^2\right)\\ &= \left(2x\exp\left(-x^2\right)\right)' + 4x^2\exp\left(-x^2\right)\\ &= \left(2-4x^2\right)\exp\left(-x^2\right) + 4x^2\exp\left(-x^2\right)\\ &= 2\exp\left(-x^2\right).\end{aligned}$$

$\exp\left(-x^2\right)$ is an eigenfunction of $\hat{C}$. The associated eigenvalue is 2.

iv.

$$\begin{aligned}\hat{D}x\exp(-x) &= -\left(x\exp(-x)\right)'' - \frac{2}{x}x\exp(-x)\\ &= -\left((1-x)\exp(-x)\right)' - 2\exp(-x)\\ &= -(-2+x)\exp(-x) - 2\exp(-x)\\ &= x\exp(-x).\end{aligned}$$

$x\exp(-x)$ is an eigenfunction of $\hat{D}$. The associated eigenvalue is 1.

2.3 Fundamental theorems of quantum mechanics

2.3.1 The eigenvalues and eigenfunctions of Hermitian operators

Theorem 1. The eigenvalues of a Hermitian operator are real.

Proof: Let $\{a_j\}$ be the eigenvalues of Hermitian operator, $\hat{A}$, with $\{\psi_j\}$ the corresponding eigenfunctions - i.e.,

$$\hat{A}\psi_j = a_j\psi_j.$$

The Hermitian property of $\hat{A}$ gives

$$\left\langle\psi\right|\left.\hat{A}\varphi\right\rangle = \left\langle\varphi\right|\left.\hat{A}\psi\right\rangle^*,$$

for any ψ and φ. If we let $\psi = \varphi = \psi_j$, then

$$\begin{aligned}\left\langle\psi_j\right|\left.\hat{A}\psi_j\right\rangle &= \left\langle\psi_j\right|\left.\hat{A}\psi_j\right\rangle^*\\ \left\langle\psi_j\right|\left.a_j\psi_j\right\rangle &= \left\langle\psi_j\right|\left.a_j\psi_j\right\rangle^*\\ a_j\left\langle\psi_j\right|\left.\psi_j\right\rangle &= a_j^*\left\langle\psi_j\right|\left.\psi_j\right\rangle\\ a_j &= a_j^* \qquad \text{i.e., } a_j \text{ is real}\end{aligned}$$

Theorem 2. Eigenfunctions associated with distinct eigenvalues of a Hermitian operator are orthogonal.

Proof: The Hermitian property of $\hat{A}$ gives

$$\langle\psi|\left.\hat{A}\varphi\right\rangle = \langle\varphi|\left.\hat{A}\psi\right\rangle^*,$$

for any ψ and φ. If we let $\psi = \psi_j$ and $\varphi = \psi_{j'}$, such that $a_j \neq a_{j'}$, then

$$\begin{aligned}
\left\langle\psi_j\right|\left.\hat{A}\psi_{j'}\right\rangle &= \left\langle\psi_{j'}\right|\left.\hat{A}\psi_j\right\rangle^* \\
\left\langle\psi_j\right|\left.a_{j'}\psi_{j'}\right\rangle &= \left\langle\psi_{j'}\right|\left.a_j\psi_j\right\rangle^* \\
a_{j'}\left\langle\psi_j\right|\left.\psi_{j'}\right\rangle &= a_j^*\left\langle\psi_{j'}\right|\left.\psi_j\right\rangle^* \\
a_{j'}\left\langle\psi_j\right|\left.\psi_{j'}\right\rangle &= a_j\left\langle\psi_j\right|\left.\psi_{j'}\right\rangle \qquad a_j \text{ is real}
\end{aligned}$$

or

$$(a_{j'} - a_j)\left\langle\psi_j\right|\left.\psi_{j'}\right\rangle = 0.$$

Since, $a_j \neq a_{j'}$, we must have

$$\left\langle\psi_j\right|\left.\psi_{j'}\right\rangle = 0;$$

i.e., ψ_j and $\psi_{j'}$ are orthogonal. If $a_j = a_{j'}$, then the eigenvalue is degenerate (there is more than one eigenfunction). In this case, the eigenfunctions associated with the same eigenvalue form a vector space. An orthonormal basis for the space can be found. Taking these basis states as the distinct eigenstates associated with degenerate eigenvalues extends this orthogonality theorem to all pairs of distinct eigenstates.

2.3.2 Commutators

Operators can be added or multiplied. Operator multiplication is like (exactly like) matrix multiplication (matrices are just operators that act on vectors). In particular, operators do not necessarily commute - specifically, in general,

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}.$$

When $\hat{A}\hat{B}$ does equal $\hat{B}\hat{A}$, we say that $\hat{A}$ and $\hat{B}$ *commute*. Now we consider another fundamental theorem of quantum mechanics.

Theorem 3. $\hat{A}$ and $\hat{B}$ commute iff they have a common set of eigenfunctions, $\{\psi_j\}$.

Proof (if): Suppose $\hat{A}$ and $\hat{B}$ have a common set of eigenfunctions, $\{\psi_j\}$. For physical observables, the set of eigenfunctions is always a complete set - i.e., any function, $\psi(x)$, can be written as a linear combination of the eigenfunctions:

$$\psi(x) = \sum_j c_j \psi_j(x).$$

Apply $\hat{A}\hat{B}$ to ψ.

$$\begin{aligned}
\hat{A}\hat{B}\psi &= \hat{A}\hat{B}\sum_j c_j\psi_j(x) \\
&= \sum_j c_j\hat{A}\hat{B}\psi_j(x) && \hat{A}\hat{B}\text{ is a linear operator} \\
&= \sum_j c_j\hat{A}b_j\psi_j(x) && \psi_j\text{ is an eigenfunction of }\hat{B} \\
&= \sum_j c_j b_j\hat{A}\psi_j(x) && \hat{A}\text{ is linear} \\
&= \sum_j c_j b_j a_j\psi_j(x) && \psi_j\text{ is an eigenfunction of }\hat{A} \\
&= \sum_j c_j a_j b_j\psi_j(x) && \text{complex numbers commute} \\
&= \sum_j c_j a_j\hat{B}\psi_j(x) && \psi_j\text{ is an eigenfunction of }\hat{B} \\
&= \sum_j c_j\hat{B}a_j\psi_j(x) && \hat{B}\text{ is linear} \\
&= \sum_j c_j\hat{B}\hat{A}\psi_j(x) && \psi_j\text{ is an eigenfunction of }\hat{A} \\
&= \hat{B}\hat{A}\sum_j c_j\psi_j(x) && \hat{B}\hat{A}\text{ is a linear operator} \\
&= \hat{B}\hat{A}\psi(x)
\end{aligned}$$

Since $\hat{A}\hat{B}\psi = \hat{B}\hat{A}\psi$ for any function, ψ, we can say that the operators, $\hat{A}\hat{B}$ and $\hat{B}\hat{A}$ are equal - i.e., $\hat{A}$ and $\hat{B}$ commute.

Proof (only if): Suppose that $\hat{A}$ and $\hat{B}$ commute. Let $\{\psi_j\}$ be the set of eigenfunctions of $\hat{A}$; i.e.,

$$\hat{A}\psi_j = a_j\psi_j.$$

Now apply $\hat{A}\hat{B}$ to any ψ_j.

$$\begin{aligned}
\hat{A}\hat{B}\psi_j &= \hat{B}\hat{A}\psi_j && \hat{A}\text{ and }\hat{B}\text{ commute} \\
&= \hat{B}a_j\psi_j && \psi_j\text{ is an eigenfunction of }\hat{A} \\
&= a_j\hat{B}\psi_j && \hat{B}\text{ is linear}
\end{aligned}$$

This means that $\hat{B}\psi_j$ is an eigenfunction of $\hat{A}$, associated with eigenvalue, a_j. However, ψ_j is the eigenfunction of $\hat{A}$, associated with eigenvalue,

a_j. If there is only one eigenfunction associated with eigenvalue, a_j, the eigenvalue is said to be *non-degenerate*. In this case, $\hat{B}\psi_j$ must be a multiple of ψ_j; i.e.,

$$\hat{B}\psi_j = b_j\psi_j.$$

The multiple is denoted by b_j and is seen here to be the eigenvalue of $\hat{B}$ associated with eigenfunction, ψ_j. Specifically, we see that ψ_j is an eigenfunction of $\hat{B}$ and b_j is the associated eigenvalue. Thus, the eigenfunctions of $\hat{A}$ are also eigenfunctions of $\hat{B}$.

Degenerate case: If the eigenvalue, a_j, is *degenerate* - i.e., there is more than one associated eigenfunction - we apply $\hat{A}\hat{B}$ to all the eigenfunctions associated with eigenvalue, a_j. The result is that $\hat{B}$ maps the subspace spanned by these eigenfunctions into itself. Specifically, $\hat{B}$ maps any function from this subspace onto an eigenfunction of $\hat{A}$ associated with eigenvalue, a_j. The eigenfunctions of $\hat{B}$ that span this subspace are thus also eigenfunctions of $\hat{A}$. $\hat{A}$ and $\hat{B}$ have a common set of eigenfunctions.

We see that commuting operators have a *common set of eigenfunctions*. This means that the associated observables have well-defined values for exactly the same states. For example, if a state has well-defined value of observable, A, then it also has well-defined value of any other observable, B, for which the operators representing A and B commute.

The commutator of two operators is simply defined as

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

If two operators commute, their commutator is zero. To evaluate commutators, consider the operators acting on some wavefunction, $\psi(x)$. $\hat{A}\hat{B}\psi$ is given by first applying $\hat{B}$ to ψ, then applying $\hat{A}$ to the result.

2.3.3 The commutator of position and momentum

For position and momentum operators, we consider the action of the position and momentum commutator on an arbitrary wavefunction, $\psi(x)$.

$$\begin{aligned}
[\hat{p}, x]\,\psi(x) &= \left[-i\hbar\frac{d}{dx}, x\right]\psi(x) && \text{definition of } \hat{p} \\
&= -i\hbar\frac{d}{dx}x\psi(x) - x\left(-i\hbar\frac{d}{dx}\right)\psi(x) && \text{definition of commutator} \\
&= -i\hbar\left(\frac{d}{dx}x\psi(x) - x\frac{d\psi(x)}{dx}\right) \\
&= -i\hbar\left(\psi(x) + x\frac{d\psi(x)}{dx} - x\frac{d\psi(x)}{dx}\right) && \text{product rule} \\
&= -i\hbar\psi(x)\,.
\end{aligned}$$

Since the above is true for all wavefunctions, $\psi(x)$, we must have

$$[\hat{p}, x] = -i\hbar \tag{2.24}$$

This means that $[\hat{p}, x]$ is just $-i\hbar\times$ the identity operator. Note that since

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

for all $\hat{A}$ and $\hat{B}$, we also have

$$[x, \hat{p}] = i\hbar. \tag{2.25}$$

Note that it is customary to leave the hat off of multiplication operators such as x.

As position and momentum do not commute, they do not have a common set of eigenfunctions. Position and momentum cannot have well-defined values for the same system state.

2.4 Uncertainty principles

2.4.1 The Heisenberg uncertainty principle

The relationship between uncertainty in position and uncertainty in momentum is quantified by the *Heisenberg uncertainty principle.* To understand this principle, we must first define *uncertainty.* In quantum mechanics, the outcomes of measurements are, in general, uncertain. The mean of such outcomes is given by the expectation value of the associated operator. In statistics, the *standard deviation*, σ, gives the spread in a distribution. It is the square root of the variance,

$$\begin{aligned}
\sigma^2 &= \langle\psi\left|\left(\hat{A}-\left\langle\hat{A}\right\rangle\right)^2\psi\right\rangle \qquad (2.26)\\
&= \langle\psi\left|\left(\hat{A}^2-2\left\langle\hat{A}\right\rangle\hat{A}+\left\langle\hat{A}\right\rangle^2\right)\psi\right\rangle\\
&= \langle\psi\left|\hat{A}^2\psi\right\rangle-2\left\langle\hat{A}\right\rangle\langle\psi\left|\hat{A}\psi\right\rangle+\left\langle\hat{A}\right\rangle^2\langle\psi\left|\psi\right\rangle \qquad \left\langle\hat{A}\right\rangle \text{ is a constant}\\
&= \langle\psi\left|\hat{A}^2\psi\right\rangle-2\left\langle\hat{A}\right\rangle^2+\left\langle\hat{A}\right\rangle^2\\
&= \langle\psi\left|\hat{A}^2\psi\right\rangle-\langle\psi\left|\hat{A}\psi\right\rangle^2,
\end{aligned}$$

where

$$\left\langle\hat{A}\right\rangle=\langle\psi\left|\hat{A}\psi\right\rangle.$$

If ψ were an eigenfunction of $\hat{A}$, then $\langle\psi\left|\hat{A}^2\psi\right\rangle=\langle\psi\left|\hat{A}\psi\right\rangle^2$ and $\sigma=0$; i.e., there is no uncertainty in the value of $\mathcal{A}$.

For position, we have

$$\sigma_x^2=\langle\psi\left|(\Delta x)^2\psi\right\rangle$$

where

$$\Delta x=x-\langle x\rangle.$$

Similarly, for momentum,

$$\sigma_p^2=\langle\psi\left|(\Delta\hat{p})^2\psi\right\rangle$$

where

$$\Delta\hat{p}=\hat{p}-\langle\hat{p}\rangle.$$

Consider the state,

$$\varphi=(\Delta\hat{p}+i\alpha\Delta x)\,\psi,$$

where α is any real number. The norm of this state is greater than or equal to zero - true for any state.

$$\begin{aligned}
\|\varphi\|^2 &= \langle\varphi\left|\varphi\right\rangle\\
&= \langle(\Delta\hat{p}+i\alpha\Delta x)\,\psi\left|(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle\\
&= \langle\Delta\hat{p}\psi\left|(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle-i\alpha\langle\Delta x\psi\left|(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle \qquad \text{definition of inner product}\\
&= \langle\psi\left|\Delta\hat{p}\,(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle-i\alpha\langle\psi\left|\Delta x\,(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle \qquad \Delta\hat{p}\text{ and }\Delta x\text{ are Hermitian}\\
&= \langle\psi\left|(\Delta\hat{p}-i\alpha\Delta x)\,(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle=f(\alpha) \qquad i\alpha\text{ is brought into right side of inner product}
\end{aligned}$$

Since $f(\alpha)$ is the norm of a function, it must be positive. Thus,

$$\begin{aligned}
f(\alpha) &= \langle\psi\left|(\Delta\hat{p}-i\alpha\Delta x)\,(\Delta\hat{p}+i\alpha\Delta x)\,\psi\right\rangle\\
&= \langle\psi\left|(\Delta\hat{p})^2\psi\right\rangle+\alpha^2\langle\psi\left|(\Delta x)^2\psi\right\rangle-i\alpha\langle\psi\left|[\Delta x,\Delta\hat{p}]\,\psi\right\rangle\geq 0
\end{aligned}$$

Since

$$\begin{aligned} \langle\psi\left|[\Delta x,\Delta\hat{p}]\,\psi\right\rangle &= \langle\psi\left|[x,\hat{p}]\,\psi\right\rangle & \text{constant terms drop out both on left} \\ && \text{and right side of commutators - check this} \\ &= \langle\psi\left|i\hbar\psi\right\rangle = i\hbar, \end{aligned}$$

the above inequality takes the form

$$\begin{aligned} f(\alpha) &= \sigma_p^2+\alpha^2\sigma_x^2+\hbar\alpha \\ &\geq \min_\alpha f(\alpha)\geq 0 \end{aligned}$$

The minimum of $f(\alpha)$ is determined by taking its derivative and setting it to zero. This gives

$$\frac{d}{d\alpha}f(\alpha)=2\alpha\sigma_x^2+\hbar=0,$$

or

$$\alpha=-\frac{\hbar}{2\sigma_x^2}.$$

Substituting this value back into $f(\alpha)$ gives

$$\begin{aligned} \min_\alpha f(\alpha) &= \sigma_p^2+\frac{\hbar^2}{4\sigma_x^4}\sigma_x^2-\frac{\hbar^2}{2\sigma_x^2} \\ &= \sigma_p^2+\frac{\hbar^2}{4\sigma_x^2}-\frac{\hbar^2}{2\sigma_x^2} \\ &= \sigma_p^2-\frac{\hbar^2}{4\sigma_x^2}\geq 0. \end{aligned}$$

Therefore,

$$\sigma_p^2\geq\frac{\hbar^2}{4\sigma_x^2},$$

$$\sigma_x^2\sigma_p^2\geq\frac{\hbar^2}{4}$$

and

$$\sigma_x\sigma_p\geq\frac{\hbar}{2}, \tag{2.27}$$

the *Heisenberg uncertainty principle*. The uncertainty in the position and momentum of a particle are reciprocally related. If one of position and momentum is precisely known, the other is less precisely known - and vice versa. $\hbar/2$ sets the scale on which the uncertainty principle plays a role. It is not relevant to the macroscopic objects normally treated using classical mechanics. It is very important for electrons on the atomic scale.

2.4.2 Angular momentum

Angular momentum is a vector observable for particles in three dimensional space. The three components of angular momentum are given by

$$\begin{aligned} \hat{L}_x &= y\hat{p}_z - z\hat{p}_y \qquad (2.28) \\ \hat{L}_y &= z\hat{p}_x - x\hat{p}_z \\ \hat{L}_z &= x\hat{p}_y - y\hat{p}_x . \end{aligned}$$

The angular momentum vector is the cross product of the position and momentum vectors,

$$\hat{\mathbf{L}} = \mathbf{x} \times \hat{\mathbf{p}}.$$

The order of the position and momentum factors in the terms on the right sides of the above equations does not matter because each position component is paired with a different momentum component - these operators commute. However, the different components of $\hat{\mathbf{L}}$ do not commute. For example,

$$\begin{aligned} \left[\hat{L}_x, \hat{L}_y\right] &= \left[y\hat{p}_z - z\hat{p}_y, z\hat{p}_x - x\hat{p}_z\right] \\ &= \left[y\hat{p}_z, z\hat{p}_x\right] - \left[y\hat{p}_z, x\hat{p}_z\right] - \left[z\hat{p}_y, z\hat{p}_x\right] + \left[z\hat{p}_y, x\hat{p}_z\right] \end{aligned}$$

commutators are linear
with respect both left and right sides

Consider the first of the four terms on the second line above.

$$\begin{aligned} \left[y\hat{p}_z, z\hat{p}_x\right] &= y\hat{p}_x \left[\hat{p}_z, z\right] \\ &= -i\hbar y\hat{p}_x \qquad \left[\hat{p}_z, z\right] = -i\hbar \end{aligned}$$

This follows because y and $\hat{p}_x$ commute with all other operators present. As such, they can be factored out of the commutator (on either side) [check this]. The second and third of the four terms above are zero because all operators in both of these commutators commute. The fourth term takes the form,

$$\begin{aligned} \left[z\hat{p}_y, x\hat{p}_z\right] &= x\hat{p}_y \left[z, \hat{p}_z\right] \\ &= i\hbar x\hat{p}_y . \end{aligned}$$

Altogether,

$$\begin{aligned} \left[\hat{L}_x, \hat{L}_y\right] &= -i\hbar y\hat{p}_x + i\hbar x\hat{p}_y \qquad (2.29) \\ &= i\hbar \left(x\hat{p}_y - y\hat{p}_x\right) \\ &= i\hbar \hat{L}_z . \end{aligned}$$

Similarly,

$$\left[\hat{L}_y, \hat{L}_z\right] = i\hbar \hat{L}_x$$

and

$$\left[\hat{L}_z, \hat{L}_x\right] = i\hbar \hat{L}_y .$$

The latter two equations are obtained by substituting x for y, y for z and z for x. The components of angular momentum advance in the same fashion.

Since no two of the components of $\hat{\mathbf{L}}$ commute, there is no common set of eigenfunctions. However, each component of $\hat{\mathbf{L}}$ commutes with the (square) magnitude of $\hat{\mathbf{L}}$. For example,

$$\begin{aligned}
\left[\hat{L}_z, \hat{\mathbf{L}}^2\right] &= \left[\hat{L}_z, \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right] \\
&= \left[\hat{L}_z, \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right] \\
&= \left[\hat{L}_z, \hat{L}_x^2 + \hat{L}_y^2\right]. \qquad \begin{array}{c}\hat{L}_z \text{ commutes with } \hat{L}_z \\ \text{or any power of } \hat{L}_z\end{array}
\end{aligned}$$

Since

$$\begin{aligned}
\left[\hat{L}_z, \hat{L}_x^2\right] &= \hat{L}_z\hat{L}_x^2 - \hat{L}_x^2\hat{L}_z \\
&= \hat{L}_z\hat{L}_x^2 - \hat{L}_x\hat{L}_z\hat{L}_x + \hat{L}_x\hat{L}_z\hat{L}_x - \hat{L}_x^2\hat{L}_z \\
&= \left(\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\right)\hat{L}_x + \hat{L}_x\left(\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\right) \\
&= \left[\hat{L}_z, \hat{L}_x\right]\hat{L}_x + \hat{L}_x\left[\hat{L}_z, \hat{L}_x\right] \\
&= \hat{L}_y\hat{L}_x + \hat{L}_x\hat{L}_y
\end{aligned}$$

and

$$\begin{aligned}
\left[\hat{L}_z, \hat{L}_y^2\right] &= \left[\hat{L}_z, \hat{L}_y\right]\hat{L}_y + \hat{L}_y\left[\hat{L}_z, \hat{L}_y\right] \\
&= -\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x,
\end{aligned}$$

clearly

$$\left[\hat{L}_z, \hat{\mathbf{L}}^2\right] = 0. \tag{2.30}$$

We also have

$$\left[\hat{L}_x, \hat{\mathbf{L}}^2\right] = 0$$

and

$$\left[\hat{L}_y, \hat{\mathbf{L}}^2\right] = 0.$$

Therefore, it is possible to find eigenfunctions common to both $\hat{\mathbf{L}}^2$ and one of the components of $\hat{\mathbf{L}}$. It is customary to choose $\hat{L}_z$ as the favored component of $\hat{\mathbf{L}}$. The eigenfunctions of $\hat{L}_z$ and $\hat{\mathbf{L}}^2$, expressed as functions of the two angles θ and ϕ in spherical coordinates are called the spherical harmonics - see Sec. XX.

Example 2.4 Evaluate the following commutators using commutators evaluated above.

(a)

$$\left[x^2, \hat{p}_x\right]$$

(b)

$$\left[x, \hat{p}_x^2\right]$$

(c)

$$\left[x^2, \hat{p}_x^2\right]$$

(d)

$$\left[\hat{L}_x, \hat{L}_y^2\right]$$

(e)

$$\left[\hat{L}_x, \hat{L}_z^2\right]$$

Solution 2.4

(a)

$$\begin{aligned}
\left[x^2, \hat{p}_x\right] &= x^2\hat{p}_x - \hat{p}_x x^2 \\
&= x^2\hat{p}_x - x\hat{p}_x x + x\hat{p}_x x - \hat{p}_x x^2 \\
&= x\left[x, \hat{p}_x\right] + \left[x, \hat{p}_x\right] x \\
&= 2i\hbar x \qquad \left[x, \hat{p}_x\right] = i\hbar
\end{aligned}$$

(b)

$$\begin{aligned}
\left[x, \hat{p}_x^2\right] &= x\hat{p}_x^2 - \hat{p}_x^2 x \\
&= x\hat{p}_x^2 - \hat{p}_x x\hat{p}_x + \hat{p}_x x\hat{p}_x - \hat{p}_x^2 x \\
&= \left[x, \hat{p}_x\right]\hat{p}_x + \hat{p}_x\left[x, \hat{p}_x\right] \\
&= 2i\hbar\hat{p}_x
\end{aligned}$$

(c)

$$\begin{aligned}
\left[x^2, \hat{p}_x^2\right] &= x^2\hat{p}_x^2 - \hat{p}_x^2 x^2 \\
&= x^2\hat{p}_x^2 - x\hat{p}_x^2 x + x\hat{p}_x^2 x - \hat{p}_x^2 x^2 \\
&= x\left[x, \hat{p}_x^2\right] + \left[x, \hat{p}_x^2\right] x \\
&= 2i\hbar\left(x\hat{p}_x + \hat{p}_x x\right) \qquad \text{see part b} \\
&= 2i\hbar\left(2x\hat{p}_x - i\hbar\right) \qquad \left[x, \hat{p}_x\right] = i\hbar
\end{aligned}$$

(d)

$$\begin{aligned}
\left[\hat{L}_x, \hat{L}_y^2\right] &= \hat{L}_x\hat{L}_y^2 - \hat{L}_y^2\hat{L}_x \\
&= \left[\hat{L}_x, \hat{L}_y\right]\hat{L}_y + \hat{L}_y\left[\hat{L}_x, \hat{L}_y\right] \qquad \left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \\
&= i\hbar\left(\hat{L}_z\hat{L}_y + \hat{L}_z\hat{L}_y\right)
\end{aligned}$$

(e)

$$\begin{aligned}\left[\hat{L}_x, \hat{L}_z^2\right] &= \hat{L}_x\hat{L}_z^2 - \hat{L}_z^2\hat{L}_x \\ &= \left[\hat{L}_x, \hat{L}_z\right]\hat{L}_z + \hat{L}_z\left[\hat{L}_x, \hat{L}_z\right] \\ &= -i\hbar\left(\hat{L}_y\hat{L}_z + \hat{L}_z\hat{L}_y\right)\end{aligned}$$

Example 2.5 Which of the following pairs of operators have a common set of eigenstates?

(a)

$$x \text{ and } \sin(x)$$

(b)

$$\sin(x) \text{ and } -i\frac{d}{dx}$$

(c)

$$\hat{p}_x \text{ and } \hat{p}_x^2$$

(d)

$$x \text{ and } \hat{p}_y$$

Solution 2.5

(a) Only commuting operators have a common set of eigenstates. Thus, we must show that the commutator is zero. To do this, apply the two operators, in the order given, to an arbitrary state, $\psi(x)$. Then apply the operators in reverse order and take the difference. This gives $\left[\hat{A}, \hat{B}\right]\psi(x)$. The operators commute only if the result is zero for any $\psi(x)$. Here,

$$\hat{A}\hat{B}\psi(x) = x\sin(x)\psi(x),$$

$$\begin{aligned}\hat{B}\hat{A}\psi(x) &= \sin(x)x\psi(x) \\ &= x\sin(x)\psi(x) \\ &= \hat{A}\hat{B}\psi(x)\end{aligned}$$

and the difference is

$$\left[\hat{A}, \hat{B}\right]\psi(x) = 0,$$

for any $\psi(x)$. The operators commute. They have a common set of eigenstates. These are both multiplication operators - i.e., functions of x. Any two multiplication operators commute because ordinary multiplication is commutative. In fact any two functions of the same operator commute. This can be proven by first proving that $\hat{A}$ and $f\left(\hat{A}\right)$ commute - expand $f\left(\hat{A}\right)$ as a Taylor series and show that $\hat{A}$ commutes with each term.

(b) Here,

$$\begin{aligned}\hat{A}\hat{B}\psi(x) &= \sin(x)\left(-i\frac{d}{dx}\psi(x)\right)\\ &= -i\sin(x)\frac{d}{dx}\psi(x),\end{aligned}$$

$$\begin{aligned}\hat{B}\hat{A}\psi(x) &= -i\frac{d}{dx}(\sin(x)\psi(x))\\ &= -i\left(\cos(x)\psi(x)+\sin(x)\frac{d}{dx}\psi(x)\right)\end{aligned}$$

and

$$\left[\hat{A},\hat{B}\right]\psi(x) = -i\cos(x)\psi(x) \neq 0.$$

Therefore, $\sin(x)$ and $-id/dx$ do not commute, and do not have a common set of eigenfunctions.

(c) Here,

$$\begin{aligned}\hat{A}\hat{B}\psi(x) &= \hat{p}_x\hat{p}_x^2\psi(x)\\ &= \hat{p}_x^3\psi(x),\end{aligned}$$

$$\begin{aligned}\hat{B}\hat{A}\psi(x) &= \hat{p}_x^2\hat{p}_x\psi(x)\\ &= \hat{p}_x^3\psi(x)\\ &= \hat{A}\hat{B}\psi(x)\end{aligned}$$

and the difference is

$$\left[\hat{A},\hat{B}\right]\psi(x) = 0,$$

for any $\psi(x)$. The operators commute. They have a common set of eigenstates. This is another example of an operator (here, $\hat{p}_x$) and a function of the operator (here, $\hat{p}_x^2$). Such pairs always commute.

(d) Since there is a $\hat{p}_y$ operator, there must also be a y coordinate. States of a particle with an x and y coordinate are represented by functions of both x and y. In this case,

$$\begin{aligned}\hat{A}\hat{B}\psi(x,y) &= x\hat{p}_y\psi(x,y)\\ &= x\left(-i\hbar\frac{\partial}{\partial y}\right)\psi(x,y)\\ &= -i\hbar x\frac{\partial}{\partial y}\psi(x,y)\end{aligned}$$

$$\begin{aligned}
\hat{B}\hat{A}\psi(x,y) &= \hat{p}_y x\psi(x,y) \\
&= -i\hbar\frac{\partial}{\partial y}x\psi(x,y) \\
&= -i\hbar x\frac{\partial}{\partial y}\psi(x,y) \\
&= \hat{A}\hat{B}\psi(x)
\end{aligned}$$

and the difference is

$$\left[\hat{A},\hat{B}\right]\psi(x) = 0.$$

These operators commute and have a common set of eigenstates. Position, x, commutes with the momenta associated with other coordinates.

Example 2.6 Uncertainty principle for angular momentum.

(a) Using the known commutator of $\hat{L}_x$ and $\hat{L}_y$, derive an uncertainty principle for the x and y components of angular momentum. To do this simply replace x and $\hat{p}_x$, in the derivation of the Heisenberg uncertainty principle, by $\hat{L}_x$ and $\hat{L}_y$. The only difference is the form of the commutator.

(b) There are subspaces of states with specific values of $\left|\left\langle\hat{L}_z\right\rangle\right|$. The associated values are 0, $\hbar$, $2\hbar$, ... Suppose the uncertainty in the x component of angular momentum is $\hbar$. For each distinct $\left|\left\langle\hat{L}_z\right\rangle\right|$ subspace of states, what does the uncertainty principle say about the uncertainty of the y component of angular momentum? Is it possible to have zero uncertainty in L_y?

Solution 2.6

(a) Following the derivation of the Heisenberg uncertainty principle, with x and $\hat{p}_x$ replaced by $\hat{L}_x$ and $\hat{L}_y$, the first difference arises because the commutator is different. Specifically, we have

$$\begin{aligned}
f(\alpha) &= \langle\psi\left|\left(\Delta\hat{L}_y - i\alpha\Delta\hat{L}_x\right)\left(\Delta\hat{L}_y + i\alpha\Delta\hat{L}_x\right)\psi\right\rangle \\
&= \langle\psi\left|\left(\Delta\hat{L}_y\right)^2\psi\right\rangle + \alpha^2\langle\psi\left|\left(\Delta\hat{L}_x\right)^2\psi\right\rangle - i\alpha\langle\psi\left|\left[\Delta\hat{L}_x,\Delta\hat{L}_y\right]\psi\right\rangle \geq 0,
\end{aligned}$$

where

$$\begin{aligned}
\langle\psi\left|\left[\Delta\hat{L}_x,\Delta\hat{L}_y\right]\psi\right\rangle &= \langle\psi\left|\left[\hat{L}_x,\hat{L}_y\right]\psi\right\rangle \\
&= \langle\psi\left|i\hbar\hat{L}_z\psi\right\rangle \\
&= i\hbar\langle\psi\left|\hat{L}_z\psi\right\rangle.
\end{aligned}$$

In the next step, we have

$$\begin{aligned} f(\alpha) &= \sigma_{L_y}^2 + \alpha^2 \sigma_{L_x}^2 + \hbar \left\langle \hat{L}_z \right\rangle \alpha \\ &\geq \min_{\alpha} f(\alpha) \geq 0. \end{aligned}$$

The minimum of $f(\alpha)$ occurs for

$$\alpha = -\frac{\hbar \left\langle \hat{L}_z \right\rangle}{2\sigma_{L_x}^2},$$

giving

$$\sigma_{L_y}^2 - \frac{\hbar^2 \left\langle \hat{L}_z \right\rangle^2}{4\sigma_{L_x}^2} \geq 0,$$

or

$$\sigma_{L_x}^2 \sigma_{L_y}^2 \geq \frac{\hbar^2 \left\langle \hat{L}_z \right\rangle^2}{4}$$

and

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{\hbar}{2} \left| \left\langle \hat{L}_z \right\rangle \right|,$$

(b) If $\sigma_{L_x} = \hbar$, then

$$\sigma_{L_y} \geq \frac{\hbar}{2\sigma_{L_x}} \left| \left\langle \hat{L}_z \right\rangle \right| = \frac{1}{2} \left| \left\langle \hat{L}_z \right\rangle \right| = 0,\ \frac{\hbar}{2},\ \hbar,\ \frac{3\hbar}{2}, \ldots.$$

For example, for the $\left| \left\langle \hat{L}_z \right\rangle \right| = 0$ subspace,

$$\sigma_{L_y} \geq 0.$$

It is possible to have zero uncertainty in L_y (or L_x) only for the $\left| \left\langle \hat{L}_z \right\rangle \right| = 0$ subspace. For the other $\left| \left\langle \hat{L}_z \right\rangle \right| = \ell\hbar$ subspaces, $\ell = 1, 2, \ldots$, we have

$$\sigma_{L_y} \geq \frac{\ell\hbar}{2} > 0.$$

Chapter 3

Model systems

Understanding quantum mechanics requires study of simple model systems. We begin with the simplest such system, the particle in a one dimensional box.

3.1 Particle in a one dimensional box

Consider a mass, m, particle constrained to move in one dimension with rigid boundaries. The position of the particle is specified by one variable, x. The state of the system is represented by a function, $\psi(x)$ - the *wavefunction*. The allowed wavefunctions must be zero on the boundaries of the box, at $x = 0$ and $x = L$. Otherwise, the wave function must be continuous. Normally we would also impose the condition that the derivative of the wavefunction is continuous at the boundary. However, in this case, the derivative of the wavefunction is not continuous at the boundary. This arises because the potential of the particle in a box has infinite discontinuities. The potential is zero within the box, but infinite outside. Specifically,

$$V(x) = \begin{cases} \infty, & x \leq 0 \\ 0, & 0 < x < L \\ \infty, & L \leq x \end{cases}, \tag{3.1}$$

as shown in Fig. 2.4. Such a potential can be viewed as the limit, as $V_{\mathrm{d}} \to \infty$, of the more realistic particle in a box potential,

$$V_{\text{finite-well-depth}}(x) = \begin{cases} V_{\mathrm{d}}, & x \leq 0 \\ 0, & 0 < x < L \\ V_{\mathrm{d}}, & L \leq x \end{cases}. \tag{3.2}$$

In the case of a particle in finite-well-depth potential, the wave function decays exponentially with distance beyond the edge of the box. As $V_{\mathrm{d}} \to \infty$ the decaying portions of the wavefunction become vanishingly small and the *boundary conditions*,

$$\psi(x) = 0 \qquad x = 0 \text{ or } x = L$$

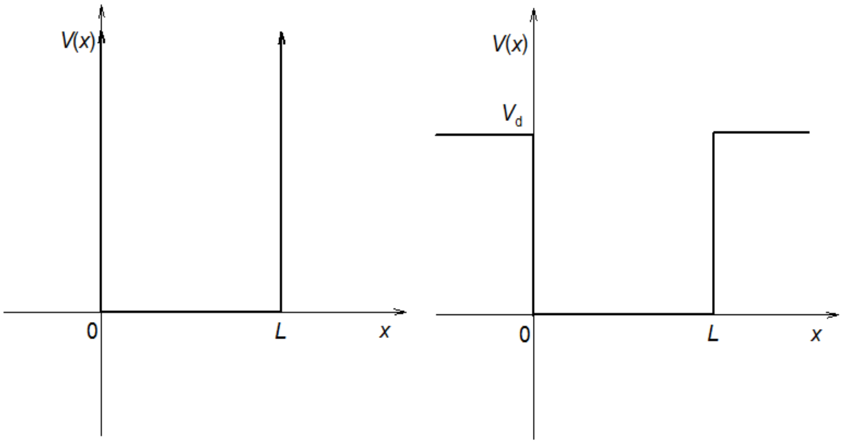


Figure 3.1: The particle in a one dimensional box potential (left), and particle in a finite one dimensional well (right). L is the width of the box and well. V_{d} is the depth of the well. The particle in a box potential is the $V_{\mathrm{d}} \to \infty$ limit of the particle in a well potential.

follow.

With the above boundary conditions, we solve the TISE for the energy eigenstates. For the particle in a one dimensional box, the TISE takes the form,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x) \tag{3.3}$$

or

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

We need the *eigenfunctions* of d^2/dx^2.

The *exponential function* is an eigenfunction of the d^2/dx^2 operator;

$$\frac{d^2}{dx^2}\exp(ax) = a^2\exp(ax).$$

For positive energy ($E > 0$), setting

$$a^2 = -\frac{2mE}{\hbar^2}$$

gives

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

There are thus two solutions to Eq. 3.3:

$$\exp\left(i\frac{\sqrt{2mE}}{\hbar}x\right) \qquad \text{and} \qquad \exp\left(-i\frac{\sqrt{2mE}}{\hbar}x\right).$$

Any combination of these two functions also solves Eq. 3.3 . The most general solution takes the form,

$$\psi(x) = A\exp\left(i\frac{\sqrt{2mE}}{\hbar}x\right) + B\exp\left(-i\frac{\sqrt{2mE}}{\hbar}x\right).$$

Now, we impose the *boundary conditions.*

1. At $x = 0$, $\psi(x) = 0$. Therefore,

$$\begin{aligned} 0 &= A\,\exp\left(i\frac{\sqrt{2mE}}{\hbar}x\right)\bigg|_{x=0} + B\,\exp\left(-i\frac{\sqrt{2mE}}{\hbar}x\right)\bigg|_{x=0} \\ &= A + B \end{aligned}$$

   or

$$B = -A$$

   and

$$\begin{aligned} \psi(x) &= A\left[\exp\left(i\frac{\sqrt{2mE}}{\hbar}x\right) - \exp\left(-i\frac{\sqrt{2mE}}{\hbar}x\right)\right] \\ &= 2iA\sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) \\ &= C\sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) \qquad C = 2iA \text{ is just an another arbitrary constant} \end{aligned}$$

2. At $x = L$, $\psi(x) = 0$. Therefore,

$$0 = C\sin\left(\frac{\sqrt{2mE}}{\hbar}L\right).$$

   This equation is satisfied if

$$C = 0$$

   or

$$\sin\left(\frac{\sqrt{2mE}}{\hbar}L\right) = 0.$$

   $C = 0$ gives $\psi(x) = 0$ which corresponds to no particle. If there is a particle, then $C \neq 0$ and

$$\sin\left(\frac{\sqrt{2mE}}{\hbar}L\right) = 0.$$

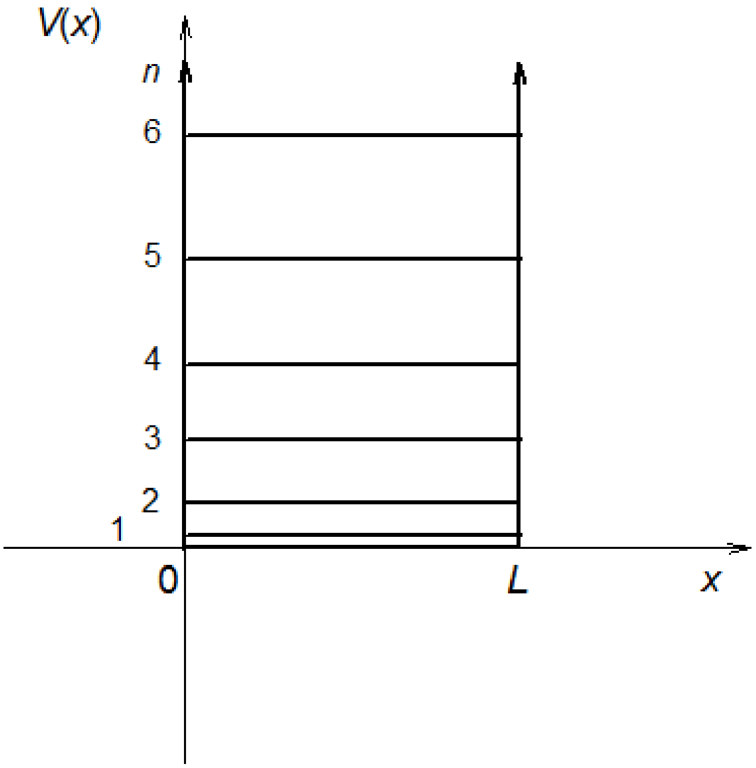


Figure 3.2: The first six energy levels of a particle in a one dimensional box. $E_n = n^2 E_1 = E_1$,$4E_1$,$9E_1$,$16E_1$,...

This equation is satisfied for energies such that

$$\frac{\sqrt{2mE}}{\hbar}L = n\pi \qquad n = 1, 2, 3, \ldots \qquad \text{the zeros of the sine function} \tag{3.4}$$

Note that $n = 0$ does not give a state. In this case, we get $E = 0$ and $\psi(x) = 0$ for all x - i.e., no particle. Also, negative n do not give real energies. The *energy eigenvalues* - the solutions to Eq. 3.4 - are given by

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2} \qquad n = 1, 2, 3, \ldots \tag{3.5}$$

These eigenvalues, also known as *energy levels* are depicted in Fig. 3.2. They are the possible outcomes of an energy measurement for a mass, m, particle in a one dimensional box of width, L.

The *energy eigenstates* are associated with a *discrete* set of energies. *Energy* *level*, E_n, is associated with energy eigenstate,

$$\psi_n(x) = C \sin\left(\frac{n\pi x}{L}\right) \qquad \text{inside box, zero outside.} \tag{3.6}$$

Discrete energy levels is a hallmark of quantum mechanics. The energy levels of any bound particle are discrete. Here, the particle is bound within a one dimensional box. The spacing between energy levels increases with increasing n, because of the n^2 energy level dependence. The spacing is reciprocal to the mass of the particle, and the square of the width of the box. A heavier particle, or a larger box, produces smaller energy level spacings.

3.1.1 Normalization of energy eigenstates

The overall mulitplicative constant, C, is determined by imposing *normalization*. Specifically, we impose the condition that the probability of finding the particle somewhere inside the box is one; i.e.,

$$\int_0^L |\psi_n(x)|^2 \, dx = 1. \qquad \text{the particle is definitely inside the box } (\psi_n(x) = 0 \text{ for } x < 0 \text{ or } L < x)$$

For $n = 1$, this equation takes the form

$$\int_0^L |C|^2 \sin^2\left(\frac{\pi x}{L}\right) dx = 1$$

$$|C|^2 \int_0^L \sin^2\left(\frac{\pi x}{L}\right) dx = 1, \qquad \text{constants can be moved outside integrals}$$

or

$$|C| = \frac{1}{\sqrt{\int_0^L \sin^2\left(\frac{\pi x}{L}\right) dx}}.$$

The integral here is given by

$$\int_0^L \sin^2\left(\frac{\pi x}{L}\right) dx = \frac{L}{\pi} \int_0^\pi \sin^2(u) \, du \qquad \text{change of variables, } u = \frac{\pi x}{L}.$$

To evaluate $\int_0^\pi \sin^2(u)\, du$, we use the trigonometric identity,

$$\cos(2u) = \cos^2(u) - \sin^2(u) = 1 - 2\sin^2(u),$$

which gives

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

and

$$\begin{aligned}
\int_0^\pi \sin^2(u)\,du &= \frac{1}{2}\int_0^\pi (1-\cos(2u))\,du \\
&= \frac{1}{2}\left(\int_0^\pi du - \int_0^\pi \cos(2u)\ du\right) \\
&= \frac{1}{2}\left(\pi - \left[\frac{1}{2}\sin(2u)\right]_0^\pi\right) \\
&= \frac{1}{2}\left(\pi - \frac{1}{2}\left[\sin(2\pi) - \sin(0)\right]\right) = \frac{\pi}{2},
\end{aligned}$$

since

$$\sin(2\pi) = \sin(0) = 0.$$

Therefore, the normalized wavefunction for $n = 1$ is given by

$$\begin{aligned}
\psi_1(x) = C\sin\left(\frac{\pi x}{L}\right) &= \frac{1}{\sqrt{\frac{L}{\pi}\frac{\pi}{2}}}\sin\left(\frac{\pi x}{L}\right) \\
&= \sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right).
\end{aligned}$$

Here, we set $C = |C|$. In principle, the above eigenstate can be multiplied by any phase factor - i.e., a complex number with modulus one - and still be a normalized energy eigenstate. However, we choose real wavefunctions whenever possible.

In the case of $\psi_n(x)$, we get

$$C = \frac{1}{\sqrt{\int_0^L \sin^2\left(\frac{n\pi x}{L}\right)dx}} \tag{3.7}$$

and

$$\begin{aligned}
\int_0^L \sin^2\left(\frac{n\pi x}{L}\right)\,dx &= \frac{L}{n\pi}\int_0^{n\pi}\sin^2(u)\ du \\
&= \frac{L}{n\pi}\ n\int_0^\pi \sin^2(u)\ du \qquad \text{integrand is periodic with period, } \pi \\
&= \frac{L}{\pi}\frac{\pi}{2} = \frac{L}{2}. \qquad \text{same as } n = 1 \text{ case}
\end{aligned}$$

So,

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right). \tag{3.8}$$

The four lowest energy eigenstates, $\psi_n(x)$ ($n = 1$ to 4), are depicted in Fig. 3.3.The energy eigenstates of the particle in a one dimensional box are waves. The $n = 1$ state has the lowest energy. It is called the *ground state*. It is the

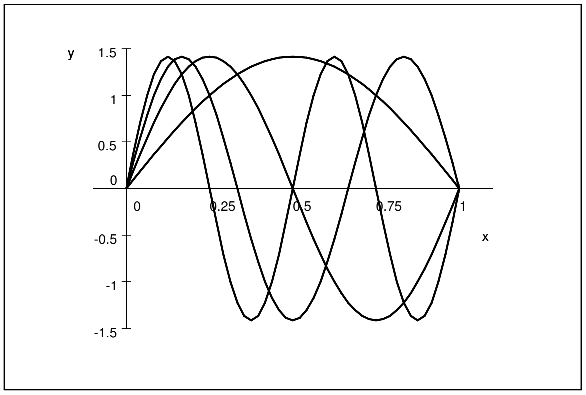


Figure 3.3: The first four energy eigenstates of a partile in a one dimensional box.

simplest standing wave subject to the boundary conditions. Each subsequent state has an additional *node* (x value such that $\psi(x) = 0$). Also, note that the odd n states are *even* about the center of the box, while the even n states are *odd*. Even states, about the center of the box, are symmetric with respect to reflection through a mirror at the center of the box. Odd states are reversed in sign when reflected through such a mirror.

3.1.2 Properties of a particle in a box

Momentum:

The expectation value of momentum for a particle in a box, in state $\psi_n(x)$:

$$\begin{aligned}
\langle\hat{p}\rangle_n &= \langle\psi_n|\,\hat{p}\psi_n\rangle = \int_0^L \psi_n^*(x)\left[-i\hbar\frac{d}{dx}\psi_n(x)\right]dx \\
&= \frac{2}{L}\int_0^L \sin\left(\frac{n\pi x}{L}\right)\left[-i\hbar\frac{n\pi}{L}\cos\left(\frac{n\pi x}{L}\right)\right]dx \\
&= -\frac{2in\pi\hbar}{L^2}\int_0^L \sin\left(\frac{n\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right)dx \\
&= -\frac{2i\hbar}{L}\int_0^{n\pi} \sin(u)\cos(u)\,du \\
&= -\frac{i\hbar}{L}\int_0^{n\pi} \sin(2u)\,du \\
&= -\frac{i\hbar n}{L}\int_0^{\pi} \sin(2u)\,du \\
&= \frac{i\hbar n}{2L}\left[\cos(2u)\right]_0^{\pi} = \frac{i\hbar n}{2L}\left[1-1\right] = 0
\end{aligned}$$

The energy eigenstates of the particle in a box are standing waves - superpositions of *equal* amounts of *left-going* and *right-going* wavefunctions. The average of many momentum measurements is zero - plus and minus contributions cancel. Note that the above conclusion could have reached much faster using symmetry. Specifically, write the above integral in the form (just consider case of $n = 1$ here),

$$\begin{aligned}
\int_0^L \sin\left(\frac{\pi x}{L}\right)\cos\left(\frac{\pi x}{L}\right)\,dx &= \int_{-L/2}^{L/2} \sin\left(\frac{\pi\left(x+L/2\right)}{L}\right)\cos\left(\frac{\pi\left(x+L/2\right)}{L}\right)\,dx \\
&= \int_{-L/2}^{L/2} \sin\left(\frac{\pi x}{L}+\frac{\pi}{2}\right)\cos\left(\frac{\pi x}{L}+\frac{\pi}{2}\right)\,dx \\
&= -\int_{-L/2}^{L/2} \cos\left(\frac{\pi x}{L}\right)\sin\left(\frac{\pi x}{L}\right)\,dx \\
&= -\frac{L}{\pi}\int_{-\pi/2}^{\pi/2} \cos\left(u\right)\sin\left(u\right)\,du = 0
\end{aligned}$$

The integral in the last line is zero because the integrand is the product of an even function, $\cos(u)$, and an odd function, $\sin(u)$, about $u = 0$, the center of the box. The product of an even and odd function is odd. The integral over the two halves of the box have equal magnitude, but opposite sign - they cancel perfectly.

Position

The expectation value of position for a particle in a box, in state $\psi_n(x)$:

$$\begin{aligned}
\langle \hat{x} \rangle_n &= \langle \psi_n | \, \hat{x}\psi_n \rangle = \int_0^L \psi_n^*\left(x\right) x \psi_n\left(x\right)\,dx \qquad (3.9)\\
&= \frac{2}{L}\int_0^L x\sin^2\left(\frac{n\pi x}{L}\right)\,dx \\
&= \frac{2L}{n^2\pi^2}\int_0^{n\pi} u\sin^2\left(u\right)\,du \\
&= \frac{2L}{n^2\pi^2}\int_0^{n\pi} u\left(\frac{1-\cos\left(2u\right)}{2}\right)\,du \\
&= \frac{L}{n^2\pi^2}\left\{\int_0^{n\pi} u\,du - \int_0^{n\pi}\cos\left(2u\right)\,du\right\} \\
&= \frac{L}{n^2\pi^2}\left\{\frac{1}{2}\left[u^2\right]_0^{n\pi} - \frac{1}{2}\left[\sin\left(2u\right)\right]_0^{n\pi}\right\} \\
&= \frac{L}{2n^2\pi^2}\left\{\left[n^2\pi^2-0\right]-\left[0-0\right]\right\} \\
&= \frac{L}{2}
\end{aligned}$$

The distribution of position is spread out within the box, symmetrically about the middle, $x = L/2$. The average of many position measurements would be $L/2$, for any energy eigenstate, ψ_n.

3.1.3 Spectroscopic transitions

Light impinging on a particle in a box (e.g., an electron in a polyene) can induce *transitions* between the states of the particle in a box. Specifically, an electron can absorb the energy of one photon of light, if its energy ($h\nu$) equals the difference in energy between the initial and final states of the electron - n_{i} and n_{f}, respectively. In this case, the *probability* $P_{n_{\mathrm{f}},n_{\mathrm{i}}}$ that the electron absorbs the light (per unit time squared, given a constant intensity of light) is proportional to the light intensity, and to the square of the expectation of the electron dipole moment, $\hat{\mu}_e = -ex$, ($-e$ is the charge on an electron) - see Sec. 6.3.1.

$$P_{n_{\mathrm{f}},n_{\mathrm{i}}} \propto \left|\langle\hat{\mu}_e\rangle_{n_{\mathrm{f}},n_{\mathrm{i}}}\right|^2 = e^2\left|\int_0^L x\psi_{n_{\mathrm{f}}}(x)\,\psi_{n_{\mathrm{i}}}(x)\;dx\right|^2 .$$

For example, in the case of $n_{\mathrm{i}} = 1$ and $n_{\mathrm{f}} = 2$, the *transition integral* (the probability is proportional to the square of this integral) takes the form,

$$\begin{aligned} I_{2,1} &= \frac{2}{L}\int_0^L x\sin\left(\frac{2\pi x}{L}\right)\sin\left(\frac{\pi x}{L}\right)dx \\ &= \frac{2L}{\pi^2}\int_0^\pi u\sin(2u)\sin(u)\;du \qquad u = \frac{\pi x}{L} \end{aligned}$$

This integral is evaluated using the trigonometric identities,

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$$

and

$$\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$$

which together give

$$\sin(a)\sin(b) = \frac{1}{2}\left[\cos(a-b) - \cos(a+b)\right]$$

or

$$\sin(2u)\sin(u) = \frac{1}{2}\left[\cos(u) - \cos(3u)\right].$$

The above integral now takes the form,

$$\begin{aligned} I_{2,1} &= \frac{L}{\pi^2}\int_0^\pi u\left[\cos(u) - \cos(3u)\right]\,du \\ &= \frac{L}{\pi^2}\left[\int_0^\pi u\cos(u)\;du - \int_0^\pi u\cos(3u)\;du\right] . \end{aligned}$$

Both integrals in the square bracket are of the form,

$$\begin{aligned}
&\int_0^{\pi} u\cos(ju)\, du \\
&= \int_{u=0}^{u=\pi} u\,\frac{1}{j}\,d\sin(ju) \\
&= \frac{1}{j}\left\{[u\,\sin(ju)]_0^{\pi} - \int_0^{\pi}\sin(ju)\,du\right\} \\
&= \frac{1}{j}\left\{\left[\underbrace{\pi\sin(j\pi)}_{=0,\text{ since } j=\text{ an integer}} - 0\right] - \left[-\frac{1}{j}\cos(ju)\right]_0^{\pi}\right\} \\
&= \frac{1}{j^2}\left[\cos(j\pi) - \cos(0)\right] \\
&= \frac{1}{j^2}\left[(-1)^j - 1\right] = \begin{cases} 0, & j \text{ even} \\ -\frac{2}{j^2}, & j \text{ odd} \end{cases} .
\end{aligned}$$

Applying this formula to the integral that determines the $n = 1$ to 2 transition integral gives

$$I_{2,1} = \frac{L}{\pi^2}\left[-\frac{2}{1^2} + \frac{2}{3^2}\right] = -\frac{2L}{\pi^2}\left[1 - \frac{1}{9}\right] = -\frac{16L}{9\pi^2}. \tag{3.10}$$

The associated transition probability is proportional to

$$P_{2,1} \propto |I_{2,1}|^2 = \frac{256L^2}{81\pi^4}.$$

If we considered the transition from $n = 1$ to 3, the integral would have been $(3 - 1 = 2$ and $3 + 1 = 4)$

$$P_{3,1} = \frac{L}{\pi^2}\int_0^{\pi} u\left[\cos(2u) - \cos(4u)\right]\,du = 0,$$

since j is even (2 and 4) in both of the (sub) integrals appearing here.

In general, the transition probability is zero for transitions from even n to even n, and for odd n to odd n. This is called a *selection rule.* In *spectroscopy,* there are selection rules which specify which transitions are possible. Transitions with *zero probability* are said to be *forbidden.* Transitions with non-zero probability are *allowed.* Since the above formula is actually an approximation - the *electric dipole approximation* - transitions that are allowed according to that formula are *dipole-allowed.*

We can use the square of the transition integral to determine ratios of transition probabilities and thereby predict the relative heights of absorption peaks

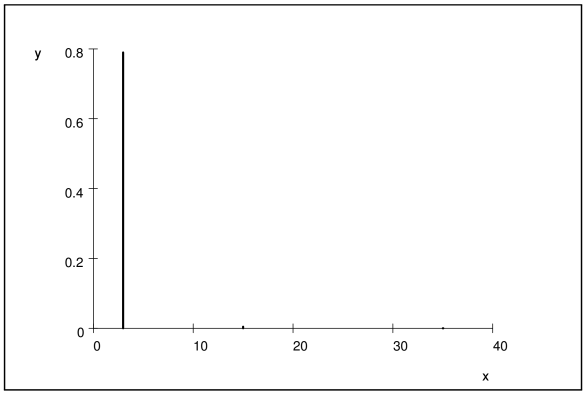


Figure 3.4: Absorption spectrum of a particle in a box in its ground state. Frequency - along horizontal axis - is in units of E_1/h.

in a measured spectrum. For example, the $n = 1$ to 4 transition integral takes the form,

$$\begin{aligned} & \frac{L}{\pi^2}\int_0^{\pi} u\left[\cos\left(3u\right)-\cos\left(5u\right)\right]\,du \\ = & \frac{L}{\pi^2}\left[-\frac{2}{3^2}+\frac{2}{5^2}\right]=-\frac{2L}{\pi^2}\left[\frac{1}{9}-\frac{1}{25}\right]=-\frac{48L}{225\pi^2}. \end{aligned}$$

and the ratio of the 1 to 2 transition peak to the 1 to 4 transition peak is

$$\frac{\frac{256}{81}}{\frac{48^2}{225^2}} = 69.4$$

The 1 to 4 *overtone* transition is much weaker than the 1 to 2 transition. Subsequent overtones (1 to 6, 1 to 8, ...) are successively weaker. The fundamental transition, 1 to 2, dominates the absorption spectrum.

Figure 3.4 shows the absorption spectrum for an electron, in its ground state, in a one dimensional box.

Example 3.1. Expectation values for *superposition states* of a particle in a one dimensional box.

(a) Determine the expectation value of x for the particle in the superposition state,

$$\psi_+\left(x\right)=\frac{1}{\sqrt{2}}\left(\psi_1\left(x\right)+\psi_2\left(x\right)\right),$$

where ψ_1 and ψ_2 are the first two energy eigenstates.

(b) Determine the expectation value of x for the particle in the superposition state,

$$\psi_- (x) = \frac{1}{\sqrt{2}} \left(\psi_1 (x) - \psi_2 (x) \right),$$

(c) Determine the expectation value of energy (i.e., $\hat{H}$) for the particle in either ψ_+ or ψ_-.

Solution 3.1.

(a) The expectation of position for this particle is

$$\begin{aligned} \langle x \rangle_+ &= \left\langle \psi_+ \left| x \psi_+ \right. \right\rangle \\ &= \left\langle \frac{1}{\sqrt{2}} \left(\psi_1 + \psi_2 \right) \right| \left. x \frac{1}{\sqrt{2}} \left(\psi_1 + \psi_2 \right) \right\rangle \\ &= \frac{1}{2} \left\langle \psi_1 + \psi_2 \left| x \left(\psi_1 + \psi_2 \right) \right. \right\rangle \\ &= \frac{1}{2} \left(\left\langle \psi_1 \left| x \psi_1 \right. \right\rangle + \left\langle \psi_2 \left| x \psi_2 \right. \right\rangle + \left\langle \psi_1 \left| x \psi_2 \right. \right\rangle + \left\langle \psi_2 \left| x \psi_1 \right. \right\rangle \right) \\ &= \frac{1}{2} \left(\langle x \rangle_1 + \langle x \rangle_2 + 2 \left\langle \psi_2 \left| x \psi_1 \right. \right\rangle \right) \qquad x \text{ is Hermitian} \end{aligned}$$

The expectation value for the particle in state ψ_1 or ψ_2 is determined above to be

$$\langle x \rangle_1 = \langle x \rangle_2 = \frac{L}{2},$$

where L is the width of the box. The matrix element, $\left\langle \psi_2 \left| x \psi_1 \right. \right\rangle$, is also determined above;

$$\left\langle \psi_2 \left| x \psi_1 \right. \right\rangle = -\frac{16L}{9\pi^2}.$$

Thus,

$$\begin{aligned} \langle x \rangle_+ &= \frac{1}{2} \left(L + 2 \left(-\frac{16L}{9\pi^2} \right) \right) \\ &= 0.320L. \end{aligned}$$

The particle is more likely to be found closer to the left side of the box than the right.

(b) Here,

$$\begin{aligned} \langle x \rangle_- &= \left\langle \psi_- \left| x \psi_- \right. \right\rangle \\ &= \left\langle \frac{1}{\sqrt{2}} \left(\psi_1 - \psi_2 \right) \right| \left. x \frac{1}{\sqrt{2}} \left(\psi_1 - \psi_2 \right) \right\rangle \\ &= \frac{1}{2} \left(\langle x \rangle_1 + \langle x \rangle_2 - 2 \left\langle \psi_2 \left| x \psi_1 \right. \right\rangle \right) \\ &= \frac{1}{2} \left(L - 2 \left(-\frac{16L}{9\pi^2} \right) \right) \\ &= 0.680L. \end{aligned}$$

In ths case, the particle is more likely to be found closer to the right side of box.

(c) The ψ_+ energy expectation is given by

$$\begin{aligned}\left\langle\hat{H}\right\rangle_+ &= \left\langle\psi_+\left|\hat{H}\psi_+\right.\right\rangle \\ &= \left\langle\frac{1}{\sqrt{2}}\left(\psi_1+\psi_2\right)\left|\hat{H}\frac{1}{\sqrt{2}}\left(\psi_1+\psi_2\right)\right.\right\rangle \\ &= \frac{1}{2}\left(\left\langle\hat{H}\right\rangle_1+\left\langle\hat{H}\right\rangle_2+2\left\langle\psi_2\left|\hat{H}\psi_1\right.\right\rangle\right).\end{aligned}$$

Since ψ_1 and ψ_2 are energy eigenstates, the terms in the last line simplify. Specifically,

$$\begin{aligned}\left\langle\hat{H}\right\rangle_1 &= \left\langle\psi_1\left|\hat{H}\psi_1\right.\right\rangle \\ &= \left\langle\psi_1\left|E_1\psi_1\right.\right\rangle && \begin{array}{c}\psi_1 \text{ is an eigenfunction of } \hat{H} \\ E_1 \text{ is the associated eigenvalue}\end{array} \\ &= E_1\left\langle\psi_1\left|\psi_1\right.\right\rangle && E_1 \text{ is a constant} \\ &= E_1, && \psi_1 \text{ is normalized}\end{aligned}$$

$$\begin{aligned}\left\langle\hat{H}\right\rangle_2 &= \left\langle\psi_2\left|\hat{H}\psi_2\right.\right\rangle \\ &= E_2 \qquad \text{same argument as above}\end{aligned}$$

and

$$\begin{aligned}\left\langle\psi_2\left|\hat{H}\psi_1\right.\right\rangle &= \left\langle\psi_2\left|E_1\psi_1\right.\right\rangle \\ &= E_1\left\langle\psi_2\left|\psi_1\right.\right\rangle \\ &= 0. \qquad \psi_1 \text{ and } \psi_2 \text{ are orthogonal}\end{aligned}$$

The off-diagonal matrix element of $\hat{H}$, $\left\langle\psi_2\left|\hat{H}\psi_1\right.\right\rangle$, is zero because ψ_1 and ψ_2 are eigenfunctions of $\hat{H}$, and distinct eigenfunctions of a Hermitian operator are orthogonal. Thus,

$$\begin{aligned}\left\langle\hat{H}\right\rangle_+ &= \frac{1}{2}\left(E_1+E_2+2\times 0\right) \\ &= \frac{1}{2}\left(E_1+E_2\right)\end{aligned}$$

Similarly,

$$\begin{aligned}\left\langle\hat{H}\right\rangle_- &= \left\langle\psi_-\left|\hat{H}\psi_-\right.\right\rangle \\ &= \left\langle\frac{1}{\sqrt{2}}\left(\psi_1-\psi_2\right)\left|\hat{H}\frac{1}{\sqrt{2}}\left(\psi_1-\psi_2\right)\right.\right\rangle \\ &= \frac{1}{2}\left(\left\langle\hat{H}\right\rangle_1+\left\langle\hat{H}\right\rangle_2-2\left\langle\psi_2\left|\hat{H}\psi_1\right.\right\rangle\right) \\ &= \frac{1}{2}\left(E_1+E_2\right).\end{aligned}$$

These two superposition states have the same energy expectation value - the average of the associated energy eigenvalues.

Example 3.2. Measurements of observables for *superposition states* of a particle in a one dimensional box.

(a) Suppose the particle is in the superposition state,

$$\varphi_1(x) = \frac{\sqrt{3}}{2}\psi_1(x) + \frac{1}{2}\psi_2(x).$$

What are the possible outcomes of an energy measurement for this particle? What are the probabilities of the possible outcomes of an energy measurement?

(b) What is the probability of a measurement of position yielding a value to the left of the center of the box?

(c) Repeat parts a and b for a particle in the superposition,

$$\varphi_2(x) = \frac{1}{2}\psi_1(x) - \frac{\sqrt{3}}{2}\psi_2(x).$$

Solution 3.2.

(a) The possible outcomes of an energy measurement are the eigenvalues of the Hamiltonian, the energy operator. For a particle in a box, these are the energies, $E_n,\ n = 1,\, 2,\, \ldots;$

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}n^2 = E_1 n^2.$$

The probability of outcome E_n is given by the *mod-squared* inner product of the eigenstate, $\psi_n(x)$, associated with eigenvalue, E_n. Specifically,

$$\begin{aligned} P(E_n) &= \left|\langle\psi_n|\varphi_1\rangle\right|^2 \\ &= \left|\left\langle\psi_n\left|\frac{\sqrt{3}}{2}\psi_1 + \frac{1}{2}\psi_2\right.\right\rangle\right|^2 \\ &= \left|\frac{\sqrt{3}}{2}\langle\psi_n|\psi_1\rangle + \frac{1}{2}\langle\psi_n|\psi_2\rangle\right|^2 \\ &= \begin{cases} \left|\frac{\sqrt{3}}{2}\right|^2, & n = 1 \\ \left|\frac{1}{2}\right|^2, & n = 2 \\ 0, & n > 2 \end{cases} \\ &= \begin{cases} \frac{3}{4}, & n = 1 \\ \frac{1}{4}, & n = 2 \\ 0, & n > 2 \end{cases}. \end{aligned}$$

These probabilities result because the distinct ψ_n are orthogonal and normalized. We see that only E_1 and E_2 are possible outcomes of an energy measurement for this state - the probability of the other outcomes is zero. The probability of outcome E_1 is 3/4 which is just the mod-squared coefficient of ψ_1 in the expansion of φ_1 in terms of the ψ_n. Similarly, the probability of outcome E_2 is 1/4, the mod-squared coefficient of ψ_2.

(b) The probability of observing position less than $L/2$ is given by integrating $|\varphi_1(x)|^2$ from 0 to $L/2$.

$$\begin{aligned}
P(0,L/2) &= \int_0^{L/2} |\varphi_1(x)|^2\,dx \\
&= \int_0^{L/2} \left|\frac{\sqrt{3}}{2}\psi_1(x)+\frac{1}{2}\psi_2(x)\right|^2 dx \\
&= \frac{3}{4}\int_0^{L/2}\psi_1^2(x)\,dx+\frac{1}{4}\int_0^{L/2}\psi_2^2(x)\,dx+ \qquad \psi_1 \text{ and } \psi_2 \text{ are real} \\
&\quad +2\frac{\sqrt{3}}{4}\int_0^{L/2}\psi_1(x)\,\psi_2(x)\,dx
\end{aligned}$$

To evaluate this probability, we must evaluate three integrals. The first two integrals give the probability of observing $x < L/2$ for the states ψ_1 and ψ_2. The associated two contributions to $P(0,L/2)$ together give the weighted average of these two probabilities, with weightings given by the probabilities determined in part a. Since $\psi_1^2(x)$ and $\psi_2^2(x)$ are both symmetric about $x = L/2$, the probability of observing $x < L/2$ for the states ψ_1 and ψ_2 is just 1/2. It is the same for both states. Thus, the first two contributions to $P(0,L/2)$ together give 1/2. The third integral is given by

$$\begin{aligned}
\int_0^{L/2}\psi_1(x)\,\psi_2(x)\,dx &= \frac{2}{L}\int_0^{L/2}\sin\left(\frac{\pi x}{L}\right)\sin\left(\frac{2\pi x}{L}\right)dx \\
&= \frac{1}{L}\int_0^{L/2}\left(\cos\left(\frac{(2-1)\,\pi x}{L}\right)-\cos\left(\frac{(2+1)\,\pi x}{L}\right)\right)dx \\
&= \frac{1}{\pi}\int_0^{\pi/2}(\cos(u)-\cos(3u))\,du \\
&= \frac{1}{\pi}\left([\sin(u)]_0^{\pi/2}-\left[\frac{1}{3}\sin(3u)\right]_0^{\pi/2}\right) \\
&= \frac{1}{\pi}\left([1-0]-\frac{1}{3}[-1-0]\right) \\
&= \frac{4}{3\pi}.
\end{aligned}$$

Therefore, the probability of observing $x < L/2$ is

$$\begin{aligned} P(0, L/2) &= \frac{3}{4}\frac{1}{2} + \frac{1}{4}\frac{1}{2} + \frac{\sqrt{3}}{2}\left(\frac{4}{3\pi}\right) \\ &= \frac{1}{2} + \frac{2\sqrt{3}}{3\pi} = 0.86755. \end{aligned}$$

In this state, the particle is more likely to be closer to the left wall at $x = 0$.

(c) The probability of an energy measurement yielding E_n is

$$\begin{aligned} P_n &= \left|\langle\psi_n \left|\varphi_2\right\rangle\right|^2 \\ &= \left|\left\langle\psi_n \left|\frac{1}{2}\psi_1 - \frac{\sqrt{3}}{2}\psi_2\right.\right\rangle\right|^2 \\ &= \left|\frac{1}{2}\langle\psi_n \left|\psi_1\right\rangle - \frac{\sqrt{3}}{2}\langle\psi_n \left|\psi_2\right\rangle\right|^2 \\ &= \begin{cases} \left|\frac{1}{2}\right|^2, & n = 1 \\ \left|\frac{\sqrt{3}}{2}\right|^2, & n = 2 \\ 0, & n > 2 \end{cases} \\ &= \begin{cases} \frac{1}{4}, & n = 1 \\ \frac{3}{4}, & n = 2 \\ 0, & n > 2 \end{cases}. \end{aligned}$$

The same two energies, E_1 and E_2, are the possible outcomes. They are observed with probabilities 1/4 and 3/4, respectively. These probabilities are the reverse of those found in part a, as the ψ_1 and ψ_2 coefficients are reversed. The minus sign makes no contribution to the probability of E_2, as probability is given by the mod-squared coefficient. The probability of a position measurement outcome less than $L/2$ is

$$\begin{aligned}
P(0,L/2) &= \int_0^{L/2} |\varphi_2(x)|^2\, dx \\
&= \int_0^{L/2} \left|\frac{1}{2}\psi_1(x) - \frac{\sqrt{3}}{2}\psi_2(x)\right|^2 dx \\
&= \frac{1}{4}\int_0^{L/2} \psi_1^2(x)\, dx + \frac{3}{4}\int_0^{L/2} \psi_2^2(x)\, dx + \qquad \psi_1 \text{ and } \psi_2 \text{ are real} \\
&\quad -2\frac{\sqrt{3}}{4}\int_0^{L/2} \psi_1(x)\psi_2(x)\, dx \\
&= \frac{1}{2} - \frac{\sqrt{3}}{2}\left(\frac{4}{3\pi}\right) \\
&= \frac{1}{2} - \frac{2\sqrt{3}}{3\pi} = 0.13245
\end{aligned}$$

In this case, the particle is more likely to nearer the right wall at $x = L$.

3.2 A well with finite binding energy

Bound particles are everywhere. However, for real systems, the particles have a finite *binding energy*. Particles can be ejected from a well by absorbing energy - most often by absorbing light. The lowest that can eject the particle is called the binding energy of the particle in the well. Here, we consider a one-sided well with finite depth, V_d;

$$V(x) = \begin{cases} \infty, & x \le 0 \\ 0, & 0 < x < L \\ V_\mathrm{d}, & L < x \end{cases} .$$

For this potential, the particle can only escape the well to the right. The well depth is larger than the binding energy because the particle has energy in its ground state - c.f., the case of a one dimensional box.

Just like the particle in the box, we solve the Schrödinger equation inside the well and apply the boundary condition, $\psi(0) = 0$ (there is still a hard wall at $x = 0$) to get

$$\psi(x) = C_1 \sin\left(\frac{\sqrt{2mE}x}{\hbar}\right),$$

where C_1 is (so far) an arbitrary constant. The boundary condition at $x = L$ is different in this case. The wavefunction is not zero at $x = L$, because the barrier on the right is finite. First, we must consider the wavefunction beyond the right edge of the box. For $x > L$, the Schrödinger equation takes the form,

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

which rearranges to

$$\begin{aligned}\frac{d^2\psi}{dx^2} &= -\frac{2m\left(E - V_{\mathrm{d}}\right)}{\hbar^2}\psi \\ &= \frac{2m\left(V_{\mathrm{d}} - E\right)}{\hbar^2}\psi.\end{aligned}$$

If $E < V_{\mathrm{d}}$ (the states of the particle bound to the well), $V_{\mathrm{d}} - E > 0$ and the general solution to this equation is

$$\psi\left(x\right) = A\exp\left(\frac{\sqrt{2m\left(V_{\mathrm{d}} - E\right)}x}{\hbar}\right) + B\exp\left(-\frac{\sqrt{2m\left(V_{\mathrm{d}} - E\right)}x}{\hbar}\right),$$

where A and B are new (so far) arbitrary constants. Because this region extends to $x = \infty$, the coefficient, A, must equal zero. Otherwise, the wavefunction would blow up as $x \to \infty$. This is a boundary condition at infinity - solutions which blow up are discarded. This leaves only the exponentially decaying solution for $x > L$. Altogether, we have

$$\begin{aligned}\psi\left(x\right) &= \begin{cases}\psi^{(1)}\left(x\right), & 0 < x < L \\ \psi^{(2)}\left(x\right), & L < x\end{cases} \\ &= \begin{cases} C_1\sin\left(\frac{\sqrt{2mE}x}{\hbar}\right), & 0 < x < L \\ C_2\exp\left(-\frac{\sqrt{2m(V_{\mathrm{d}}-E)}(x-L)}{\hbar}\right), & L < x\end{cases},\end{aligned}$$

in terms of the more convenient abitrary constant,

$$C_2 = B\exp\left(-\frac{\sqrt{2m\left(V_{\mathrm{d}} - E\right)}L}{\hbar}\right).$$

The exponential decay portion of the wavefunction is called *barrier penetration.* In quantum mechanics, particles penetrate into regions where they are *classically forbidden.* In classical mechanics, a particle cannot be in a place, x, where $E < V\left(x\right)$. This is because kinetic energy cannot be negative. When wavefunctions penetrate into classically forbidden regions, they decay exponentially. A larger difference between the barrier and particle energy, or larger mass of particle, produces a faster exponential decay.

3.2.1 Energy levels

Since the wavefunction must be smooth, it is subject to boundary conditions at $x = L$. First, the values of $\psi^{(1)}\left(x\right)$ and $\psi^{(2)}\left(x\right)$ must match at $x = L$;

$$\psi^{(1)}\left(L\right) = \psi^{(2)}\left(L\right),$$

or

$$C_1 \sin\left(\frac{\sqrt{2mE}L}{\hbar}\right) = C_2. \tag{3.11}$$

This reduces the number of arbitrary constants back to one. Next, the derivatives of $\psi^{(1)}(x)$ and $\psi^{(2)}(x)$ must match at $x = L$. This gives the smoothest possible wavefunction at a discontinuity in the potential.

$$\left.\frac{d\psi^{(1)}}{dx}\right|_{x=L} = \left.\frac{d\psi^{(2)}}{dx}\right|_{x=L},$$

or

$$\frac{\sqrt{2mE}}{\hbar} C_1 \cos\left(\frac{\sqrt{2mE}L}{\hbar}\right) = -\frac{\sqrt{2m\left(V_{\mathrm{d}} - E\right)}}{\hbar} C_2. \tag{3.12}$$

Dividing Eq. 3.11 by Eq. 3.12 gives

$$\frac{\hbar}{\sqrt{2mE}} \tan\left(\frac{\sqrt{2mE}L}{\hbar}\right) = -\frac{\hbar}{\sqrt{2m\left(V_{\mathrm{d}} - E\right)}}$$

or

$$\tan\left(\frac{\sqrt{2mE}L}{\hbar}\right) = -\sqrt{\frac{E}{V_{\mathrm{d}} - E}}. \tag{3.13}$$

This is an equation only for the energy. It is the *quantization condition* for the particle in a (one-sided) well. It is a non-linear equation that requires a non-linear equation solver - see below. However, its solutions can be seen graphically - see Fig. 3.5. The intersections of the tangent function on the left of Eq. 3.13 with the square root function on the right determine the energy levels of the particle in a well. The functions are plotted versus the square root of energy in units of ϵ_1, where ϵ_1 is the ground state energy of the particle in the same width box - i.e.,

$$\epsilon_1 = \frac{\pi^2 \hbar^2}{2mL^2}.$$

The number of intersections is finite because the square root function on the right of Eq. 3.13 is real only if $E < V_{\mathrm{d}}$. There are no bound states with energy above the well depth. This is the biggest contrast with the particle in a box. A finite well depth produces a finite number of bound states. The figure shows that the number of bound states is just $\left(V_{\mathrm{d}}/\epsilon_1\right)^{1/2}$, rounded to the nearest whole number. The energy levels of the particle in a well are lower than the corresponding particle in a box levels - i.e.,

$$E_n \lesssim n^2 \epsilon_1, \qquad n = 1, 2, \ldots [(V_{\mathrm{d}}/\epsilon_1)^{1/2}].$$

A more detailed analysis can provide formulas. In terms of

$$\eta = \frac{\sqrt{2mE}}{\pi\hbar} L,$$

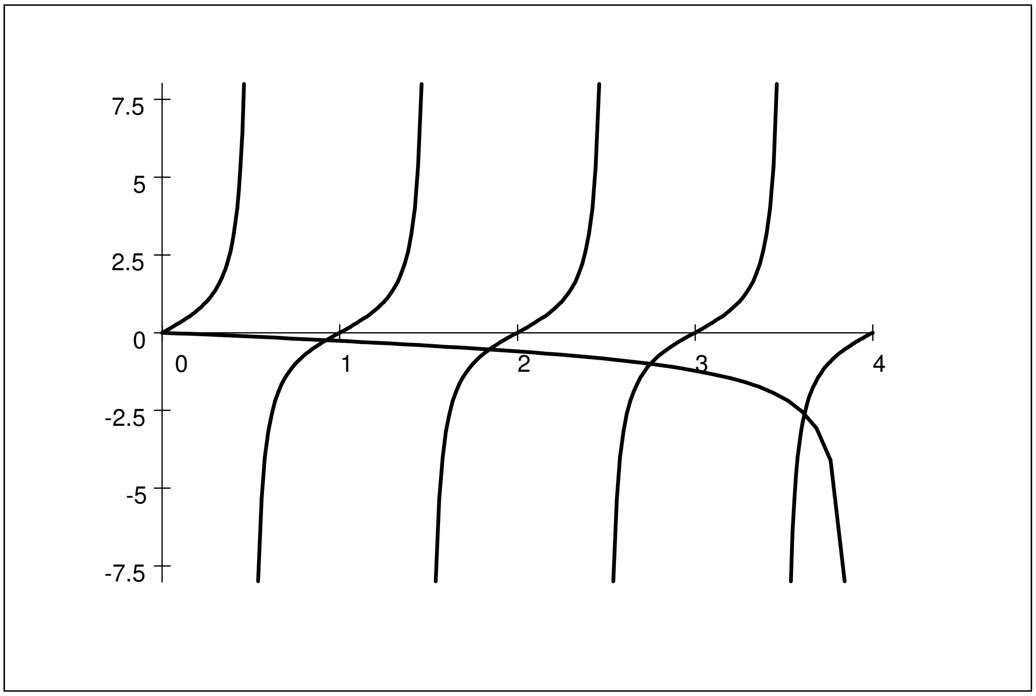


Figure 3.5: The quantization condition for a particle in a well. The intersections of the tangent function (solid) and the square root function (dashed) determine the energy levels. The well depth, $V_0 = 15\epsilon_1$, where $\epsilon_1 = \pi^2\hbar^2/\left(2mL^2\right)$ is the ground state the particle in a box with the same width. The horizontal axis here is $\left(E/\epsilon_1\right)^{1/2}$.

Eq. 3.13 takes the form,

$$\tan(\pi\eta) = -\frac{\eta}{\sqrt{\eta_{\mathrm{d}}^2 - \eta^2}}, \tag{3.14}$$

where

$$\eta_{\mathrm{d}} = \frac{\sqrt{2mV_{\mathrm{d}}}}{\pi\hbar}L.$$

The particle in a box energy levels arises if $V_0 \to \infty$. In this case, Eq. 3.14 reduces to

$$\tan\left(\pi\eta^{(0)}\right) = 0,$$

with solutions

$$\eta_n^{(0)} = n, \qquad n = 1, 2, 3, \ldots.$$

Figure 3.5 shows that the energy levels of the well are lower than, but close to, the corresponding energy levels of the box. It also shows that the intersections can be approximated via linear approximation of the tangent function. Specifically, we let

$$\begin{aligned} \eta_n &= \eta_n^{(0)} - \delta_n \\ &= n - \delta_n \end{aligned}$$

aniticipating that δ_n will be positive. Since the derivative of $\tan(x)$ is $\sec^2(x)$, the left side of Eq. 3.14 takes the form

$$\begin{aligned} \tan(\pi\eta_n) &\cong \tan\left(\pi\eta_n^{(0)}\right) - \pi\sec^2\left(\pi\eta_n^{(0)}\right)\delta_n \\ &= 0 - \pi\sec^2(n\pi)\,\delta_n \\ &= -\pi\delta_n. \end{aligned}$$

With this approximation, Eq. 3.14 becomes

$$-\pi\delta_n \cong -\frac{n - \delta_n}{\sqrt{\eta_{\mathrm{d}}^2 - (n - \delta_n)^2}},$$

which rearranges to

$$(n - \delta_n)^2 \cong \frac{\pi^2\delta_n^2\eta_{\mathrm{d}}^2}{\left(1 + \pi^2\delta_n^2\right)},$$

or

$$\begin{aligned} \delta_n &\cong n - \frac{\pi\delta_n\eta_{\mathrm{d}}}{\left(1 + \pi^2\delta_n^2\right)^{1/2}} \\ &\cong n - \pi\delta_n\eta_{\mathrm{d}} \qquad \delta_n^2 \text{ is much smaller than } 1/\pi^2 \end{aligned} \tag{3.15}$$

This equation is easily solved to give

$$\delta_n \cong \frac{n}{1 + \pi\eta_{\mathrm{d}}} < \frac{1}{\pi}. \tag{3.16}$$

This approximation can be refined by iteration of the first line of Eq. 3.15 . We are content with this lowest order approximation for the sake of simplicity. The associated energy values - the energy eigenvalues - are given by

$$\begin{aligned} E_n &= \eta_n^2 \epsilon_1 \\ &\cong (n - \delta_n)^2 \epsilon_1 \\ &= n^2 \left(\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}} + 1/\pi} \right)^2 \epsilon_1 . \end{aligned} \tag{3.17}$$

They still vary in proportion to n^2. All the energy levels - except the highest - have approximately the same *relative* shift. Here, we see that the energy levels of a finite well are lower than those of the particle in a box. All levels are lowered. The highest energy levels are lowered the most. In general, energy levels are lowered when a barrier is lowered. This is because of barrier penetration. The particle is able to access a wider space. Just as a wider box has lower energy levels than a narrower box, a well that has barrier penetration has lower energy levels.

3.2.2 Bound state energy eigenfunctions

Consider a finite well wavefunction with energy much lower than V_0. In this case, the energy is close to a particle in a box energy eigenvalue, and the wavefunction is similarly close to the particle in a box wavefunction. Even the normalization factor is almost the same - it is close to $\sqrt{2/L}$. It is a little smaller because the integral of $\sin^2(\pi\eta_n x/L)$ over $x \in (0, L)$ is larger when $\eta_n < n$, and there is an additional contribution for $x > L$. The latter is the integral of the exponentially decaying function with a small prefactor,

$$|C_2|^2 \exp\left(-2\sqrt{2m\left(V_{\mathrm{d}} - E_n\right)}\left(x - L\right)/\hbar \right),$$

from $x = L$ to $x = \infty$. Since both corrections are small, we approximate the energy eigenfunction of the particle in a finite well using the particle in a box normalization constant, and let $C_1 = \sqrt{2/L}$. In this case,

$$\begin{aligned} C_2 &= \sqrt{\frac{2}{L}}\sin\left(\frac{\sqrt{2mE_n}L}{\hbar}\right) \\ &= \sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right) && (3.18) \\ &\cong \sqrt{\frac{2}{L}}\left(\sin\left(\eta_n^{(0)}\pi\right) - \pi\cos\left(\eta_n^{(0)}\pi\right)\delta_n\right) && (3.19) \\ &= \sqrt{\frac{2}{L}}\left(-1\right)^{n+1}\pi\delta_n && (3.20) \\ &= \sqrt{\frac{2}{L}}\left(-1\right)^{n+1}\frac{\pi n}{\pi\eta_{\mathrm{d}}+1} && (3.21) \\ &\cong \sqrt{\frac{2}{L}}\frac{\left(-1\right)^{n+1}n}{\eta_{\mathrm{d}}} \qquad \text{if } 1 \ll \eta_{\mathrm{d}} && (3.22) \end{aligned}$$

and

$$\psi_n\left(x\right) \cong \begin{cases} \sqrt{\frac{2}{L}}\sin\left(\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}}+1/\pi}\frac{n\pi x}{L}\right), & 0 < x < L \\ \sqrt{\frac{2}{L}}\frac{(-1)^{n+1}n}{\eta_{\mathrm{d}}+1/\pi}\exp\left(-\left(\eta_{\mathrm{d}}^2 - n^2\left(\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}}+1/\pi}\right)^2\right)^{1/2}\frac{\pi(x-L)}{L}\right), & L < x \end{cases} . \tag{3.23}$$

These wavefunctions are depicted in Fig. 3.6. The error in the approximations used to get Eqs. 3.17 and 3.23 , is evident only in the top level wavefunction which does not quite satisfy the boundary condition at $x = 1$ - if you look closely.

3.2.3 State with energy above binding energy

For energy greater than the binding energy, the particle is not bound to the box. If $E > V_{\mathrm{d}}$, $E - V_{\mathrm{d}} > 0$ and the general solution to the Schrödinger equation for $x > L$ is

$$\psi\left(x\right) = A\exp\left(i\frac{\sqrt{2m\left(E-V_{\mathrm{d}}\right)}x}{\hbar}\right) + B\exp\left(-i\frac{\sqrt{2m\left(E-V_{\mathrm{d}}\right)}x}{\hbar}\right),$$

or

$$\psi\left(x\right) = C_2\exp\left(i\frac{\sqrt{2m\left(E-V_{\mathrm{d}}\right)}\left(x-L\right)}{\hbar}\right) + C_3\exp\left(-i\frac{\sqrt{2m\left(E-V_{\mathrm{d}}\right)}\left(x-L\right)}{\hbar}\right),$$

where C_2 and C_3 (or A and B) are new (so far) arbitrary constants. Neither of the two solutions blows up as $x \to \infty$. However, neither do they decay. As such, the integral of $\left|\psi\left(x\right)\right|^2$ over all x (up to $x = \infty$) cannot be finite. It is not possible to normalize this wavefunction. When dealing with unbound particles,

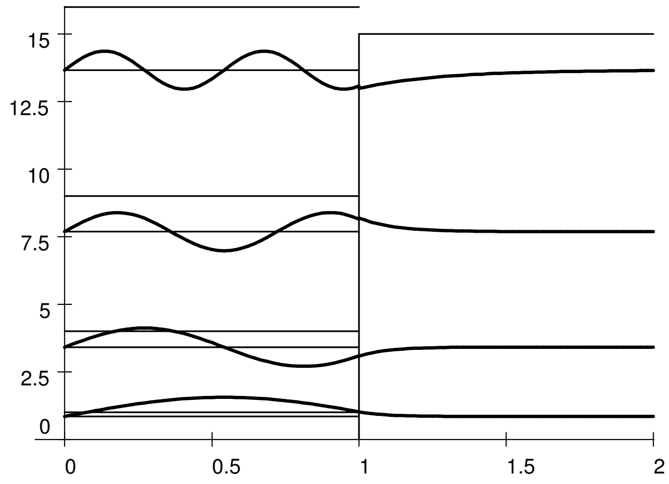


Figure 3.6: The four bound energy eigenstates of a particle in a one dimensional well. The wavefunctions (solid) are show superimposed on a potential energy plot (dashed). There is separate vertical axis for each wavefunction - the wavefunction zero is translated vertically to match the associated energy eigenvalue. The wavefunctions are scaled by 1/2, so they do not overlap. The energy levels are depicted with thin dashed lines, while the corresponding particle in a box energy levels are shown as dotted lines. Eigenfunctions and eigenvalues are computed using Eqs. 3.23 and 3.17 , respectively.

the postulate of quantum mechanics requiring that $\|\psi\|^2 = 1$ is not satisfied by the energy eigenstates. Unbound energy eigenstates are unnormalizable. They do not describe single particles. They describe *beams* of particles. These *beam states* are spread over an infinite distance, and they have infinite norm. However, they are *beam-normalized*: $|\psi(x)|^2$ gives the number of particles per unit distance, at x.

Because neither of the two arbitrary constants, C_2 and C_3, is eliminated by the boundary condition at $x = \infty$, there is one less constraint on the coefficients than there were in the case of bound states. Energy is not quantized for states above the binding energy - it varies continuously, just as in classical mechanics.

The boundary conditions at $x = L$ relate unknown coefficient C_1 to unknown coefficients, C_2 and C_3. Specifically,

$$\psi^{(1)}(L) = \psi^{(2)}(L),$$

gives

$$C_1 \sin(kL) = C_2 + C_3, \tag{3.24}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}$$

is the wavenumber that correponds with energy, E. Also,

$$\left.\frac{d\psi^{(1)}}{dx}\right|_{x=L} = \left.\frac{d\psi^{(2)}}{dx}\right|_{x=L},$$

gives

$$kC_1 \cos(kL) = i\sqrt{k^2 - k_{\mathrm{d}}^2}\,(C_2 - C_3), \tag{3.25}$$

where

$$k_{\mathrm{d}} = \frac{\sqrt{2mV_{\mathrm{d}}}}{\hbar}$$

is the well-depth wavenumber.

Equations 3.24 and 3.25 constrain two of the three unknowns. However, there is always an overall undetermined constant in the solution to the Schrödinger equation. It is fixed (to within a phase factor, $e^{i\theta}$) by the normalization condition. Here, the beam normalization condition is simply to set one of C_2 and C_3 to one. We choose $C_3 = 1$. This corresponds to a unit intensity beam (one particle per unit distance) incoming from ∞. The wavefunction,

$$\exp(-ikx),$$

is a unit-intensity incoming beam wavefunction. It is an eigenfunction of the momentum operator (check this) associated with eigenvalue,

$$-\hbar k.$$

Negative momentum corresponds to motion from right to left. C_2 is the coefficient of the right-going beam wavefunction. It corresponds to the reflected (outgoing) beam that results subsequent to the incoming beam encountering the well and associated left wall.

With $C_3 = 1$, Eqs. 3.24 and 3.25 take the form,

$$C_1 \sin(kL) = C_2 + 1$$

and

$$kC_1 \cos(kL) = i\sqrt{k^2 - k_{\mathrm{d}}^2}\,(C_2 - 1)\,.$$

These two equations can be solved as two linear equations with two unknowns. Alternatively, note that

$$\frac{1}{k}\tan(kL) = -i\frac{1}{\sqrt{k^2 - k_{\mathrm{d}}^2}}\left(\frac{C_2 + 1}{C_2 - 1}\right).$$

So,

$$\frac{C_2 + 1}{C_2 - 1} = i\tau = i\frac{\sqrt{k^2 - k_{\mathrm{d}}^2}}{k}\tan(kL)$$

and

$$\begin{aligned}
C_2 &= -\frac{1 + i\tau}{1 - i\tau} && (3.26)\\
&= -\frac{k + i\sqrt{k^2 - k_{\mathrm{d}}^2}\tan(kL)}{k - i\sqrt{k^2 - k_{\mathrm{d}}^2}\tan(kL)} && (3.27)\\
&= -\frac{k\cos(kL) + i\sqrt{k^2 - k_{\mathrm{d}}^2}\sin(kL)}{k\cos(kL) - i\sqrt{k^2 - k_{\mathrm{d}}^2}\sin(kL)} && (3.28)\\
&= -\frac{\kappa + i\sigma}{\kappa - i\sigma} && (3.29)
\end{aligned}$$

where

$$\kappa = k\cos(kL)$$

and

$$\sigma = \sqrt{k^2 - k_{\mathrm{d}}^2}\sin(kL)\,.$$

Note that $|C_2|^2 = 1$; i.e., the intensity of the outgoing beam equals that of the incoming beam. This corresponds to particle conservation - there is an outgoing particle for every incoming particle.

Determining C_1 completes specification of the wavefunction.

$$\begin{aligned}
C_1 &= -\frac{2i\sigma}{(\kappa - i\sigma)\sin(kL)}\\
&= -2i\frac{\sqrt{k^2 - k_{\mathrm{d}}^2}}{\kappa - i\sigma}
\end{aligned}$$

and

$$\psi_E(x) = \begin{cases} 2i\frac{\sqrt{k^2-k_\mathrm{d}^2}}{\kappa-i\sigma}\sin(kx), & 0<x<L \\ -\frac{\kappa+i\sigma}{\kappa-i\sigma}\exp\left(i\sqrt{k^2-k_\mathrm{d}^2}\,(x-L)\right) & \\ +\exp\left(-i\sqrt{k^2-k_\mathrm{d}^2}\,(x-L)\right), & L<x \end{cases}. \tag{3.30}$$

It is instructive to evaluate the number of particles within the well. For beam normalized wavefunctions, the integral of $|\psi_E(x)|^2$ from 0 to L gives the number of particles in the well (associated with a unit intensity incoming beam). The integral of $\sin^2(kx)$ varies around, but always close to, $L/2$ (it equals $L/2$ exactly when $kL = n\pi$). The number of particles in the well is thus close to $|C_1|^2 L/2$ (with deviations diminishing as k increases). The the mean particle density within the well is

$$\begin{aligned} |C_1|^2/2 &= 2\frac{k^2-k_\mathrm{d}^2}{\kappa^2+\sigma^2} \\ &= 2\frac{k^2-k_\mathrm{d}^2}{k^2\cos^2(kL)+(k^2-k_\mathrm{d}^2)\sin^2(kL)} \\ &= 2\frac{k^2-k_\mathrm{d}^2}{k^2-k_\mathrm{d}^2\sin^2(kL)} \end{aligned}$$

It has a maximum of 2, when $\sin(kL) = \pm 1$, or $kL = (n+1/2)\pi$. In this case, $C_2 = 1$ - i.e., the phase of the outgoing wave is the same as the incoming wave, at the well boundary. Such energies are called *resonances*. They arise here when kL is a half-integral multiple of π. For unbound particles, the quantization condition is replaced by a resonance condition. A maximum mean particle density of 2 matches that of the incoming and outgoing beam mean particle density - the incoming and outgoing beam interfere to the right of the well producing a particle density varying from 0 to 4, with a mean of 2. Between the resonances, $|C_1|^2/2$ is a minimum when $\sin(kL) = 0$, or $kL = n\pi$. For these energies, $C_2 = -1$ - the incoming and outgoing waves are exactly out of phase at the well boundary, and the particle density inside the well is a minimum. At these energies, the incoming beam mostly reflects directly off the right boundary of the well, and fewer particles enter the well. Figure 3.7 shows energy eigenstates for beams at five different energies, $E = 4^2$, 4.25^2, 4.5^2, 5^2 and 5.5^2, in units of ϵ_1 the associated particle in a box ground state energy. The top and third-from-top of these eigenstates are resonance states, showing maximum well penetration. Between resonance energies - at the particle in a box energy levels - well penetration has local minima, most notably just above the well depth.

If the well were *two-sided* - i.e., the infinite barrier on the left is replaced by a V_d barrier, as on the right - then the resonance condition reverts back to the particle in a box quantization condition, $kL = n\pi$. At the particle in a box energies, a wave incoming from the right passes through the barrier with 100% likelihood - i.e., 100% transmission probability. At other energies, there

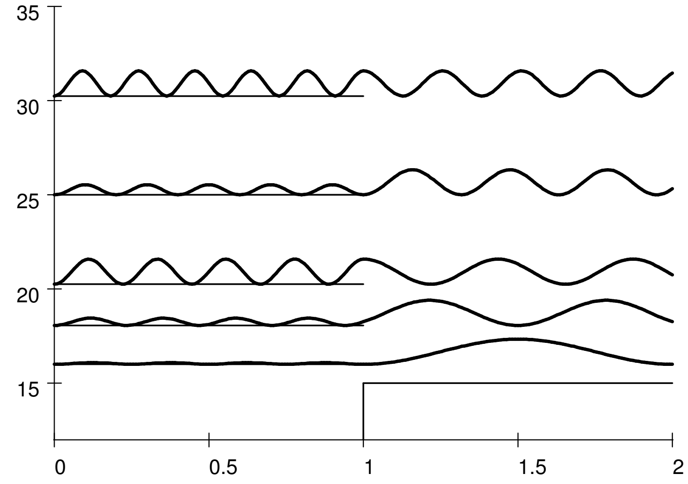


Figure 3.7: Some unbound energy eigenstates of a beam of particles encountering a one dimensional well. The wavefunctions (solid) are show superimposed on a potential energy plot (dashed). Here, the top of the well is just visible. There is separate vertical axis for each wavefunction - the wavefunction zero is translated vertically to match the associated energy eigenvalue. Since all energies above $V_{\rm d}$ are eigenvalues, only a few examples are shown. The wavefunctions are scaled by 1/3, so they do not overlap. The energy levels are depicted with thin dashed lines. Eigenfunctions are computed using Eq. 3.30 .

is some reflection. Reflection is a maximum at half integral multiples of π, in this case.

3.2.4 Tunneling

We now suppose that the barrier on the right side of the well just described has a finite width, w. Here, there is a second boundary at $x = L+w$, beyond which $V(x) = V_{\mathrm{e}} < V_{\mathrm{d}}$, the particle escape potential. Specifically,

$$V(x) = \begin{cases} \infty, & x \leq 0 \\ 0, & 0 < x < L \\ V_{\mathrm{d}}, & L < x < L+w \\ V_{\mathrm{e}}, & L+w < x \end{cases} .$$

We consider the below barrier - but above escape potential - case with $V_{\mathrm{e}} < E < V_{\mathrm{d}}$. In this case, were it not for the escape channel on right, the eigenstates would be the particle in a well bound states found in Sec. 3.2.2 above. To find the new energy eigenstates requires introducing two new constants, and satisfying two new boundary conditions. Anticipating that there will be resonances at the well energy eigenvalues between V_{e} and V_{d}, we consider the beam states of the particle only at these energies. In addition, rather than imposing a unit intensity incoming (or outgoing) beam from the left, we normalize by imposing that there be one particle in well. The resulting outgoing beam intensity is related to probability (per unit time) with which the particle tunnels from the well. Consequently, we adopt the following wavefunction - expressed in terms of four unknown coefficients - with the coefficient in the well specified by the one particle normalization - the wavefunction integrates to (close to) 1 within the well.

$$\psi_{E_n}(x) \cong \begin{cases} \sqrt{\frac{2}{L}} \sin\left(\frac{\eta_n \pi x}{L}\right), & 0 < x < L \\ C_1 \exp\left(-\left(\eta_{\mathrm{d}}^2 - \eta_n^2\right)^{1/2} \frac{\pi(x-L)}{L}\right) & \\ +C_2 \exp\left(\left(\eta_{\mathrm{d}}^2 - \eta_n^2\right)^{1/2} \frac{\pi(x-L)}{L}\right) & L < x < L+w \\ C_3 \exp\left(i\left(\eta_n^2 - \eta_{\mathrm{e}}^2\right)^{1/2} \frac{\pi(x-L-w)}{L}\right) & \\ +C_4 \exp\left(-i\left(\eta_n^2 - \eta_{\mathrm{e}}^2\right)^{1/2} \frac{\pi(x-L-w)}{L}\right) & L+w < x \end{cases} . \tag{3.31}$$

where

$$\eta_n = n \frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}} + 1/\pi}$$

and

$$\eta_{\mathrm{e}} = \frac{\sqrt{2mV_{\mathrm{e}}}}{\pi\hbar} L.$$

The boundary conditions at $x = L$ are the continuity equation,

$$\sqrt{\frac{2}{L}} \sin(\eta_n \pi) = C_1 + C_2,$$

and the continuity of derivative equation,

$$\sqrt{\frac{2}{L}}\eta_n \frac{\pi}{L}\cos\left(\eta_n\pi\right) = -\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi}{L}\left(C_1-C_2\right)$$

or (cancel π/L from the second equation and divide the equations)

$$\frac{C_1+C_2}{C_1-C_2} = -\frac{\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}}{\eta_n}\tan\left(\eta_n\pi\right) = -\tau$$

and

$$C_2 = -\frac{\tau+1}{\tau-1}C_1.$$

Substituting this back into the first equation gives,

$$\begin{aligned}\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right) &= \left(1-\frac{\tau+1}{\tau-1}\right)C_1\\ &= -\frac{2}{\tau-1}C_1\end{aligned}$$

or

$$C_1 = -\frac{\tau-1}{2}\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right)$$

and

$$C_2 = \frac{\tau+1}{2}\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right).$$

At the resonance energies, $\tau = -1$. In this case,

$$C_1 = \sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right)$$

and

$$C_2 = 0.$$

The energy eigenstate reduces to the form of the well bound state, at resonance energies.

The boundary conditions at $x = L + w$ are the continuity equation,

$$C_1\gamma + C_2\gamma^{-1} = C_3 + C_4$$

where

$$\gamma = \exp\left(-\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi w}{L}\right)$$

and the continuity of derivative equation,

$$-\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi}{L}\left(C_1\gamma-C_2\gamma^{-1}\right) = i\left(\eta_n^2-\eta_{\mathrm{e}}^2\right)^{1/2}\frac{\pi}{L}\left(C_3-C_4\right).$$

Under resonance conditions, these two equations become

$$\gamma\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right) = C_3 + C_4$$

and

$$i\left(\frac{\eta_{\mathrm{d}}^2-\eta_n^2}{\eta_n^2-\eta_{\mathrm{e}}^2}\right)^{1/2}\gamma\sin\left(\eta_n\pi\right) = C_3 - C_4,$$

or

$$\frac{C_3-C_4}{C_3+C_4} = i\left(\frac{\eta_{\mathrm{d}}^2-\eta_n^2}{\eta_n^2-\eta_{\mathrm{e}}^2}\right)^{1/2} = i\beta$$

and

$$C_4 = \frac{1-i\beta}{1+i\beta}C_3,$$

which gives

$$\begin{aligned}\gamma\sin\left(\eta_n\pi\right) &= C_3 + \frac{1-i\beta}{1+i\beta}C_3\\ &= \frac{2}{1+i\beta}C_3,\end{aligned}$$

or

$$C_3 = \frac{1+i\beta}{2}\gamma\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right)$$

and

$$C_4 = \frac{1-i\beta}{2}\gamma\sqrt{\frac{2}{L}}\sin\left(\eta_n\pi\right).$$

The intensity of the outgoing beam - i.e., the probability, per unit distance, of finding a particle in the beam - is just

$$\begin{aligned}\left|C_3\right|^2 &= \frac{1+\beta^2}{2L}\gamma^2\sin\left(\eta_n\pi\right) &(3.32)\\ &= \frac{\eta_{\mathrm{d}}^2-\eta_{\mathrm{e}}^2}{2L\left(\eta_n^2-\eta_{\mathrm{e}}^2\right)}\exp\left(-2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi w}{L}\right)\sin\left(\eta_n\pi\right)\\ &= \frac{V_{\mathrm{d}}-V_{\mathrm{e}}}{2L\left(E_n-V_{\mathrm{e}}\right)}\exp\left(-2\left(2m\left(V_{\mathrm{d}}-E_n\right)\right)^{1/2}\frac{w}{\hbar}\right)\sin\left(\left(2mE_n\right)^{1/2}\frac{L}{\hbar}\right).\end{aligned}$$

$\left|C_4\right|^2 = \left|C_3\right|^2$ gives the intensity of the associated incoming beam that maintains one particle in the well. If we multiply $\left|C_3\right|^2$ by the velocity of the outgoing particles, $\left(2\left(E_n-V_{\mathrm{e}}\right)/m\right)^{1/2}$, we get the outgoing flux - the probability per unit time of a particle *tunneling* from the n th state of the well,

$$\begin{aligned}P_{\text{tunnel}} &= \left(2\left(E_n-V_{\mathrm{e}}\right)/m\right)^{1/2}\left|C_3\right|^2 &(3.33)\\ &= \frac{V_{\mathrm{d}}-V_{\mathrm{e}}}{L\left(2m\left(E_n-V_{\mathrm{e}}\right)\right)^{1/2}}\exp\left(-2\left(2m\left(V_{\mathrm{d}}-E_n\right)\right)^{1/2}\frac{w}{\hbar}\right)\sin\left(\left(2mE_n\right)^{1/2}\frac{L}{\hbar}\right)\end{aligned}$$

The tunneling probability (flux) decreases exponentially with increasing barrier width, and the square root of the energy difference to the top of the barrier.

The system just studied is a simple model for *scanning tunneling microscopy (STM)*. In STM, a very sharp metal tip (represented above by the escape region to the right) is brought very close to the surface of another metal. Atoms on the surface are like wells for electrons. Tunneling current, $-eP_{\text{tunnel}}$, is measured as the tip is scanned across the metal surface. The distance between the tip and an atom on the surface is represented here by w. Because the exponential function is so sensitive to w, a distinctly larger current is measured when the tip is directly over an atom, than when it is over the space between atoms. The exponential sensitivity of STM gives it *atomic level resolution*. Individual atoms can be seen on the surface when the tunneling current, as a function of tip position over the surface, is itself rendered as a surface.

Figure 3.8 shows an STM image of iron atoms - arranged in the shape of a stadium - on a copper (111) surface. The stadium-shaped arrangement of iron atoms was achieved using the STM with high voltage, scanning very close to the surface. Under these conditions, the tip exerts sufficient force to move atoms around on the surface. The copper atoms in the surface are not well resolved because it is further from the tip and the current is weak. More importantly, because copper is a metal, it has delocalized valence electrons. These electrons are sensitive to the arrangement of iron atoms. These correspond to the highest energy resonance states of copper electrons on the surface. Consequently, they make the principle contribution to the STM current. These electrons appear as a wave pattern both inside and outside the stadium structure.

To better understand STM, we need more realistic models. In particular, we need three dimensional models. After all, ,space has (at least) three dimensions.

3.3 Particle in a three dimensional box

We return now to the particle in a box problem, except this time the particle is in a three dimensional box. It has three coordinates and its state is represented by a wavefunction, $\psi(x,y,z)$. The Hamiltonian for this problem is pure kinetic energy, with three components,

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\left(\hat{p}_x^2+\hat{p}_y^2+\hat{p}_z^2\right)\\ &= \frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right).\end{aligned}$$

The potential energy is zero inside the box. We account for the hard wall potential ($V=\infty$) at the walls of the box by the boundary condition that $\psi(x,y,z)=0$ on the walls.

The TISE for this problem is

$$\begin{aligned}\hat{H}\psi(x,y,z) &= E\psi(x,y,z) \qquad (3.34)\\ &= \frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi(x,y,z)\end{aligned}$$

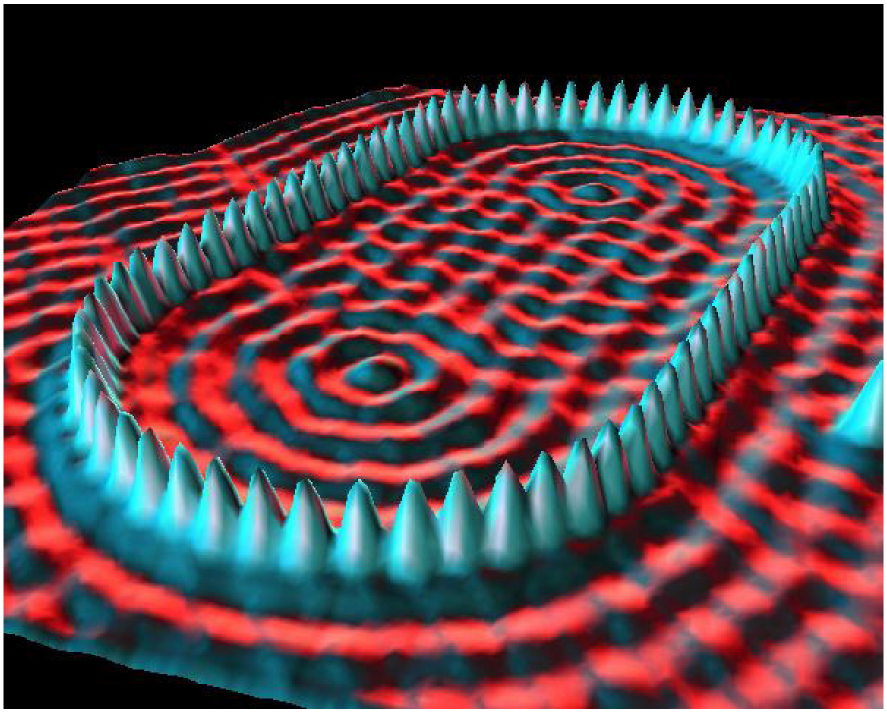


Figure 3.8: An STM image of iron atoms on a copper (111) surface. The iron atoms are arranged in the shape of a stadium. This was accomplished by the STM tip using high voltage and very close proximity to the surface. [From: IBM STM image gallery, http://researcher.watson.ibm.com/researcher/files/us-flinte/stm15.jpg, retrieved May 8, 2015]

Now, we have a partial differential equation. This one is soluble because each term in the Hamiltonian probes the dependence of ψ on just one of the three variables - i.e., one variable at a time. Suppose we can write $\psi(x, y, z)$ in the form,

$$\psi(x, y, z) = \psi_x(x)\,\psi_y(y)\,\psi_z(z)\,.$$

Substitute this *ansatz* into the TISE, and divide both sides by $\psi(x, y, z)$, to get

$$\begin{aligned} E &= \frac{-\hbar^2}{2m}\frac{\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi_x(x)\,\psi_y(y)\,\psi_z(z)}{\psi_x(x)\,\psi_y(y)\,\psi_z(z)} \\ &= \frac{-\hbar^2}{2m}\frac{\left(\psi_y(y)\,\psi_z(z)\,\frac{\partial^2}{\partial x^2}\psi_x(x)+\psi_x(x)\,\psi_z(z)\,\frac{\partial^2}{\partial y^2}\psi_y(y)+\psi_x(x)\,\psi_y(y)\,\frac{\partial^2}{\partial z^2}\psi_z(z)\right)}{\psi_x(x)\,\psi_y(y)\,\psi_z(z)} \\ &= \frac{-\hbar^2}{2m}\left(\frac{1}{\psi_x(x)}\frac{\partial^2}{\partial x^2}\psi_x(x)+\frac{1}{\psi_y(y)}\frac{\partial^2}{\partial y^2}\psi_y(y)+\frac{1}{\psi_z(z)}\frac{\partial^2}{\partial z^2}\psi_z(z)\right) \\ &= f(x)+g(y)+h(z)\,, \end{aligned}$$

where $f(x)$, $g(y)$ and $h(z)$ functions only of x, y and z, respectively. This equation can be rearranged to take the form

$$f(x) = E - g(y) - h(z)$$

Since E is a constant and neither g nor h is a function of x, $f(x)$ must be independent of x. We write

$$f(x) = E_x,$$

where E_x is some constant. Similarly,

$$g(y) = E_y$$

and

$$h(z) = E_z$$

are constants, independent of y and z, respectively.

Thus,

$$\frac{-\hbar^2}{2m}\frac{1}{\psi_x(x)}\frac{\partial^2}{\partial x^2}\psi_x(x) = E_x$$

or

$$\frac{-\hbar^2}{2m}\frac{d^2}{dx^2}\psi_x(x) = E_x\psi_x(x)\,. \qquad \text{since } \psi_x(x) \text{ only depends on } x$$

Also,

$$\frac{-\hbar^2}{2m}\frac{d^2}{dy^2}\psi_y(y) = E_y\psi_y(y)$$

and

$$\frac{-\hbar^2}{2m}\frac{d^2}{dz^2}\psi_z(z) = E_z\psi_z(z)\,.$$

We see that the partial differential equation in three variables separates into three independent ordinary differential equations. This is called *separation of variables.*

The total energy is the sum of three energy eigenvalues - one for each degree of freedom:

$$E = E_x + E_y + E_z.$$

Each of the individual energies is an eigenvalue of a particle in a one dimensional box. The energy eigenvalues are labeled by three quantum numbers, n_x, n_y and n_z, each varying from 1 to ∞ in unit steps (1, 2, 3, ...). If the box has length a in the x direction, b in the y direction and c in the z direction, the energy eigenvalues are

$$E_{n_x,n_y,n_z} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2}\right). \tag{3.35}$$

Each energy eigenvalue is the sum of three particle-in-a-one-dimensional-box eigenvalues. In general, every distinct set of natural numbers, (n_x, n_y, n_z), corresponds to a distinct energy. However, if the lengths satisfy certain relationships, there can be degenerate energy levels - i.e., more than one set of (n_x, n_y, n_z) giving the same energy. Most notably, if $a = b = c$, then

$$E_{n_x,n_y,n_z} = \frac{\hbar^2\pi^2}{2ma^2}\left(n_x^2 + n_y^2 + n_z^2\right),$$

and there are many degeneracies. First, any distinct permutation of the three quantum numbers produces the same energy. If all three quantum numbers are different, then there are six such permutations. There are even more degeneracies. For example, since $36 + 49 = 4 + 81$, the quantum numbers $(6, 7, n)$ (or any permutation) produce the same energy as $(2, 9, n)$ (or any permutation).

The eigenfunctions associated with the above energy eigenvalues are given by

$$\begin{aligned}\psi_{n_x,n_y,n_z}(x,y,z) &= \sqrt{\frac{2}{a}}\sin\left(\frac{n_x\pi x}{a}\right)\sqrt{\frac{2}{b}}\sin\left(\frac{n_y\pi y}{b}\right)\sqrt{\frac{2}{c}}\sin\left(\frac{n_z\pi z}{c}\right)\\ &= \frac{2^{3/2}}{\sqrt{abc}}\sin\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right)\sin\left(\frac{n_z\pi z}{c}\right).\end{aligned}$$

It is the product of three particle-in-a-one-dimensional-box eigenfunctions. That these are eigenfunctions of the Hamiltonian is easily verified by substituting this function into the TISE and noting that it satisfies the boundary conditions. To further support the initial *ansatz*, we note that it can be shown that any function of x, y and z can be written as a linear combination (summing over all values of n_x, n_y and n_z) of the above eigenfunctions. Thus, the ansatz is justified. All eigenfunctions either satisfy the ansatz, or are sums of such functions (only in case of degenerate eigenvalues). In the latter case, the product functions considered here provide a complete basis for the associated eigenspace - sums of product functions are not required.

3.3.1 Properties of a particle in three dimensional box

The inner product for states of a particle in three dimensions is an integral over all three coordinates, x, y and z. For example, the *expectation value* of x for a particle in the above state, $\psi_{n_x,n_y,n_z}(x,y,z)$, is

$$\begin{aligned}
\langle x\rangle_{n_x,n_y,n_z} &= \int_0^a\int_0^b\int_0^c \psi^*_{n_x,n_y,n_z}(x,y,z)\, x\psi_{n_x,n_y,n_z}(x,y,z)\; dxdydz \\
&= \frac{2}{a}\int_0^a x\sin^2\left(\frac{n_x\pi x}{a}\right) dx\; \frac{2}{b}\int_0^b \sin^2\left(\frac{n_y\pi y}{b}\right) dy\; \frac{2}{c}\int_0^c \sin^2\left(\frac{n_z\pi z}{c}\right) dz \\
&= \frac{a}{2}\times 1\times 1 = \frac{a}{2}. \qquad \text{see Eq. 3.9}
\end{aligned}$$

Because the integrand in the first line is a product of factors - each depending on only one of the three variables - the integral factors into three one dimensional integrals. The three factors are particle-in-a-one-dimensional-box expectation values. Since this is the expectation value of x, the y and z integrals are both just the expectation value of 1 - i.e., they are both 1. The expectation value of any operator that does not depend on two of the three variables, or their associated partial derivatives, reduces to a particle-in-a-one-dimensional-box expectation value. In this case, we get $a/2$; i.e., the particle is - on average - in the middle of the interval of possible x values, for all the energy eigenstates.

3.4 Many independent electron systems

So far, we have considered only single particle systems. Particles abound in nature - most notably, electrons. When we treat a multiparticle system, the wavefunction is a function of the coordinates of all the particles. However, if the particle-particle interaction is negligible, or can be approximated by a *mean field*, the energy eigenstates are products of one-particle states. In this case, the energy eigenvalues are sums of one-particle eigenvalues. When the particles are electrons, the energy eigenstates must satisfy the *Pauli principle* - see Sec. XX. This means that only two electrons - one with spin up, the other with spin down - can occupy each one-particle state. This limits the number of energy levels of an n electron system. These points are illustrated by considering many electrons in one and three dimensional boxes, and in a one dimensional well.

3.4.1 Many electrons in a one dimensional box

Suppose there are many electrons in a one dimensional box, large enough so that the electrons can be treated as independent via a *mean field* approximation. The mean field is already incorporated into the model potential - in this case, the one dimensional box potential. The states of the many electron system can be viewed in terms of spin up and spin down electrons occupying the one-particle states, $\psi_n(x)$, found above (Eq. 3.8). These one-particle states are filled, two electrons each, starting with the lowest energy level, then moving to the next lowest energy level - until all there are no more electrons. For example, four electrons in a box produces the ground state energy,

$$\begin{aligned} E_{\text{g.s.}} &= 2E_1 + 2E_2 \\ &= \left(2 + 2 \times 2^2\right) E_1 \\ &= 10E_1 \end{aligned}$$

because the two lowest energy one-particle states accomodate the four electrons. The first excited state energy for this system corresponds to one of the $n = 2$ electrons excited to the $n = 3$ level, and its energy is given by

$$\begin{aligned} E_{\text{1st e.s.}} &= 2E_1 + E_2 + E_3 \\ &= \left(2 + 2^2 + 3^2\right) E_1 \\ &= 15E_1 . \end{aligned}$$

The transition energy, $5E_1$, is just the difference in energy between the $n = 3$ and $n = 2$ levels. It is an even to odd transition that is *dipole-allowed*.

The next three excited states of the four electron system have energy,

$$\begin{aligned} E_{\text{2nd e.s.}} &= E_1 + 2E_2 + E_3 \\ &= \left(1 + 2 \times 2^2 + 3^2\right) E_1 \\ &= 18E_1 , \end{aligned}$$

$$\begin{aligned} E_{\text{3rd e.s.}} &= 2E_1 + 2E_3 \\ &= \left(2 + 2 \times 3^2\right) E_1 \\ &= 20E_1 \end{aligned}$$

and

$$\begin{aligned} E_{\text{4th e.s.}} &= 2E_1 + E_2 + E_4 \\ &= \left(2 + 2^2 + 4^2\right) E_1 \\ &= 22E_1 . \end{aligned}$$

However, transitions to these levels are *dipole-forbidden*. The first of these requires a 1 to 3 transition - an odd to odd - while the last requires a 2 to 4 transition - an even to even. The third excited state is also dipole-forbidden - it requires two electrons to make transitions simultaneously. This is dipole-forbidden for independent electron systems.

3.4.2 Many electrons in a three dimensional box

Now, we let the box be three dimensional - in particular a cubic box; $a = b = c$. In this case, the first three one-particle energy levels are

$$\begin{aligned} E_1 &= E_{1,1,1} = \left(1^2 + 1^2 + 1^2\right) \epsilon_1 \\ &= 3\epsilon_1 , \end{aligned}$$

$$\begin{aligned} E_2 &= E_{1,1,2} = \left(1^2+1^2+2^2\right)\epsilon_1 \\ &= 6\epsilon_1 \end{aligned}$$

and

$$\begin{aligned} E_3 &= E_{1,2,2} = \left(1^2+2^2+2^2\right)\epsilon_1 \\ &= 9\epsilon_1 \end{aligned}$$

where

$$\epsilon_1 = \frac{\hbar^2\pi^2}{2ma^2}.$$

The second and third energy levels are triply degenerate. The three distinct permutations of the indices 1,1,2 and 1,2,2 produce the same energy. Thus, the first two energy levels can accomodate eight electrons. The ground and first excited state energies of eight electron in this box are

$$\begin{aligned} E_{\text{g.s.}} &= 2E_1 + 6E_2 \\ &= (2\times 3 + 6\times 6)\,\epsilon_1 \\ &= 42\epsilon_1 \end{aligned}$$

and

$$\begin{aligned} E_{\text{1st e.s.}} &= 2E_1 + 5E_2 + E_3 \\ &= (2\times 3 + 5\times 6 + 9)\,E_1 \\ &= 45\epsilon_1. \end{aligned}$$

Transitions to the first excited state are dipole-allowed, involving an $n_x = 1$ to 2 (or, for n_y or n_z) transition of one electron from the E_2 to the E_3 level. The energy difference is $3\epsilon_1$.

3.4.3 Many electrons in a one dimensional well

We return to electrons in one dimension - this time in a well with finite depth. The finite well depth allows the system to model ionization processes - the electrons can escape. We determine the ionization energy and electron affinity for a system of four independent electrons in the well depicted in Figure 3.6. These electrons occupy the bottom two one-electron levels - two electrons in each. The ground state energy is

$$\begin{aligned} E_{\text{g.s.}} &= 2E_1 + 2E_2 \\ &= 2\left(\eta_1^2+\eta_2^2\right)\epsilon_1 \\ &= 2\left(1^2+2^2\right)\left(\frac{\eta_{\text{d}}}{\eta_{\text{d}}+1/\pi}\right)^2\epsilon_1 \\ &= 10E_1 \end{aligned}$$

Suppose the four electron system is neutral - i.e., the well has a net charge of $+4e$ that binds the electrons. The (first) ionization energy, I_1, is the difference in energy between the cation ground state plus one electron at the ionization threshold, and the ground state of the neutral system. Here, the ionization threshold is $V_\mathrm{d} = \eta_\mathrm{d}^2\ \epsilon_1$ ($= 15\epsilon_1$ for the well depicted in Fig. 3.6).

$$\begin{aligned}
I_1 &= E_{\text{cation g.s.}} + V_\mathrm{d} - E_{\text{g.s.}} \\
&= 2E_1 + E_2 + V_\mathrm{d} - 10E_1 \\
&= 6E_1 + V_\mathrm{d} - 10E_1 \\
&= \eta_\mathrm{d}^2\epsilon_1 - 4\left(\frac{\eta_\mathrm{d}}{\eta_\mathrm{d} + 1/\pi}\right)^2 \epsilon_1 \\
&= \left(15 - 4\frac{15}{\left(\sqrt{15} + 1/\pi\right)^2}\right)\epsilon_1 \\
&= 11.584\epsilon_1
\end{aligned}$$

The (first) electron affinity, A_1, is the difference in energy between the anion ground state, and the neutral system plus one electron at the ionization threshold.

$$\begin{aligned}
A_1 &= E_{\text{anion g.s.}} - (E_{\text{g.s.}} + V_\mathrm{d}) \\
&= 2E_1 + 2E_2 + E_3 - 10E_1 - V_\mathrm{d} \\
&= 19E_1 - 10E_1 - V_\mathrm{d} \\
&= 9\left(\frac{\eta_\mathrm{d}}{\eta_\mathrm{d} + 1/\pi}\right)^2 \epsilon_1 - \eta_\mathrm{d}^2\epsilon_1 \\
&= \left(9\frac{15}{\left(\sqrt{15} + 1/\pi\right)^2} - 15\right)\epsilon_1 \\
&= -7.3151\epsilon_1
\end{aligned}$$

The electron affinity is negative - the third one-electron level is below the ionization threshold. If the neutral system had eight electrons, than it would have a positive electron affinity. There is no fifth one-electron level for the well under consideration. However, there are resonance states - as seen in Sec. 3.2.3. These resonance states decay rapidly, as they connect to a unit intensity outgoing beam. However, if there were a barrier - as in Sec. 3.2.4 - then the lowest energy resonance state could decay slowly, and a positive electron affinity can be defined. This is analogous to the electron affinities of alkaline earth metals, and nitrogen. The positive electron affinity results because, for these atoms, the anion ground state is actually a resonance state which *autoionizes* with a certain rate.

Many electrons in boxes or wells provide models for *quantum dots* - nanocrystals of a semiconductor. While the electrons are delocalized within the quantum dot, they are confined to the dot. The quantum dot can be a colloidal particle, or a nano-sized island etched into the surface of an oxide coated semiconductor. Quantum dots can also exist in semiconductor heterostructures - layered

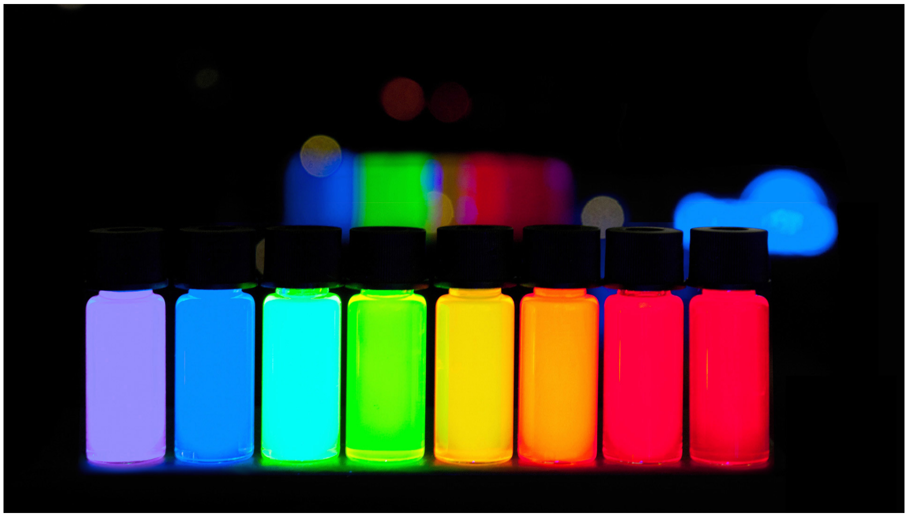


Figure 3.9: Quantum dots, in aqueous suspension, emitting light. Frequency varies continuously with the size of the dots - increasing from left to right. [From: http://en.wikipedia.org/wiki/Quantum_dot#/media/File:Quantum_Dots_with_emission_maxima_in_a_10-nm_step_are_being_produced_at_PlasmaChem_in_a_kg_scale.jpg. Retrieved May 8, 2015]

materials that can produce confinement in the direction of layering. Quantum dots are so small that hold a limited number of electrons, with well separated energy levels. Transition frequencies of quantum dots can be finely tuned by adjusting the size of the dot - L or a in the models considered above. The principle transition frequency in the above models is inversely proportional to L^2 (or a^2) - recall, $\epsilon_1 = \hbar^2\pi^2/\left(2mL^2\right)$. Figure 3.9 shows ZnCdSeS colloidal quantum dots in a liquid suspension. The quantum dots are larger in each successive suspension - from left to right. Since these quantum dots are emitting light (they have been excited), we see the color shift from violet for the smallest quantum dots - with the largest energy spacings - to red for the largest quantum dots - with the smallest energy spacings.

Example 3.3. Consider a particle in the n th energy eigenstate of a one dimensional well. Approximate energy eigenvalues and eigenfunctions are given in Eqs. 3.17 and 3.23 , respectively. Note that the normalization constant, $\sqrt{2/L}$, in only an approximation. For probability calculations, this constant must be corrected.

(a) Determine the normalization contant, C, given the form of the n th energy eigenstate,

$$\psi_n(x) \cong \begin{cases} C \sin\left(\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}}+1/\pi}\frac{n\pi x}{L}\right), & 0 < x < L \\ C\frac{(-1)^{n+1}n}{\eta_{\mathrm{d}}+1/\pi}\exp\left(-\left(\eta_{\mathrm{d}}^2 - n^2\left(\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}}+1/\pi}\right)^2\right)^{1/2}\frac{\pi(x-L)}{L}\right), & L < x \end{cases}.$$

(b) Determine a formula (depending on n, L and η_{d}) for the probability that a measurement of the position of the particle will yield a value outside the well.

(c) Evaluate the probability determined in part b, for the $n = 1$ and 4 states of the particle in a well depicted in Fig. 3.6 - $L = 1$, $V_{\mathrm{d}} = 15\epsilon_1$.

(d) How do you expect the expectation of x for a particle in the well to compare with the expectation value of x in the corresponding particle in a box state - i.e., $L/2$? Which state - the ground state, $n = 1$, or the top state, $n = 4$ - has the largest x expectation value?

Solution 3.3. (a) The normalization condition takes the form,

$$\begin{aligned} 1 &= \int_0^\infty |\psi_n|^2\,dx \\ &= \int_0^L |\psi_n|^2\,dx + \int_L^\infty |\psi_n|^2\,dx \\ &= |C|^2\left(\int_0^L \sin^2\left(\frac{\eta_n\pi x}{L}\right)dx + \left(\frac{\eta_n}{\eta_{\mathrm{d}}}\right)^2\int_L^\infty \exp\left(-2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi(x-L)}{L}\right)dx\right) \end{aligned}$$

The first integral in the second line is determined as follows:

$$\begin{aligned} &\int_0^L \sin^2\left(\frac{\eta_n\pi x}{L}\right)dx \\ =\ & \frac{1}{2}\int_0^L\left(1-\cos\left(\frac{2\eta_n\pi x}{L}\right)\right)dx \\ =\ & \frac{1}{2}\left(L - \int_0^L \cos\left(\frac{2\eta_n\pi x}{L}\right)dx\right) \\ =\ & \frac{1}{2}\left(L - \frac{L}{2\eta_n\pi}\left[\sin\left(\frac{2\eta_n\pi x}{L}\right)\right]_0^L\right) \\ =\ & \frac{L}{2}\left(1-\frac{[\sin(2\eta_n\pi)-0]}{2\eta_n\pi}\right) \\ =\ & \frac{L}{2}\left(1-\frac{\sin(2\eta_n\pi)}{2\eta_n\pi}\right). \end{aligned}$$

The second integral gives

$$\begin{aligned}
&\int_L^\infty \exp\left(-2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi(x-L)}{L}\right)\\
= \quad & \frac{-L}{2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\left[\exp\left(-2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi(x-L)}{L}\right)\right]_L^\infty\\
= \quad & \frac{L}{2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}.
\end{aligned}$$

Altogether, we have

$$\begin{aligned}
1 &= \int_0^\infty |\psi_n|^2\,dx\\
&= |C|^2\left(\frac{L}{2}\left(1-\frac{\sin(2\eta_n\pi)}{2\eta_n\pi}\right)+\left(\frac{\eta_n}{\eta_{\mathrm{d}}}\right)^2\frac{L}{\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\right)\\
&= \frac{L}{2}|C|^2\left(1-\frac{\sin(2\eta_n\pi)}{2\eta_n\pi}+\frac{\eta_n^2}{\eta_{\mathrm{d}}^2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\right),
\end{aligned}$$

or

$$C=\left(\frac{2}{L}\right)^{1/2}\left(1-\frac{\sin(2\eta_n\pi)}{2\eta_n\pi}+\frac{\eta_n^2}{\eta_{\mathrm{d}}^2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\right)^{-1/2}.$$

(b) The probability that the particle is outside the well is given by

$$\begin{aligned}
P_{\text{outside}} &= \int_L^\infty |\psi_n|^2\,dx\\
&= C^2\left(\frac{\eta_n}{\eta_{\mathrm{d}}}\right)^2\int_L^\infty \exp\left(-2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\frac{\pi(x-L)}{L}\right)dx\\
&= \frac{L}{2}C^2\frac{\eta_n^2}{\eta_{\mathrm{d}}^2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi} \qquad \text{from part a}\\
&= \frac{\eta_n^2}{\eta_{\mathrm{d}}^2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\left(1-\frac{\sin(2\eta_n\pi)}{2\eta_n\pi}+\frac{\eta_n^2}{\eta_{\mathrm{d}}^2\left(\eta_{\mathrm{d}}^2-\eta_n^2\right)^{1/2}\pi}\right)^{-1}.
\end{aligned}$$

(c) For the well with $L=1$ and $V_{\mathrm{d}}=15\epsilon_1$, $\eta_{\mathrm{d}}^2=15$, and

$$\frac{\eta_{\mathrm{d}}}{\eta_{\mathrm{d}}+1/\pi}=\frac{\sqrt{15}}{\sqrt{15}+1/\pi}=0.92405.$$

In this case,

$$\eta_1=0.92405$$

and

$$\begin{aligned}\eta_4 &= 4 \times 0.92405\\ &= 3.6962.\end{aligned}$$

For the ground state, $n = 1$,

$$\begin{aligned}& P_{\text{outside}}\\ = \; & \frac{0.92405^2}{15\left(15 - 0.92405^2\right)^{1/2}\pi}\left(1 - \frac{\sin\left(2 \times 0.92405\pi\right)}{2 \times 0.92405\pi} + \frac{0.92405^2}{15\left(15 - 0.92405^2\right)^{1/2}\pi}\right)^{-1}\\ = \; & \frac{4.82 \times 10^{-3}}{1.0791 + 4.82 \times 10^{-3}} = 4.45 \times 10^{-3}.\end{aligned}$$

For the top state, $n = 4$,

$$\begin{aligned}& P_{\text{outside}}\\ = \; & \frac{16 \times 0.92405^2}{15\left(15 - 16 \times 0.92405^2\right)^{1/2}\pi}\left(1 - \frac{\sin\left(8 \times 0.92405\pi\right)}{8 \times 0.92405\pi} + \frac{16 \times 0.92405^2}{15\left(15 - 16 \times 0.92405^2\right)^{1/2}\pi}\right)^{-1}\\ = \; & \frac{0.25063}{1.0406 + 0.25063} = 0.1941.\end{aligned}$$

The particle is much more likely to be found outside the well when in the top state, rather than the ground state.

(d) The expectation of position for a particle in the n th state is

$$\begin{aligned}& \int_0^\infty x\left|\psi_n\right|^2 dx\\ = \; & \int_0^L x\psi_n^2 dx + \int_L^\infty x\psi_n^2 dx\\ = \; & C^2\left(\int_0^L x\sin^2\left(\frac{\eta_n\pi x}{L}\right)dx + \left(\frac{\eta_n}{\eta_{\mathrm{d}}}\right)^2\int_L^\infty x\exp\left(-2\left(\eta_{\mathrm{d}}^2 - \eta_n^2\right)^{1/2}\frac{\pi\left(x - L\right)}{L}\right)dx\right).\end{aligned}$$

The first integral is like the expectation value of a particle in a box, except that the sine function does not quite reach zero at the right boundary - the miss is more prounounced for larger n values. This means the position probability density is shifted a higher values - in comparison with the particle in a box state. The associated mean - the x expectation - is larger for this distribution. The second integral is an exponential integral that can be evaluated using integration by parts. Here, we note that it shifts the x expectation to higher values. The shift increases with energy, as the prefactor increases and the exponential decay is slower. Overall, the x expectation associated with a particle in a well energy eigenstates increases with energy, and is always

Example 3.4. Consider a three dimensional box with sides, $a = 2b$ and $c = b$. Suppose $\hbar^2\pi^2/\left(2mb^2\right) = 1.00$ eV, where m is electron mass.

(a) Determine the lowest five energy levels and the associated degeneracies for an electron in this box.

(b) What are the ground and first excited state energies of ten independent electrons in the box?

(c) Suppose now that there are eight independent electrons in the box. What are the two lowest dipole-allowed absorption frequencies?

Solution 3.4.

(a) The energy levels have given by

$$\begin{aligned} E_{n_x,n_y,n_z} &= n_x^2\frac{\hbar^2\pi^2}{2ma^2} + n_y^2\frac{\hbar^2\pi^2}{2mb^2} + n_z^2\frac{\hbar^2\pi^2}{2mb^2} \\ &= \frac{n_x^2}{4} + n_y^2 + n_z^2 \text{ eV}. \end{aligned}$$

The ground state energy,

$$E_{1,1,1} = 2.25 \text{ eV},$$

is non-degenerate. The next three energy levels are

$$E_{2,1,1} = 3.00 \text{ eV},$$

$$E_{3,1,1} = 4.25 \text{ eV},$$

$$E_{1,2,1} = E_{1,1,2} = 5.25 \text{ eV}$$

and

$$E_{4,1,1} = 6.00 \text{ eV}.$$

Only the second highest of these, 5.25 eV, is degenerate - it is doubly degenerate.

(b) Ten electrons would completely fill the first four energy levels found in part a. The associated ground state energy is

$$\begin{aligned} E_{\text{g.s.}} &= 2\times E_{1,1,1} + 2\times E_{2,1,1} + 2\times E_{3,1,1} + 4\times E_{1,2,1} \\ &= 40.00 \text{ eV}. \end{aligned}$$

The first excited state energy is

$$\begin{aligned} E_{\text{1st e.s.}} &= 2\times E_{1,1,1} + 2\times E_{2,1,1} + 2\times E_{3,1,1} + 3\times E_{1,2,1} + E_{4,1,1} \\ &= 40.00 \; - 5.25 + 6.00 \text{ eV} \\ &= 40.75 \text{ eV}. \end{aligned}$$

This excited state here corresponds to an excitation of an electron from $1, 2, 1$ (or $1, 1, 2$) to $3, 1, 1$.

(c) If there are eight electrons in the box, then the ground state energy would be

$$\begin{aligned} E_{\mathrm{g.s.}} &= 2\times E_{1,1,1}+2\times E_{2,1,1}+2\times E_{3,1,1}+2\times E_{1,2,1} \\ &= 29.50 \text{ eV}. \end{aligned}$$

Excited states that are dipole-allowed involve the transition of one electron. It must also involve the change of only one quantum number. The transition integral is zero if more than one more quantum number changes. This results because the three dimensional transition integral factors into three one dimensional integrals. In three dimensions, the dipole moment is linear in x, y and z. The x term connects (i.e., produces a non-zero integral) states differing in only n_x. If the n_y or n_z quantum numbers are different in the two states, the y or z integral is zero - the distinct states of a particle in a one dimension are orthogonal. Similarly, the y and z tems connect states differing in n_y and n_z, respectively - while the other quantum numbers do not change. Thus, we look for allowed changes in a quantum number. For the particle in a box, the lowest frequency dipole-allowed transitions are for $\Delta n = 1$ and 3. The energy spacings are smallest for the x degree of freedom. However, the lowest energy transition available is $n_x = 3$ to 4; i.e., $3,1,1$ to $4,1,1$. This transition produces the higher energy excited state,

$$\begin{aligned} E_{\mathrm{e.s.}} &= 2\times E_{1,1,1}+2\times E_{2,1,1}+E_{3,1,1}+2\times E_{1,2,1}+E_{4,1,1} \\ &= 29.50-4.25+6.00 \text{ eV} \\ &= 31.25 \text{ eV}. \end{aligned}$$

Here, $\Delta E = 6.00 - 4.25$ eV $= 1.75$ eV. The $3,1,1$ to $6,1,1$ is much higher in energy, $\Delta E = 9.00 - 4.25$ eV $= 4.75$ eV. Because there is still room for two more electrons at the $1,2,1$ energy level, the lowest energy $\Delta n = 1$ transition for n_y (or n_z), is $1,1,1$ to $1,2,1$, with energy, $\Delta E = 4.00 - 1.00$ eV $= 3.00$ eV. This is the lowest energy dipole-allowed transition. It results in the excited state with energy,

$$\begin{aligned} E_{\mathrm{e.s.}} &= E_{1,1,1}+2\times E_{2,1,1}+2\times E_{3,1,1}+3\times E_{1,2,1} \\ &= 29.50-1.00+4.00 \text{ eV} \\ &= 32.50 \text{ eV}. \\ &= 30.50 \text{ eV}. \end{aligned}$$

The two lowest allowed transition frequencies are given by $\nu = \Delta E/h$.

Chapter 4

Vibrations of a diatomic molecule

Here, we jump to the *diatomic molecule* - two atoms bound together by shared electrons in covalent bond(s). Diatomic molecules vibrate and rotate, while their enter of mass travels freely. To treat these motions using quantum mechanics, we must treat all the particles - all the electrons, and the two nuclei. However, because electrons are so much lighter than nuclei, they move much faster. The diatomic molecule is well-modeled by first solving the Schrödinger equation for the electrons in the diatomic molecule - see Sec. XX - for all distances between the nuclei. The resulting electronic ground state energy, as a function of distance between the nuclei, becomes the potential energy that binds the nuclei together. We suppose that the electronic portion of the Schrödinger equation has been solved, and the Schrödinger equation for two nuclei has a potential energy that depends only on the distance between the nuclei. Since we have two particles in three dimensional space, there are six degrees of freedom in total - three coordinates for each nucleus. The TISE is a partial differential equation in six coordinates. However, it is subject to *separation of variables.*

4.1 Relative and center of mass motion

The Hamiltonian has the form,

$$\hat{H} = \frac{1}{2m_1}\hat{\mathbf{p}}_1^2 + \frac{1}{2m_2}\hat{\mathbf{p}}_2^2 + V\left(r\right),$$

where $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ are the momentum operators (each a three component vector) for nucleus 1 and nucleus 2. These momentum operators are conjugate to the position vectors, $\mathbf{x}_1$ and $\mathbf{x}_2$ of the two nuclei;

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

is the gradient with respect to the coordinates of the j th atom (nucleus). Since the potential depends only on

$$r = \|\mathbf{x}_2 - \mathbf{x}_1\|,$$

we change variables to *relative* and *center of mass* variables,

$$\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$$

and

$$\mathbf{X} = \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2}{M},$$

respectively. $M = m_1 + m_2$ is the total mass. To transform the momentum operators, we must express $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ is terms of

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

and

$$\hat{\mathbf{P}} = -i\hbar\frac{\partial}{\partial \mathbf{X}},$$

the *momenta conjugate* to the relative and center of mass coordinates, respectively. Using the chain rule,

$$\begin{aligned}
\hat{\mathbf{p}}_1 &= -i\hbar\frac{\partial}{\partial \mathbf{x}_1} \\
&= -i\hbar\left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_1}\frac{\partial}{\partial \mathbf{x}} + \frac{\partial \mathbf{X}}{\partial \mathbf{x}_1}\frac{\partial}{\partial \mathbf{X}}\right) \\
&= -i\hbar\left(-\frac{\partial}{\partial \mathbf{x}} + \frac{m_1}{M}\frac{\partial}{\partial \mathbf{X}}\right) \\
&= -\hat{\mathbf{p}} + \frac{m_1}{M}\hat{\mathbf{P}}
\end{aligned}$$

and

$$\begin{aligned}
\hat{\mathbf{p}}_2 &= -i\hbar\frac{\partial}{\partial \mathbf{x}_2} \\
&= -i\hbar\left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_2}\frac{\partial}{\partial \mathbf{x}} + \frac{\partial \mathbf{X}}{\partial \mathbf{x}_2}\frac{\partial}{\partial \mathbf{X}}\right) \\
&= -i\hbar\left(\frac{\partial}{\partial \mathbf{x}} + \frac{m_2}{M}\frac{\partial}{\partial \mathbf{X}}\right) \\
&= \hat{\mathbf{p}} + \frac{m_2}{M}\hat{\mathbf{P}}.
\end{aligned}$$

Here, the derivatives, $\partial\mathbf{x}/\partial\mathbf{x}_1$, $\partial\mathbf{x}/\partial\mathbf{x}_2$, $\partial\mathbf{X}/\partial\mathbf{x}_1$ and $\partial\mathbf{X}/\partial\mathbf{x}_2$ are Jacobi matrices, and terms like $\frac{\partial \mathbf{x}}{\partial \mathbf{x}_2}\frac{\partial}{\partial \mathbf{x}}$ are matrix vector products.

Substituting the above expressions into the Hamiltonian expression gives

$$\begin{aligned}
\hat{H} &= \frac{1}{2m_1}\left(-\hat{\mathbf{p}}+\frac{m_1}{M}\hat{\mathbf{P}}\right)^2+\frac{1}{2m_2}\left(\hat{\mathbf{p}}+\frac{m_2}{M}\hat{\mathbf{P}}\right)^2+V(r) \\
&= \frac{1}{2m_1}\left(\hat{\mathbf{p}}^2-\frac{2m_1}{M}\hat{\mathbf{p}}^T\hat{\mathbf{P}}+\frac{m_1^2}{M^2}\hat{\mathbf{P}}^2\right)+\frac{1}{2m_2}\left(\hat{\mathbf{p}}^2+\frac{2m_2}{M}\hat{\mathbf{p}}^T\hat{\mathbf{P}}+\frac{m_2^2}{M^2}\hat{\mathbf{P}}^2\right)+V(r) \\
&= \frac{1}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\hat{\mathbf{p}}^2+\frac{(m_1+m_2)}{2M^2}\hat{\mathbf{P}}^2+V(r) \\
&= \frac{1}{2m}\hat{\mathbf{p}}^2+\frac{1}{2M}\hat{\mathbf{P}}^2+V(\|\mathbf{x}\|),
\end{aligned}$$

where

$$m=\frac{m_1m_2}{M}$$

is called the reduced mass. Since the Hamiltonian is the sum of two operators - one that acts only affects functions of $\mathbf{x}$, and another that affects only functions of $\mathbf{X}$ - the TISE separates into two independent equations - one for the relative configuration of atoms, and another for the center of mass. Specifically, let

$$\psi(\mathbf{x},\mathbf{X})=\psi_{\text{rel}}(\mathbf{x})\,\psi_{\text{cen}}(\mathbf{X})$$

and note that the TISE separates into

$$\left(\frac{1}{2m}\hat{\mathbf{p}}^2+V(\|\mathbf{x}\|)\right)\psi_{\text{rel}}(\mathbf{x})=E_{\text{rel}}\psi_{\text{rel}}(\mathbf{x}) \tag{4.1}$$

and

$$\frac{1}{2M}\hat{\mathbf{P}}^2\psi_{\text{cen}}(\mathbf{X})=E_{\text{cen}}\psi_{\text{cen}}(\mathbf{X}). \tag{4.2}$$

$$E=E_{\text{cen}}+E_{\text{rel}}.$$

4.1.1 Center of mass motion

The center of mass TISE - Eq. 4.2 above - in turn separates into three independent TISE's - one for each dimension of space;

$$\psi_{\text{cen}}(\mathbf{X})=\psi_{\text{cen},\,X}(X)\,\psi_{\text{cen},\,Y}(Y)\,\psi_{\text{cen},\,Z}(Z),$$

$$\frac{1}{2M}\hat{P}_X^2\psi_{\text{cen},\,X}(X)=E_{\text{cen},\,X}\psi_{\text{cen},\,X}(X),\qquad\text{etc.} \tag{4.3}$$

and

$$E_{\text{cen}}=E_{\text{cen},\,X}+E_{\text{cen},\,Y}+E_{\text{cen},\,Z}.$$

The energy eigenvalues are *not quantized*, as the center of mass motion is *unbound*. The center of mass motion is the motion of a free particle in three dimensions - there is no potential energy. For each coordinate, there are two states - plane waves corresponding to forward and backward motion - for each energy eigenvalue, $E_{\text{cen},\,X,\,Y\text{ or }Z}>0$. The only boundary condition here is the

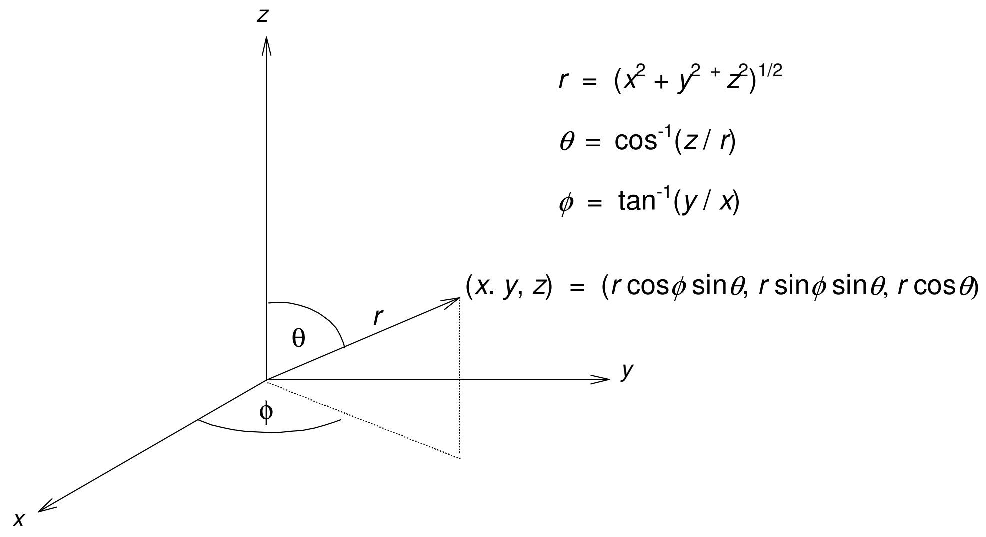


Figure 4.1: Cartesian coordinates in terms of spherical coordinates - and vice versa.

requirement that the wavefunction not blow up (i.e., diverge to ∞) for all X (Y or Z) from $-\infty$ to $+\infty$. For energy less than zero one solution blows up at $-\infty$ while the other blows up at $+\infty$. Since both of these solutions are inadmissable, negative energies are not eigenvalues of the Hamiltonian. For energy greater than or equal to zero, the *plane wave solutions* are

$$\psi_{\text{cen}}(\mathbf{X}) = \exp\left(i\frac{\sqrt{2ME_{\text{cen, }X}}}{\hbar}X\right)\exp\left(i\frac{\sqrt{2ME_{\text{cen, }Y}}}{\hbar}Y\right)\exp\left(i\frac{\sqrt{2ME_{\text{cen, }Z}}}{\hbar}Z\right).$$

4.2 Relative motion - angular and radial

The relative TISE - Eq. 4.1 above - also further separates. Here, we must change variables - this time from Cartesian (x, y and z) to *spherical coordinates*, (r, θ and ϕ) defined in Figure 4.1.

This transformation requires many steps - see the Appendix, Sec. 4.6. The key results are Eqs. 4.13 and 4.14 which express the relative kinetic energy in spherical coordinates. Since the potential depends only on r, the relative Hamiltonian is now expressed as the sum of angular and radial Hamiltonians. The only trouble here is that the angular Hamiltonian depends on r (it has a $1/r^2$ prefactor). Nevertheless, the relative energy eigenfunctions are still poducts of radial and angular parts,

$$\psi(r,\theta,\phi) = \psi_r(r)\,Y(\theta,\phi).$$

Substituting this form into the TISE for the relative motion gives

$$\left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+V\left(r\right)+\frac{1}{2mr^2}\hat{\mathbf{L}}^2\right)\psi_r\left(r\right)Y\left(\theta,\phi\right)=E\psi_r\left(r\right)Y\left(\theta,\phi\right).$$

If $Y\left(\theta,\phi\right)$ is an eigenfunction of $\hat{\mathbf{L}}^2$, i.e.,

$$\hat{\mathbf{L}}^2Y\left(\theta,\phi\right)=L^2Y\left(\theta,\phi\right),$$

then we can divide the above equation by $Y\left(\theta,\phi\right)$ to get

$$\begin{aligned}E\psi_r\left(r\right) &= \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+V\left(r\right)\right]\psi_r\left(r\right)+\psi_r\left(r\right)\frac{1}{2mr^2}\frac{\hat{\mathbf{L}}^2Y\left(\theta,\phi\right)}{Y\left(\theta,\phi\right)}\\ &= \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+V\left(r\right)\right]\psi_r\left(r\right)+\psi_r\left(r\right)\frac{1}{2mr^2}\frac{L^2Y\left(\theta,\phi\right)}{Y\left(\theta,\phi\right)}\\ &= \left[-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}\right)+V\left(r\right)+\frac{L^2}{2mr^2}\right]\psi_r\left(r\right).\end{aligned}$$

This is almost the TISE for a particle in one dimension - the kinetic energy has an extra term. The potential energy consists of the potential, $V\left(r\right)$, which binds the atoms, and an effective potential - the *centrifugal potential* energy - which pulls the atoms apart. The centrifugal potential is proportional to angular momentum squared. It it inversely proportional to r^2. The centrifugal force - minus the derivative of the centrifugal potential - drops off as $1/r^3$. It is largest when the rotating atoms are closest together (c.f., a figure skater pulling their arms and leg in to spin faster).

The kinetic energy portion of the above radial TISE can be simplified by making the substitution,

$$\psi_r\left(r\right)=\frac{1}{r}\varphi\left(r\right),$$

to get

$$\left[-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}\right)+V\left(r\right)+\frac{L^2}{2mr^2}\right]\frac{1}{r}\varphi\left(r\right)=E\frac{1}{r}\varphi\left(r\right)$$

or

$$\begin{aligned}
E\varphi(r) &= -\frac{\hbar^2 r}{2m}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}\right)\frac{1}{r}\varphi(r)+\left(V(r)+\frac{L^2}{2mr^2}\right)\varphi(r)\\
&= -\frac{\hbar^2 r}{2m}\left(\frac{d}{dr}+\frac{2}{r}\right)\frac{d}{dr}\frac{1}{r}\varphi(r)+\left(V(r)+\frac{L^2}{2mr^2}\right)\varphi(r)\\
&= -\frac{\hbar^2 r}{2m}\left(\frac{d}{dr}+\frac{2}{r}\right)\left(-\frac{1}{r^2}\varphi(r)+\frac{1}{r}\frac{d}{dr}\varphi(r)\right)+\left(V(r)+\frac{L^2}{2mr^2}\right)\varphi(r)\\
&= -\frac{\hbar^2 r}{2m}\left(\frac{2}{r^3}\varphi(r)-\frac{2}{r^2}\frac{d}{dr}\varphi(r)+\frac{1}{r}\frac{d^2}{dr^2}\varphi(r)-\frac{2}{r^3}\varphi(r)+\frac{2}{r^2}\frac{d}{dr}\varphi(r)\right)\\
&\quad+\left(V(r)+\frac{L^2}{2mr^2}\right)\varphi(r)\\
&= -\frac{\hbar^2}{2m}\frac{d^2}{dr^2}\varphi(r)+\left(V(r)+\frac{L^2}{2mr^2}\right)\varphi(r)\\
&= \left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2}+V(r)+\frac{L^2}{2mr^2}\right)\varphi(r),
\end{aligned}$$

the *TISE for a mass, m, particle in one dimension* subject to the potential, $V(r)+L^2/\left(2mr^2\right)$.

Example 4.1. Consider the following Hamiltonians:

$$\hat{H}_1=\frac{\hat{p}_x^2}{2}+\frac{\hat{p}_y^2}{2},$$

$$\hat{H}_2=\frac{\hat{p}_x^2}{2}+\frac{\hat{p}_y^2}{2x}$$

and

$$\hat{H}_3=\frac{\hat{p}_x^2}{2}+\frac{\hat{p}_y^2}{2}+xy.$$

(a) For which of these Hamiltonians can the eigenfunctions can be written as a product of a function of x and a function of y?

(b) For those that can, determine the eigenvalue equations satisfied by the functions of x and y. What is the relationship between the associated eigenvalues, and the eigenvalues of the Hamiltonian?

Solution 4.1.

(a) The first Hamiltonian is separable. It is a sum of operators that each only affect functions of one variable. Its eigenfunctions are products of functions of just x and just y. The second Hamiltonian has a $1/x$ factor in front of the operator that, otherwise, acts on functions of just y. Nevertheless, the eigenfunctions are still products of functions of just x and just y. The third Hamiltonian has a potential energy term that couples x and y. The eigenfunctions of this Hamiltonian are not products of one variable wavefunctions.

(b) For the first Hamiltonian, the eigenfunctions have the form

$$\psi(x,y) = \psi_x(x)\,\psi_y(y),$$

where

$$\frac{\hat{p}_x^2}{2}\psi_x(x) = E_x\psi_x(x),$$

$$\frac{\hat{p}_y^2}{2}\psi_y(y) = E_y\psi_y(y)$$

and the eigenvalues of $\hat{H}_1$ are sums of the $\hat{p}_x^2/2$ and $\hat{p}_y^2/2$ eigenvalues;

$$E = E_x + E_y.$$

For the second Hamiltonian, we still have

$$\psi(x,y) = \psi_x(x)\,\psi_y(y),$$

where

$$\frac{\hat{p}_y^2}{2}\psi_y(y) = E_y\psi_y(y).$$

Substituting this into the $\hat{H}_2$ eigenvalue equation gives

$$\begin{aligned}
\hat{H}_2\psi(x,y) &= E\psi(x,y)\\
\left(\frac{\hat{p}_x^2}{2}\psi_x(x)\right)\psi_y(y) + \frac{1}{x}\psi_x(x)\left(\frac{\hat{p}_y^2}{2}\psi_y(y)\right) &= E\psi_x(x)\,\psi_y(y)\\
\left(\frac{\hat{p}_x^2}{2}\psi_x(x) + \frac{E_y}{x}\psi_x(x)\right)\psi_y(y) &= E\psi_x(x)\,\psi_y(y)\\
\left(\frac{\hat{p}_x^2}{2} + \frac{E_y}{x}\right)\psi_x(x) &= E\psi_x(x)
\end{aligned}$$

The eigenvalues of $\hat{H}_2$ are the eigenvalues of the $\psi_x(x)$ eigenvalue equation. The eigenvalues of $\hat{p}_y^2/2$ appear as a parameter - the coefficient of the $1/x$ effective potential term. There is a set of E eigenvalues for each E_y eigenvalue.

4.3 The quantum harmonic oscillator

The radial relative motion of the diatomic molecule is the *molecular vibration.* Typically, the binding potential, $V(r)$, consists of a narrow well restricting r to a narrow range about the mean bond distance between the two atoms - as long as there is limited vibrational energy. In this case, the centrifugal potential can be approximated as a constant,

$$\frac{L^2}{2mr^2} \cong \frac{L^2}{2mr_0^2} = \frac{L^2}{2I_0},$$

where r_0 is *equilibrium bond distance* (where $V(r)$ is a minimum), and $I_0 = mr_0^2$ is the associated equilibrium *moment of inertia* of the diatomic molecule. Now, the angular kinetic energy (which we investigate later - see Chapter 5) simply adds a constant term to the total relative energy. For now, we leave out this energy and solve for the vibrational energy eigenvalues and associated eigenfunction, $\varphi(r)$. Further, we Taylor-expand $V(r)$ about r_0 and keep only the terms up to quadratic. Here, we set $V(r_0) = 0$, which corresponds to choosing the equilibrium bond distance as the reference configuration for the vibrational energy. Also, because $V(r)$ has a minimum at $r = r_0$, the first derivative of $V(r)$ is zero at $r = r_0$. Thus, only the second derivative term remains and the vibrational TISE takes the form,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + \frac{1}{2}V''(r_0)(r - r_0)^2\right)\varphi(r) = E_{\text{vib}}\varphi(r)\,.$$

Now, we make the substitution, $x = r - r_0$, which gives the TISE for the *harmonic oscillator*,

$$\hat{H}_{\text{vib}}\psi_{\text{vib}}(x) = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{k}{2}x^2\right)\psi_{\text{vib}}(x) = E_{\text{vib}}\psi_{\text{vib}}(x)\,, \tag{4.4}$$

where

$$k = V''(r_0)$$

is the *force constant* of the harmonic oscillator, and

$$\psi_{\text{vib}}(x) = \varphi(r_0 + x)$$

is just the wavefunction expressed in terms of displacement from equilibrium. Note that variable x is re-purposed here, as the displacement of bond length from equilbium.

There is yet another simplification. Multiplying Eq. 4.4 by $m^{1/2}/\left(\hbar k^{1/2}\right)$ gives

$$\frac{1}{2}\left(-\frac{\hbar}{(mk)^{1/2}}\frac{d^2}{dx^2} + \frac{(mk)^{1/2}}{\hbar}x^2\right)\psi_{\text{vib}}(x) = \frac{E_{\text{vib}}}{\hbar\omega}\psi_{\text{vib}}(x)$$

or

$$\hat{H}\psi(y) = \frac{1}{2}\left(-\frac{d^2}{dy^2} + y^2\right)\psi(y) = \epsilon\psi(y) \tag{4.5}$$

where

$$y = \frac{x}{\alpha},$$

$$\alpha = \frac{\hbar^{1/2}}{(mk)^{1/4}},$$

$$\epsilon = \frac{E_{\text{vib}}}{\hbar\omega}$$

and

$$\omega = \left(\frac{k}{m}\right)^{1/2}.$$

Here, we re-purpose y and ψ.

Equation 4.5 is the scaled TISE for the harmonic oscillator. It is corresponds to solving Eq. 4.4 with displacement from equilibrium measured in units of α, and energy measured in units of $\hbar\omega$.

4.3.1 Raising and lowering operators

We solve the vibrational TISE using *raising and lowering operators*, $\hat{a}^\dagger$ and $\hat{a}$, respectively:

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(-\frac{d}{dy} + y\right)$$

and

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\frac{d}{dy} + y\right).$$

The idea is to try to factor the vibrational Hamiltonian - it is the sum of two squares. The only twist is that the operators that are squared, d/dy and y, do not commute. Consider the product, $\hat{a}^\dagger\hat{a}$ - expressed here with a *test function*, $\varphi(y)$, to avoid errors in operator evaluation.

$$\begin{aligned}
\hat{a}^\dagger\hat{a}\varphi(y) &= \frac{1}{\sqrt{2}}\left(-\frac{d}{dy} + y\right)\frac{1}{\sqrt{2}}\left(\frac{d}{dy} + y\right)\varphi(y) \\
&= \frac{1}{2}\left(-\frac{d^2}{dy^2} + y\frac{d}{dy} - \frac{d}{dy}y + y^2\right)\varphi(y) \\
&= \frac{1}{2}\left(\left(-\frac{d^2}{dy^2} + y^2\right)\varphi(y) + y\frac{d}{dy}\varphi(y) - \frac{d}{dy}y\varphi(y)\right) \\
&= \frac{1}{2}\left(\left(-\frac{d^2}{dy^2} + y^2\right)\varphi(y) + y\frac{d}{dy}\varphi(y) - y\frac{d}{dy}\varphi(y) - \varphi(y)\right) \\
&= \hat{H}\varphi(y) - \frac{1}{2}\varphi(y).
\end{aligned}$$

Therefore,

$$\hat{a}^\dagger\hat{a} = \hat{H} - \frac{1}{2}$$

or

$$\hat{H} = \hat{a}^\dagger\hat{a} + \frac{1}{2}.$$

To use the raising and lowering operators to solve the TISE associated with $\hat{H}$, we need the commutation relations of $\hat{a}^\dagger$ and $\hat{a}$ with each other and with the Hamiltonian.

First,

$$\begin{aligned}
\left[\hat{a}^{\dagger},\hat{a}\right] &= \left[\frac{1}{\sqrt{2}}\left(-\frac{d}{dy}+y\right),\frac{1}{\sqrt{2}}\left(\frac{d}{dy}+y\right)\right] \\
&= \frac{1}{2}\left[-\frac{d}{dy}+y,\frac{d}{dy}+y\right] \\
&= \frac{1}{2}\left(\left[-\frac{d}{dy},y\right]+\left[y,\frac{d}{dy}\right]\right) \\
&= \left[y,\frac{d}{dy}\right] \\
&= -1.
\end{aligned}$$

Next,

$$\begin{aligned}
\left[\hat{H},\hat{a}\right] &= \left[\hat{a}^{\dagger}\hat{a}+\frac{1}{2},\hat{a}\right] \\
&= \left[\hat{a}^{\dagger}\hat{a},\hat{a}\right] \\
&= \hat{a}^{\dagger}\hat{a}^{2}-\hat{a}\hat{a}^{\dagger}\hat{a} \\
&= \left[\hat{a}^{\dagger},\hat{a}\right]\hat{a} \\
&= -\hat{a}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left[\hat{H},\hat{a}^{\dagger}\right] &= \left[\hat{a}^{\dagger}\hat{a}+\frac{1}{2},\hat{a}^{\dagger}\right] \\
&= \left[\hat{a}^{\dagger}\hat{a},\hat{a}^{\dagger}\right] \\
&= \hat{a}^{\dagger}\hat{a}\hat{a}^{\dagger}-\left(\hat{a}^{\dagger}\right)^{2}\hat{a} \\
&= \hat{a}^{\dagger}\left[\hat{a},\hat{a}^{\dagger}\right] \\
&= -\hat{a}^{\dagger}\left[\hat{a}^{\dagger},\hat{a}\right] \\
&= \hat{a}^{\dagger}.
\end{aligned}$$

The significance of the last two commutators is made clear in the following two arguments: Suppose $\psi(y)$ is an eigenstate of $\hat{H}$ associated with eigenvalue, ϵ; i.e.,

$$\hat{H}\psi(y)=\epsilon\psi(y).$$

Now consider $\hat{H}$ acting on $\hat{a}\psi(y)$:

$$\begin{aligned}
\hat{H}\hat{a}\psi(y) &= \left[\hat{H},\hat{a}\right]\psi(y)+\hat{a}\hat{H}\psi(y) \\
&= -\hat{a}\psi(y)+\hat{a}\epsilon\psi(y) \\
&= (\epsilon-1)\,\hat{a}\psi(y).
\end{aligned}$$

Thus, $\hat{a}\psi(x)$ is also an eigenstate of the Hamiltonian, except that it is associated with energy eigenvalue, $\epsilon - 1$. This is why $\hat{a}$ is called a *lowering operator* - it transforms $\psi(x)$ into a lower energy eigenstate. Similarly,

$$\begin{aligned}\hat{H}\hat{a}^{\dagger}\psi(y) &= \left[\hat{H}, \hat{a}^{\dagger}\right]\psi(y) + \hat{a}^{\dagger}\hat{H}\psi(y) \\ &= \hat{a}^{\dagger}\psi(y) + \hat{a}^{\dagger}\epsilon\psi(y) \\ &= (\epsilon + 1)\,\hat{a}^{\dagger}\psi(y)\,;\end{aligned}$$

i.e., $\hat{a}^{\dagger}$ is a *raising operator.*

4.3.2 The ground state

With the raising and lowering operators, once we have one solution to the TISE, we can generate a ladder of solutions with higher or lower energies. So, what remains is to find one solution. We do this by noting that there must be a lowest energy state. In particular, there can be no solutions associated with negative energy - i.e., below the minimum of the potential. It is possible to prove this without raising and lowering operators. However, we need only suppose that there cannot be states extending to arbitrarily low energies. As such, there must be an eigenstate (or eigenstates) that cannot be lowered in energy by the lowering operator, $\hat{a}$. The relationship,

$$\hat{H}\hat{a}\psi(y) = (\epsilon - 1)\,\hat{a}\psi(y)\,,$$

is reconciled with this observation by noting that the relationship is still satisfied if

$$\hat{a}\psi(y) = 0. \tag{4.6}$$

This must be the case for any state that cannot be subject to further lowering. The above equation is a first order differential equation that is easily solved. It takes the form,

$$\frac{1}{\sqrt{2}}\left(\frac{d}{dy} + y\right)\psi(y) = 0$$

or

$$\frac{d}{dy}\psi(y) = -y\psi(y)$$

which gives

$$\frac{d\psi}{\psi} = -ydy.$$

Integrating both sides of this equation (from $y = 0$) gives

$$\begin{aligned}
\int_{\psi(0)}^{\psi(y)} \frac{d\psi}{\psi} &= -\int_0^y y' dy' \\
[\ln \psi]_{\psi(0)}^{\psi(y)} &= -\left[\frac{1}{2} y'^2\right]_0^y \\
\ln \psi\,(y) - \ln \psi\,(0) &= -\frac{y^2}{2} \\
\ln \frac{\psi\,(y)}{\psi\,(0)} &= -\frac{y^2}{2} \\
\psi\,(y) &= \psi\,(0) \exp\left(-\frac{y^2}{2}\right).
\end{aligned}$$

This solution is unique to the multiplicative constant, $\psi\,(0)$. This is because solutions to the TISE (an eigenvalue equation) are only determined up to a multiplicative constant. Imposing the normalization condition on $\psi\,(y)$ determines $\psi\,(0)$ - up to an arbitrary phase factor that we set equal to one. Up to the arbitrary phase factor (present for any eigenfunction solution), the lowest energy state is determined uniquely - i.e., there is only one lowest energy state.

The multiplicative constant, $C_0 = \psi\,(0)$, is determined by setting

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} |\psi\,(y)|^2 \, dy \\
&= |C_0|^2 \int_{-\infty}^{\infty} \left(\exp\left(-\frac{y^2}{2}\right)\right)^2 dy \\
&= |C_0|^2 \int_{-\infty}^{\infty} \exp\left(-y^2\right) \, dy \\
&= |C_0|^2 \, \pi^{1/2},
\end{aligned}$$

or

$$|C_0| = \pi^{-1/4}.$$

Here, we used the standard Gaussian integral,

$$\int_{-\infty}^{\infty} \exp\left(-y^2\right) \, dy = \pi^{1/2}.$$

We now have the lowest energy state - the ground state - of the harmonic oscillator,

$$\psi_0\,(y) = \pi^{-1/4} \exp\left(-\frac{y^2}{2}\right).$$

The arbitrary phase factor is set to one to give a real ground state. The associated energy eigenvalue is determined by applying $\hat{H}$. Specifically,

$$\begin{aligned}\hat{H}\psi_0(y) &= \left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)\psi_0(y)\\ &= \hat{a}^\dagger\hat{a}\psi_0(y)+\frac{1}{2}\psi_0(y)\\ &= \hat{a}^\dagger(0)+\frac{1}{2}\psi_0(y)\\ &= \frac{1}{2}\psi_0(y)\\ &= \epsilon_0\psi_0(y).\end{aligned}$$

The ground state energy, ϵ_0, is given by

$$\epsilon_0=\frac{1}{2}.$$

This corresponds to

$$E_{\text{vib},0}=\frac{\hbar\omega}{2},$$

with the units of energy made explicit. This energy is called the zero point energy - it is the lowest energy attainable, and it is greater than the minimum in the potential energy.

4.3.3 The excited states

The remaining eigenfunctions are determined by repeated application of the raising operator. For example,

$$\psi_1(y)=C_1\hat{a}^\dagger\psi_0(y),$$

where

$$\begin{aligned}\hat{H}\psi_1(y) &= (\epsilon_0+1)\psi_1(y)\\ &= \frac{3}{2}\psi_1(y)\\ &= \epsilon_1\psi_1(y).\end{aligned}$$

In general,

$$\epsilon_\upsilon=\upsilon+\frac{1}{2},\qquad \upsilon=0,1,2,\ldots,$$

and

$$E_{\text{vib},\upsilon}=\hbar\omega\left(\upsilon+\frac{1}{2}\right),\qquad \upsilon=0,1,2,\ldots.$$

Note that while $\hat{a}^\dagger\psi_0(x)$ is an eigenfunction of $\hat{H}$ associated with eigenvalue, ϵ_1, it is not necessarily normalized. The constant, C_1, is introduced to account for normalization. In general, we let

$$\psi_\upsilon(y)=C_\upsilon\hat{a}^\dagger\psi_{\upsilon-1}(y),\tag{4.7}$$

and determine C_υ by requiring that

$$\langle \psi_\upsilon \left| \psi_\upsilon \right\rangle = 1$$

which gives

$$\begin{aligned}
\left\langle C_\upsilon \hat{a}^\dagger \psi_{\upsilon-1} \left| C_\upsilon \hat{a}^\dagger \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \left\langle \hat{a}^\dagger \psi_{\upsilon-1} \left| \hat{a}^\dagger \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \left\langle \psi_{\upsilon-1} \left| \hat{a} \hat{a}^\dagger \psi_{\upsilon-1} \right.\right\rangle &= 1 \qquad \hat{a}^\dagger \text{ is the hermitian conjugate of } \hat{a} \\
\left|C_\upsilon\right|^2 \left\langle \psi_{\upsilon-1} \left| \left(\left[\hat{a}, \hat{a}^\dagger\right] + \hat{a}^\dagger \hat{a} \right) \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \left\langle \psi_{\upsilon-1} \left| \left(1 + \hat{H} - \frac{1}{2} \right) \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \left\langle \psi_{\upsilon-1} \left| \left(\hat{H} + \frac{1}{2} \right) \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \left\langle \psi_{\upsilon-1} \left| \left(\epsilon_{\upsilon-1} + \frac{1}{2} \right) \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 \upsilon \left\langle \psi_{\upsilon-1} \left| \psi_{\upsilon-1} \right.\right\rangle &= 1 \\
\left|C_\upsilon\right|^2 &= \frac{1}{\upsilon}
\end{aligned}$$

or

$$\left|C_\upsilon\right| = \upsilon^{-1/2}.$$

We set

$$C_{\upsilon} = \upsilon^{-1/2} \tag{4.8}$$

to obtain real wavefunctions. Here, we used the fact that $\hat{a}$ is the Hermitian conjugate of $\hat{a}^\dagger$. This follows because $-d/dy$ is the Hermitian conjugate of d/dy, which follows from integration by parts - see Example 2.2c.

Example 4.2. Evaluate $\psi_1(y)$ and $\psi_2(y)$ by successive applications of the raising operator to $\psi_0(y)$.

Solution 4.2. Apply $\hat{a}^\dagger$ to $\psi_0(y)$ to get $\psi_1(y)$:

$$\begin{aligned}
\psi_1(y) &= C_1 \hat{a}^\dagger \psi_0(y) \\
&= 1^{-1/2} \frac{1}{\sqrt{2}} \left(-\frac{d}{dy} + y \right) \psi_0(y) \\
&= \frac{1}{\sqrt{2}} \left(-\frac{d}{dy} + y \right) \pi^{-1/2} \exp\left(-\frac{y^2}{2} \right) \\
&= (2\pi)^{-1/2} \left(-\frac{d}{dy} \exp\left(-\frac{y^2}{2} \right) + y \exp\left(-\frac{y^2}{2} \right) \right) \\
&= (2\pi)^{-1/2} (y + y) \exp\left(-\frac{y^2}{2} \right) \\
&= (2\pi)^{-1/2} \, 2y \exp\left(-\frac{y^2}{2} \right).
\end{aligned}$$

Apply $\hat{a}^\dagger$ to $\psi_1(y)$ to get $\psi_2(y)$:

$$\begin{aligned}
\psi_2(y) &= C_2\hat{a}^\dagger\psi_1(y) \\
&= 2^{-1/2}\frac{1}{\sqrt{2}}\left(-\frac{d}{dy}+y\right)\psi_1(y) \\
&= \frac{1}{\sqrt{2}}\left(-\frac{d}{dy}+y\right)(2\pi)^{-1/2}\,2y\exp\left(-\frac{y^2}{2}\right) \\
&= (4\pi)^{-1/2}\left(-\frac{d}{dy}2y\exp\left(-\frac{y^2}{2}\right)+2y^2\exp\left(-\frac{y^2}{2}\right)\right) \\
&= (4\pi)^{-1/2}\left(-2+2y^2+2y^2\right)\exp\left(-\frac{y^2}{2}\right) \\
&= (4\pi)^{-1/2}\left(4y^2-2\right)\exp\left(-\frac{y^2}{2}\right).
\end{aligned}$$

The υ th energy eigenstate of the harmonic oscillator is given by

$$\psi_\upsilon(y) = A_\upsilon H_\upsilon(y)\exp\left(-\frac{y^2}{2}\right),$$

where $H_\upsilon(y)$ is the υ th Hermite polynomial. $H_\upsilon(y)=1$, $H_1(y)=2y$, and all remaining Hermite polynomials are obtained from the recursion,

$$H_{\upsilon+1}(y) = 2yH_\upsilon(y) - 2\upsilon H_{\upsilon-1}(y).$$

For example, $H_1(y)=4y^2-2$. The normalization constant, A_υ, is given by

$$A_\upsilon = \left(2^\upsilon\pi\upsilon!\right)^{-1/2}.$$

4.3.4 Expectation values

Raising and lowering operators can be used to determine position and momentum expectation values. For example, forming sums and differences of the raising and lowering operators gives

$$\begin{aligned}
\hat{a}+\hat{a}^\dagger &= \frac{1}{\sqrt{2}}\left(\frac{d}{dy}+y\right)+\frac{1}{\sqrt{2}}\left(-\frac{d}{dy}+y\right) \\
&= \sqrt{2}y
\end{aligned}$$

and

$$\begin{aligned}
\hat{a}-\hat{a}^\dagger &= \frac{1}{\sqrt{2}}\left(\frac{d}{dy}+y\right)-\frac{1}{\sqrt{2}}\left(-\frac{d}{dy}+y\right) \\
&= \sqrt{2}\frac{d}{dy} \\
&= \frac{\sqrt{2}i}{\hbar}\hat{p}.
\end{aligned}$$

So

$$y = \frac{1}{\sqrt{2}} \left(\hat{a} + \hat{a}^{\dagger} \right)$$

and

$$\hat{p} = -i \frac{\hbar}{\sqrt{2}} \left(\hat{a} - \hat{a}^{\dagger} \right).$$

The position expectation value is

$$\begin{aligned} \langle \psi_{\upsilon} \left| y \psi_{\upsilon} \right\rangle &= \frac{1}{\sqrt{2}} \left\langle \psi_{\upsilon} \left| \left(\hat{a} + \hat{a}^{\dagger} \right) \psi_{\upsilon} \right\rangle \right. \\ &= \frac{1}{\sqrt{2}} \left(\langle \psi_{\upsilon} \left| \hat{a} \psi_{\upsilon} \right\rangle + \langle \psi_{\upsilon} \left| \hat{a}^{\dagger} \psi_{\upsilon} \right\rangle \right) \end{aligned}$$

The action of the raising operator is provided by Eqs. 4.7 and 4.8 . Specifically,

$$\begin{aligned} \hat{a}^{\dagger} \psi_{\upsilon} &= C_{\upsilon+1}^{-1} \psi_{\upsilon+1} \qquad (4.9) \\ &= (\upsilon + 1)^{1/2} \psi_{\upsilon+1}. \end{aligned}$$

To get the action of the lowering operator, use the fact that $\hat{a}$ is the Hermitian conjugate of $\hat{a}^{\dagger}$. Specifcally, we consider

$$\begin{aligned} \left\langle \psi_{\upsilon-1} \left| \hat{a} \psi_{\upsilon} \right\rangle \right. &= \left\langle \hat{a}^{\dagger} \psi_{\upsilon-1} \left| \psi_{\upsilon} \right\rangle \right. \\ &= \left(\upsilon^{1/2} \right)^{*} \langle \psi_{\upsilon} \left| \psi_{\upsilon} \right\rangle \\ &= \upsilon^{1/2} \\ &= \upsilon^{1/2} \left\langle \psi_{\upsilon-1} \left| \psi_{\upsilon-1} \right\rangle \right., \end{aligned}$$

and note that

$$\hat{a} \psi_{\upsilon} = \upsilon^{1/2} \psi_{\upsilon-1}. \qquad (4.10)$$

Now we evaluate the position expectation

$$\begin{aligned} \langle \psi_{\upsilon} \left| y \psi_{\upsilon} \right\rangle &= \frac{1}{\sqrt{2}} \left(\langle \psi_{\upsilon} \left| \hat{a} \psi_{\upsilon} \right\rangle + \langle \psi_{\upsilon} \left| \hat{a}^{\dagger} \psi_{\upsilon} \right\rangle \right) \\ &= \frac{1}{\sqrt{2}} \left(\upsilon^{1/2} \underbrace{\langle \psi_{\upsilon} \left| \psi_{\upsilon-1} \right\rangle}_{=0} + (\upsilon + 1)^{1/2} \underbrace{\langle \psi_{\upsilon} \left| \psi_{\upsilon+1} \right\rangle}_{=0} \right) \qquad \text{the } \psi_{\upsilon} \text{ are mutually orthogonal} \\ &= 0. \end{aligned}$$

This comes as no surpise, since y is the scaled displacement from equilibrium - its expectation value is zero for an energy eigenstate. The harmonic oscillator oscillates symmetrically about equilibrium. The expectation of y^2 gives the variance of the scaled displacement;

$$\begin{aligned}
\langle\psi_\upsilon\,|y^2\psi_\upsilon\rangle &= \frac{1}{2}\langle\psi_\upsilon\,\left|\left(\hat{a}+\hat{a}^\dagger\right)^2\psi_\upsilon\right\rangle \\
&= \frac{1}{2}\left(\langle\psi_\upsilon\,|\hat{a}^2\psi_\upsilon\rangle+\langle\psi_\upsilon\,|\hat{a}^{\dagger 2}\psi_\upsilon\rangle+\langle\psi_\upsilon\,|\hat{a}\hat{a}^\dagger\psi_\upsilon\rangle+\langle\psi_\upsilon\,|\hat{a}^\dagger\hat{a}\psi_\upsilon\rangle\right) \\
&= \frac{1}{2}\Bigg((\upsilon\,(\upsilon-1))^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon-2}\rangle}_{=0}+(\upsilon\,(\upsilon+1))^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon+2}\rangle}_{=0} \\
&\qquad (\upsilon+1)\underbrace{\langle\psi_\upsilon\,|\psi_\upsilon\rangle}_{=1}+\upsilon\underbrace{\langle\psi_\upsilon\,|\psi_\upsilon\rangle}_{=1}\Bigg) \\
&= \frac{1}{2}(2\upsilon+1)=\upsilon+\frac{1}{2}.
\end{aligned}$$

The standard deviation of the displacement, $x=\alpha y$, is just

$$\sigma_x=\langle\psi_\upsilon\,\left|x^2\psi_\upsilon\right\rangle^{1/2}=\left(\upsilon+\frac{1}{2}\right)^{1/2}\alpha.$$

$\alpha=\hbar^{1/2}(mk)^{-1/4}$ is seen to be a characteristic distance scale - smaller for larger masses and stiffer bonds (i.e., large k).

The expectation of momentum is

$$\begin{aligned}
\langle\psi_\upsilon\,|\hat{p}\psi_\upsilon\rangle &= -i\frac{\hbar}{\sqrt{2}}\left(\langle\psi_\upsilon\,|\hat{a}\psi_\upsilon\rangle-\langle\psi_\upsilon\,|\hat{a}^\dagger\psi_\upsilon\rangle\right) \\
&= -i\frac{\hbar}{\sqrt{2}}\left(\upsilon^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon-1}\rangle}_{=0}+(\upsilon+1)^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon+1}\rangle}_{=0}\right) \qquad \text{the } \psi_\upsilon \text{ are mutually orthogonal} \\
&= 0.
\end{aligned}$$

The expectation value of momentum is always zero for bound energy eigenstates. The expectation of $\hat{p}^2$ gives the variance of scaled-displacement momentum,

$$\begin{aligned}
\langle\psi_\upsilon\,|\hat{p}^2\psi_\upsilon\rangle &= -\frac{\hbar^2}{2}\langle\psi_\upsilon\,\left|\left(\hat{a}-\hat{a}^\dagger\right)^2\psi_\upsilon\right\rangle \\
&= -\frac{\hbar^2}{2}\left(\langle\psi_\upsilon\,|\hat{a}^2\psi_\upsilon\rangle+\langle\psi_\upsilon\,|\hat{a}^{\dagger 2}\psi_\upsilon\rangle-\langle\psi_\upsilon\,|\hat{a}\hat{a}^\dagger\psi_\upsilon\rangle-\langle\psi_\upsilon\,|\hat{a}^\dagger\hat{a}\psi_\upsilon\rangle\right) \\
&= -\frac{\hbar^2}{2}\Bigg((\upsilon\,(\upsilon-1))^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon-2}\rangle}_{=0}+(\upsilon\,(\upsilon+1))^{1/2}\underbrace{\langle\psi_\upsilon\,|\psi_{\upsilon+2}\rangle}_{=0} \\
&\qquad -(\upsilon+1)\underbrace{\langle\psi_\upsilon\,|\psi_\upsilon\rangle}_{=1}-\upsilon\underbrace{\langle\psi_\upsilon\,|\psi_\upsilon\rangle}_{=1}\Bigg) \\
&= \frac{\hbar^2}{2}(2\upsilon+1)=\hbar^2\left(\upsilon+\frac{1}{2}\right).
\end{aligned}$$

Consequently, the standard deviation of the displacement momentum, $\hat{p}_x=-i\hbar d/dx=\hat{p}/\alpha$, is just

$$\sigma_{p_x}=\frac{\hbar}{\alpha}\left(\upsilon+\frac{1}{2}\right)^{1/2}.$$

The product of σ_x and σ_{p_x} is given by

$$\sigma_x \sigma_{p_x} = \hbar \left(\upsilon + \frac{1}{2} \right) \geq \frac{\hbar}{2},$$

in accord with the Heisenberg uncertainty principle. For the ground state, the equality holds - the ground state is a *minimum uncertainty* wavefunction.

Example 4.3. Suppose

$$\psi(y) = \frac{1}{\sqrt{2}} \left(\psi_0(y) + i\psi_1(y) \right)$$

where $\psi_0(y)$ and $\psi_1(y)$ are the first two harmonic oscillator energy eigenstates.

(a) What is the scaled energy expectation value, $\langle \psi \left| \hat{H} \psi \right\rangle$, for this state?

(b) What is the scaled displacement, y, expectation value for this state?

(c) What is the uncertainty in scaled energy,

$$\sigma_\epsilon = \left(\langle \psi \left| \hat{H}^2 \psi \right\rangle - \langle \psi \left| \hat{H} \psi \right\rangle^2 \right)^{1/2},$$

for this state?

(d) What are the scaled displacement, y, expectation values for the states,

$$\psi_+(y) = \frac{1}{\sqrt{2}} \left(\psi_0(y) + \psi_1(y) \right)$$

and

$$\psi_-(y) = \frac{1}{\sqrt{2}} \left(\psi_0(y) - \psi_1(y) \right)?$$

Solution 4.3.

(a) The energy expectation is

$$\begin{aligned}
\langle \psi \left| \hat{H} \psi \right\rangle &= \left\langle \frac{1}{\sqrt{2}} \left(\psi_0 + i\psi_1 \right) \left| \hat{H} \frac{1}{\sqrt{2}} \left(\psi_0 + i\psi_1 \right) \right. \right\rangle \\
&= \frac{1}{2} \langle \psi_0 + i\psi_1 \left| \hat{H}\psi_0 + i\hat{H}\psi_1 \right\rangle \\
&= \frac{1}{2} \langle \psi_0 + i\psi_1 \left| \epsilon_0 \psi_0 + i\epsilon_1 \psi_1 \right\rangle \\
&= \frac{1}{2} \left(\epsilon_0 \underbrace{\langle \psi_0 \left| \psi_0 \right\rangle}_{=1} + \epsilon_1 \underbrace{\langle i\psi_1 \left| i\psi_1 \right\rangle}_{=i^* i \times 1 = 1} \right) \qquad \langle \psi_0 \left| \psi_1 \right\rangle = 0 \\
&= \frac{1}{2} \left(\epsilon_0 + \epsilon_1 \right) \\
&= \frac{1}{2} \left(\frac{1}{2} + \frac{3}{2} \right) = 1
\end{aligned}$$

(b) The y expectation is

$$\begin{aligned}
\langle \psi \left| y\psi \right\rangle &= \left\langle \frac{1}{\sqrt{2}}\left(\psi_0 + i\psi_1\right) \middle| \frac{1}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right)\frac{1}{\sqrt{2}}\left(\psi_0 + i\psi_1\right) \right\rangle \\
&= \frac{1}{2\sqrt{2}}\left\langle \psi_0 + i\psi_1 \left| \left(\hat{a} + \hat{a}^\dagger\right)\psi_0 + i\left(\hat{a} + \hat{a}^\dagger\right)\psi_1 \right.\right\rangle \\
&= \frac{1}{2\sqrt{2}}\left\langle \psi_0 + i\psi_1 \left| \psi_1 + i\left(\psi_0 + \sqrt{2}\psi_2\right) \right.\right\rangle \\
&= \frac{1}{2\sqrt{2}}\left(\langle \psi_0 \left| i\psi_0 \right\rangle + \langle i\psi_1 \left| \psi_1 \right\rangle\right) && \text{mutually orthogonal states} \\
&= \frac{1}{2\sqrt{2}}\left(i - i\right) = 0.
\end{aligned}$$

(c) The calculation of the expectation of $\hat{H}^2$ is exactly the same as the calculation of $\hat{H}$ in part a, except that ϵ_0 is replaced by ϵ_0^2 and ϵ_1 is replaced by ϵ_1^2.

$$\begin{aligned}
\left\langle \psi \middle| \hat{H}^2\psi \right\rangle &= \left\langle \frac{1}{\sqrt{2}}\left(\psi_0 + i\psi_1\right) \middle| \hat{H}^2\frac{1}{\sqrt{2}}\left(\psi_0 + i\psi_1\right) \right\rangle \\
&= \frac{1}{2}\left\langle \psi_0 + i\psi_1 \left| \epsilon_0^2\psi_0 + i\epsilon_1^2\psi_1 \right.\right\rangle \\
&= \frac{1}{2}\left(\epsilon_0^2 + \epsilon_1^2\right) \\
&= \frac{1}{2}\left(\frac{1}{4} + \frac{9}{4}\right) = \frac{5}{4}
\end{aligned}$$

The uncertainty in scaled energy is

$$\begin{aligned}
\sigma_\epsilon &= \left(\left\langle \psi \middle| \hat{H}^2\psi \right\rangle - \left\langle \psi \middle| \hat{H}\psi \right\rangle^2\right)^{1/2} \\
&= \left(\frac{\epsilon_0^2 + \epsilon_1^2}{2} - \left(\frac{\epsilon_0 + \epsilon_1}{2}\right)^2\right)^{1/2} \\
&= \left(\frac{\epsilon_0^2 + \epsilon_1^2}{2} - \frac{\epsilon_0^2 + 2\epsilon_0\epsilon_1 + \epsilon_1^2}{4}\right)^{1/2} \\
&= \left(\frac{\epsilon_0^2 - 2\epsilon_0\epsilon_1 + \epsilon_1^2}{4}\right)^{1/2} \\
&= \frac{\left|\epsilon_0 - \epsilon_1\right|}{2} = \frac{\epsilon_1 - \epsilon_0}{2} = \frac{1}{2}.
\end{aligned}$$

(d) Calculation of the y expectation for states, $\psi_\pm$, is analogous to the calcuation for ψ in part b. Simply substitute ±1 for i in the second last line.

$$\begin{aligned}\langle\psi_\pm\,|y\psi_\pm\rangle &= \left\langle\frac{1}{\sqrt{2}}(\psi_0\pm\psi_1)\,\middle|\,\frac{1}{\sqrt{2}}\left(\hat{a}+\hat{a}^\dagger\right)\frac{1}{\sqrt{2}}(\psi_0\pm\psi_1)\right\rangle\\ &= \frac{1}{2\sqrt{2}}\left(\langle\psi_0\,|\pm\psi_0\rangle+\langle\pm\psi_1\,|\psi_1\rangle\right)\\ &= \pm\frac{1}{2\sqrt{2}}(1+1)=\pm\frac{1}{\sqrt{2}}.\end{aligned}$$

Example 4.4. Consider a harmonic oscillator initially in the state,

$$\psi_+(y)=\frac{1}{\sqrt{2}}(\psi_0(y)+\psi_1(y)).$$

(a) Write down the state of the system after the passage of time, t. Show that the time evolving state is periodic. What is the frequency of its oscillations?

(b) Show that the expectation of y is periodic. What are the amplitude and frequency of these oscillations?

Solution 4.4.

(a) The time evolution of a wavefunction is given by writing the wavefunction as a linear combination of energy eigenstates. The coefficient of each energy eigenstate is multiplied by the phase factor, $\exp(-iE_{\mathrm{vib},\,v}t/\hbar)$, which depends on time. This is the solution to the *time dependent Schrodinger equation* - see Sec. 2.2. The current intial state is already written as a combination of the bottom two energy eigenstates. At time t, we have

$$\begin{aligned}\Psi(y,t) &= \frac{1}{\sqrt{2}}\left(\exp\left(-i\frac{\hbar\omega t}{2\hbar}\right)\psi_0(y)+\exp\left(-i\frac{3\hbar\omega t}{2\hbar}\right)\psi_1(y)\right)\\ &= \frac{\exp\left(-i\frac{\omega t}{2}\right)}{\sqrt{2}}(\psi_0(y)+\exp(-i\omega t)\,\psi_1(y)).\end{aligned}$$

The function $\exp(-i\omega t)$ is periodic with period, $\tau=2\pi/\omega$, while $\exp\left(-i\frac{\omega t}{2}\right)$ has twice this period. Thus, the frequency of the oscillations is $\omega/4\pi$.

(b) See Example 4.3 b and d. The expectation of y, at time t, is given by

$$\begin{aligned}
\langle\psi\left|y\psi\right\rangle &= \left\langle\frac{\exp\left(-i\frac{\omega t}{2}\right)}{\sqrt{2}}\left(\psi_0+\exp\left(-i\omega t\right)\psi_1\right)\right. \\
&\quad \left|\frac{1}{\sqrt{2}}\left(\hat{a}+\hat{a}^\dagger\right)\frac{\exp\left(-i\frac{\omega t}{2}\right)}{\sqrt{2}}\left(\psi_0+\exp\left(-i\omega t\right)\psi_1\right)\right\rangle \\
&= \frac{\exp\left(i\frac{\omega t}{2}-i\frac{\omega t}{2}\right)}{2\sqrt{2}}\left\langle\psi_0+\exp\left(-i\omega t\right)\psi_1\left|\left(\hat{a}+\hat{a}^\dagger\right)\psi_0+\exp\left(-i\omega t\right)\left(\hat{a}+\hat{a}^\dagger\right)\psi_1\right.\right\rangle \\
&= \frac{1}{2\sqrt{2}}\left\langle\psi_0+\exp\left(-i\omega t\right)\psi_1\left|\psi_1+\exp\left(-i\omega t\right)\left(\psi_0+\sqrt{2}\psi_2\right)\right.\right\rangle \\
&= \frac{1}{2\sqrt{2}}\left(\left\langle\psi_0\left|\exp\left(-i\omega t\right)\psi_0\right.\right\rangle+\left\langle\exp\left(-i\omega t\right)\psi_1\left|\psi_1\right.\right\rangle\right) \qquad \text{mutually orthogonal states} \\
&= \frac{1}{2\sqrt{2}}\left(\exp\left(-i\omega t\right)+\exp\left(i\omega t\right)\right) \\
&= \frac{1}{\sqrt{2}}\cos\left(\omega t\right).
\end{aligned}$$

This function is periodic with frequency, $\omega/2\pi$. The amplitude of the oscillations is $1/\sqrt{2}$.

4.4 The classical harmonic oscillator

To best understand the properties of the harmonic oscillator eigenstates, it is instructive to consider a classical treatment of the harmonic oscillator. In classical mechanics, the Hamiltonian is a function of variables, x and p_x. The shape of the path in the x, p_x plane is determined by energy conservation.

$$H=\frac{1}{2m}p_x^2+\frac{k}{2}x^2=E.$$

The oscillator follows a path along the ellipse determined by this equation. This ellipse extends from $-\sqrt{2E/k}$ to $\sqrt{2E/k}$ in the x direction, and $-\sqrt{2mE}$ to $\sqrt{2mE}$ in the p_x direction. The points,

$$x_{\min}=-\sqrt{2E/k}$$

and

$$x_{\max}=\sqrt{2E/k}$$

are called the *turning points* of the oscillator. They are found by setting $p=0$ into the above energy conservation equation. When the oscillator reaches $x=x_{\min}$ or $x_{\max}$, the sign of p_x reverses and x begins to change in the reverse direction. Positions beyond $x=x_{\min}$ or $x_{\max}$ (i.e., $x<x_{\min}$ or $x>x_{\max}$) are not accesible to oscillatory. These regions are said to *classically forbidden.* Since the quantum mechanical energy eigenstates are non-zero for all x from

$-\infty$ to ∞, the quantum harmonic oscillator exhibits *barrier penetration.* The wavefunctions do, however, decay rapidly in the classically forbidden regions. Nevertheless, there is non-zero probability for finding the oscillator in classically forbidden configurations.

In classical mechanics, the probability density associated with position, is just the relative amount of time spent by the oscillator at each position - over one complete oscillation. This is proportional to the absolute value of the velocity, at each point,

$$\rho_{\mathrm{CM}}(x) \propto \frac{1}{\sqrt{2\left(E - kx^2/2\right)/m}}.$$

The classical harmonic oscillator is more likely found near its turning points, since the velocity goes to zero there. It is least likely to be found at the equilibrium position where it is moving fastest.

For the ground state, the quantum position distribution (shown here in terms of unscaled displacement, x),

$$|\psi_0(x)|^2 = \frac{\alpha^2}{\pi}\exp\left(-\frac{x^2}{\alpha^2}\right),$$

is very different from the classical distribution - see Fig. 4.2 below. The quantum ground state distribution is peaked at the equilibrium position, and it exhibits significant penetration into classically forbidden regions.For the fifth excited state of the harmonic oscillator, the quantum distribution begins to resemble the classical distribution, except that it has an interference pattern - associated with inteference between left and right going waves.For the twelfth excited state of the harmonic oscillator, the quantum distribution oscillates - with more oscillations - about the classical distribution.

In the high quantum number limit, the quantum distribution oscillates rapidly between zero and twice the classical distribution. Averaging over small x intervals makes the quantum distribution indistinguishable from the quantum distribution, in the large v limit.

The time dependence of x and p_x, in a classical treatement is given by solving Hamilton's equations:

$$\begin{aligned}\frac{dx}{dt} &= \frac{\partial H}{\partial p_x} \\ &= \frac{p_x}{m}\end{aligned}$$

and

$$\begin{aligned}\frac{dp_x}{dt} &= -\frac{\partial H}{\partial x} \\ &= -kx.\end{aligned}$$

Taking the time derivative of the first equation gives

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x$$

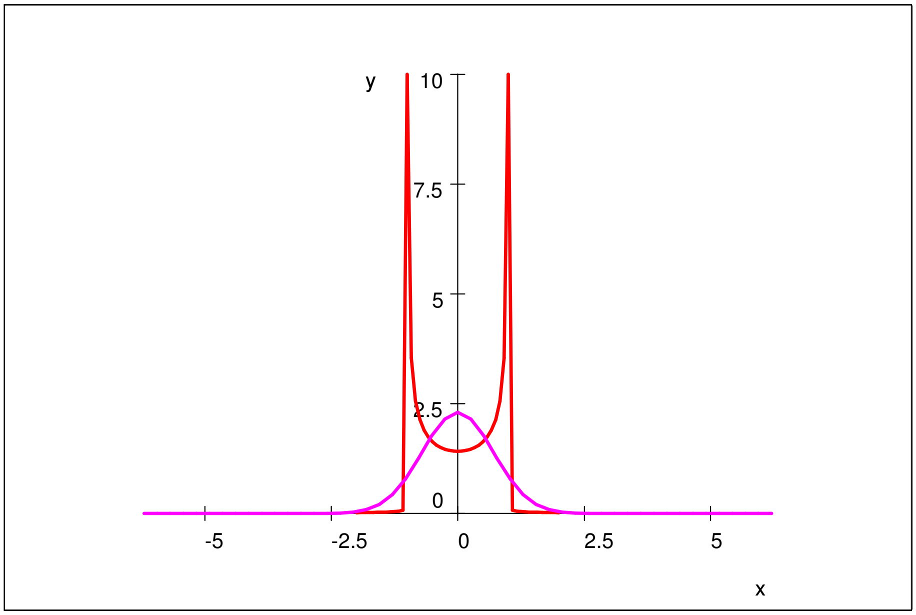


Figure 4.2: The quantum ground state position probability density (solid) of the harmonic oscillator, together with the classical probability density (dashed) - for the same energy.

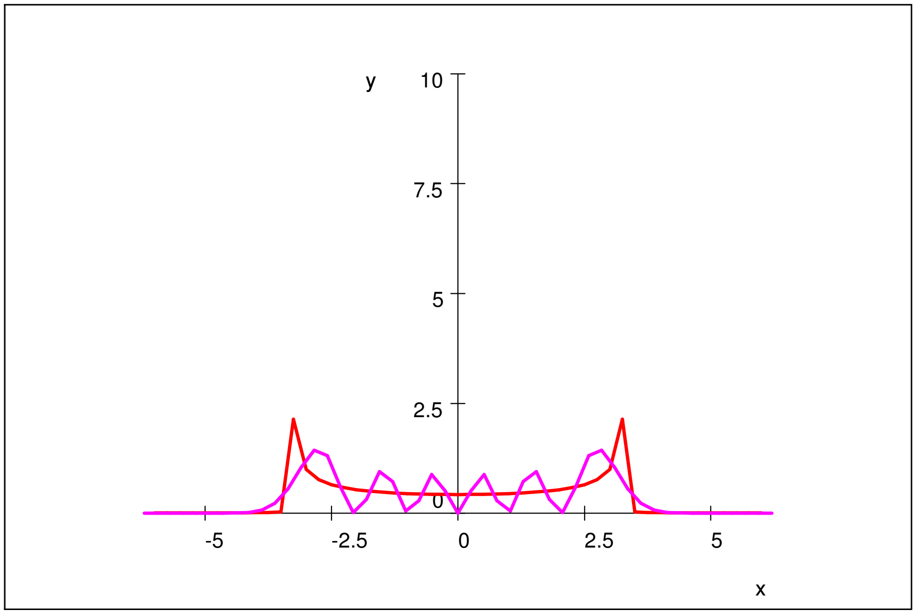


Figure 4.3: The position probability density (solid) for the fifth excited state of the harmonic oscillator, together with the corresponding classical probability density (dashed).

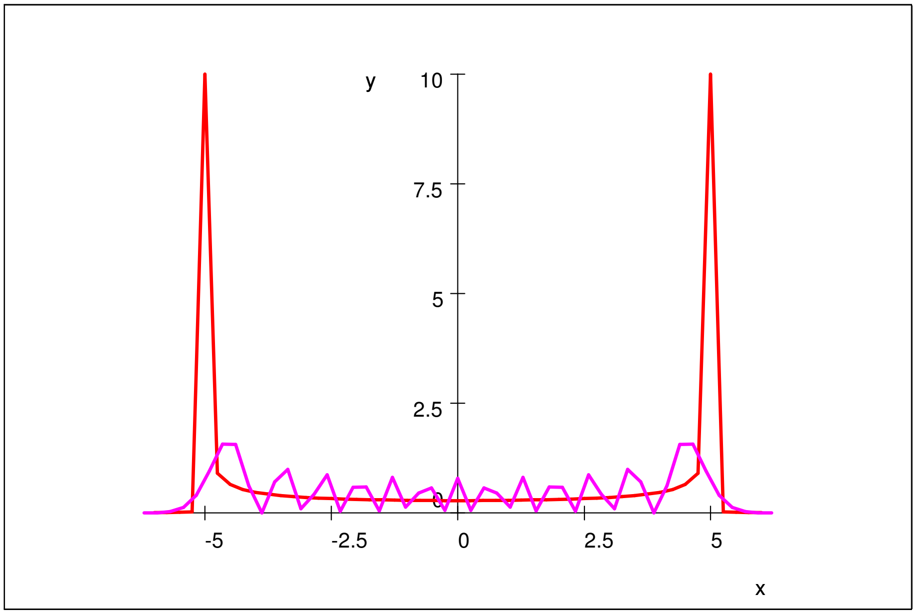


Figure 4.4: The position probability density (solid) for the twelfth excited state of the harmonic oscillator, together with the corresponding classical probability density (dashed).

or

$$\frac{d^2x}{dt^2} = -\omega^2 x$$

which has solutions

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

with corresponding momentum $(p_x = mdx/dt)$,

$$p_x(t) = \omega B\cos(\omega t) - \omega A\sin(\omega t).$$

The coefficients are determiend by the initial conditions,

$$A = x(0)$$

and

$$B = p_x(0)/\omega.$$

Note the appearance of the angular frequency $\omega = (k/m)^{1/2}$, which arose in the quantum treatment through the spacing of the vibrational energy levels, $\hbar\omega$. The vibrational frequency, in cycles per second, is just

$$\nu_{\text{vib}} = \frac{\omega}{2\pi} = \frac{1}{2\pi}\left(\frac{k}{m}\right)^{1/2}.$$

4.5 Infrared spectroscopy

A diatomic molecule can make vibrational transitions by absorbing or emitting a photon with energy,

$$h\nu = \Delta E,$$

where $\Delta E = E_{\text{final}} - E_{\text{intial}}$ is the transition energy difference. Transitions arise because of coupling of the molecular *electric dipole* to the oscillating electric field of light - see Sec. 6.3.1. The *transition probability* is proportional to the mod-square of the *transition matrix element*,

$$\langle \psi_{\upsilon'} | \hat{\mu}_e \psi_\upsilon \rangle,$$

where $\hat{\mu}_e$ is the electric dipole of the diatomic molecule. Vibrational transitions depend upon $\hat{\mu}_e$ varying with deviation from the equilibrium geometry; i.e., $\hat{\mu}_e$ must vary with x. Note that for small x (the usual realm of the harmonic approximation),

$$\begin{aligned} \hat{\mu}_e(x) &= \mu(x) && \hat{\mu}_e \text{ is a multiplication operator} \\ &= \mu_{\text{eq}} + \frac{d\mu}{dx}x, && \text{small } x \end{aligned}$$

where $\mu_{\text{eq}} = \mu(0)$ is the equilibrium dipole moment of the molecule.

The transition matrix element takes the form,

$$\begin{aligned}
\langle \psi_{\upsilon'} \left| \hat{\mu}_e \, \psi_{\upsilon} \right\rangle &= \left\langle \psi_{\upsilon'} \left| \left(\mu_{\text{eq}} + \frac{d\mu}{dx} x \right) \psi_{\upsilon} \right. \right\rangle \\
&= \mu_{\text{eq}} \langle \psi_{\upsilon'} | \psi_{\upsilon} \rangle + \frac{d\mu}{dx} \langle \psi_{\upsilon'} | x \, \psi_{\upsilon} \rangle \qquad \mu_{\text{eq}} \text{ and } \frac{d\mu}{dx} \text{ are constants} \\
&= \frac{d\mu}{dx} \langle \psi_{\upsilon'} | x \, \psi_{\upsilon} \rangle . \qquad \langle \psi_{\upsilon'} | \psi_{\upsilon} \rangle = 0 \text{ for } \upsilon' \neq \upsilon
\end{aligned}$$

Thus, the transition vanishes if $d\mu/dx = 0$ (i.e., the dipole moment does not change with vibrational motion). For example, our atmosphere consists predominantly of nitrogen and oxygen. Because these diatomic molecules are homonuclear, they have no electric dipole. This is true for any extension, x, away from equilibrium. Thus, for N_2 and O_2, $\mu_{\text{eq}} = 0$ and $d\mu/dx = 0$. On account of the latter observation, the vibrational transitions of N_2 and O_2 are electric dipole forbidden. Consequently, N_2 and O_2 are not greenhouse gases. Greenhouse gases absorb infrared radiation emitted from the surface of the earth. They re-emit the radiation. However, one half of the re-emitted radiation is returns to the earth, mitigating the cooling associated with the infrared emission into space. The principle greenhouses in the earth's atmosphere are H_2O and CO_2 (CH_4 is also important). Note that although CO_2 has no net electric dipole at equilibrium, bending motion (a type of vibration) gives rise to an electric dipole. Thus, in this case, $d\mu/dx \neq 0$, where x equals displacement of the bond angle from 180°. CO_2 exhibits electric-dipole allowed bending transitions shifting the energy balance at the earth's surface to make it warmer.

The transition matrix element is proportional to

$$\langle \psi_{\upsilon'} | x \, \psi_{\upsilon} \rangle .$$

Since $x = \alpha y$ can be written as the difference of the raising and lowering operators, the right side of the inner product takes the form,

$$x \, \psi_{\upsilon} = \frac{\alpha}{\sqrt{2}} \left((\upsilon + 1)^{1/2} \, \psi_{\upsilon+1} + \upsilon^{1/2} \psi_{\upsilon-1} \right)$$

and

$$\begin{aligned}
\langle \psi_{\upsilon'} | x \, \psi_{\upsilon} \rangle &= \left\langle \psi_{\upsilon'} \left| \frac{\alpha}{\sqrt{2}} \left((\upsilon + 1)^{1/2} \, \psi_{\upsilon+1} + \upsilon^{1/2} \psi_{\upsilon-1} \right) \right. \right\rangle \\
&= \frac{\alpha}{\sqrt{2}} \left((\upsilon + 1)^{1/2} \langle \psi_{\upsilon'} | \psi_{\upsilon+1} \rangle + \upsilon^{1/2} \langle \psi_{\upsilon'} | \psi_{\upsilon-1} \rangle \right) \\
&= \frac{\alpha}{\sqrt{2}} \left((\upsilon + 1)^{1/2} \, \delta_{\upsilon', \upsilon+1} + \upsilon^{1/2} \delta_{\upsilon', \upsilon-1} \right) \\
&= \begin{cases} \frac{\alpha}{\sqrt{2}} (\upsilon + 1)^{1/2}, & \upsilon' = \upsilon + 1 \\ \frac{\alpha}{\sqrt{2}} \upsilon^{1/2}, & \upsilon' = \upsilon - 1 \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{4.11}$$

This gives the selection rule for vibrational transitions:

$$\Delta\upsilon = \pm 1. \tag{4.12}$$

Vibrational transitions increase (case of absorption) or decrease (case of emission) the vibrational quantum number only by 1. The associated vibrational transition frequency is given by setting the photon energy to the change in energy of the oscillator. For absorption from state υ, or emission from $\upsilon + 1$, we have

$$\begin{aligned} h\nu &= E_{\upsilon+1} - E_{\upsilon} \\ &= \hbar\omega\left(\upsilon + \frac{3}{2}\right) - \hbar\omega\left(\upsilon + \frac{1}{2}\right) \\ &= \hbar\omega = h\nu_{\text{vib}}. \end{aligned}$$

Thus, the frequency of the light that induces a transition equals the oscillation frequency of the classical harmonic oscillator. It is also the frequency of light emitted by the oscillator. Because of the above selection rule, and the equal spacing of harmonic oscillator energy eigenvalues, there is only one vibrational transition frequency. The selection rule ensures there is only one transition frequency for a given level. The equal spacing of levels makes this frequency the same for all levels. Thus, excited states have the transition frequency as the ground state. In any case, vibrational level spacings are large enough such that at ordinary temperatures most molecules are in the vibrational ground state. Only for low frequency vibrations, is there significant population of first excited states (or higher).

Real diatomic molecule potentials are only quadratic close to the equilibrium geometry. Typically, the potential is less steep than quadratic for stretched bonds, and steeper than quadratic for compressed bonds. Deviation from a quadratic potential is called *anharmonicity.* It typically results in transition frequencies that are slightly smaller for excited states - the energy spacing actually decreases with increasing υ. Since vibrational excited state populations are usually small at ordinary temperatures, transitions originating in vibrational excited states - appearing at lower frequency - have small intensity. The associated absorption or emission bands are called *hot bands.* Note that we use the term *band* to describe the manifestation of vibrational transitions in absorption of emission spectra because there is a range of frequencies absorbed/emitted about the vibrational frequency. The structure of a band results from rotatonal transitions that accompany vibrational transitions. The rotational transitions are associated with the angular kinetic energy term we left out at the beginning of this chapter - see Chapter 5.

Polyatomic molecules have multiple vibrational modes. For example, CO_2 has two stretch modes - symmetric and antisymmetric combinations of the two bond stretch vibrations - and two degenerate (they have the same frequency) bends. Only the bending modes have $d\mu/dx \neq 0$ and produce infrared absorption spectra. In general, molecules have vibrational degrees of freedom that couple the displacements of all atoms from equilibrium - each degree of freedom couples the atomic motions in specific ways. However, some of these degrees of freedom are primarily associated with specific bond vibrations. This occurs when the bond frequency is different from those of neighboring bonds. Otherwise, bonds with the same frequency couple - their vibrations get mixed together. The isolated degrees of freedom have frequencies characteristic of specific bonds. This allows functional groups to be identified from an infrared spectrum. Also, bonds between atoms with different electronegativity - such as a ketone C=O - have larger electric dipoles, and larger electric dipole derivative. They appear as strong absorption bands in infrared spectra.

Example 4.5. Which of the following matrix elements of harmonic oscillator eigenstates are non-zero?

(a) $\langle \psi_1 \left| \hat{p}\, \psi_1 \right\rangle$
(b) $\langle \psi_2 \left| \hat{p}\, \psi_1 \right\rangle$
(c) $\langle \psi_3 \left| \hat{p}\, \psi_1 \right\rangle$
(d) $\langle \psi_2 \left| \hat{p}^2\, \psi_1 \right\rangle$
(e) $\langle \psi_3 \left| \hat{p}^2\, \psi_1 \right\rangle$

Solution 4.5. Matrix elements b and e are non-zero. $\hat{p}$ is a linear combination of the lowering and raising operators. It acts on ψ_1 to give a combination of ψ_0 and ψ_2. Thus, of the first three matrix elements, only b is non-zero. A second action of $\hat{p}$ produces a combination of ψ_1 and ψ_3. Thus, of the remaining matrix elements, only e is non-zero.

Example 4.6. Suppose you have a sample of diatomic molecules in their vibrational ground and first exciting states, with populations, ρ_0 and ρ_1, respectively. Using the harmonic approximation, determine the peak intensity of the 1 to 2 dipole transition relative to the 0 to 1 peak intensity.

Solution 4.6. The transition integrals we need are already evaluated in Eq. 4.11 . We repeat the computation here for the specific transitions in question.

$$\begin{aligned}
\langle \psi_1 \left| x\, \psi_0 \right\rangle &= \left\langle \psi_1 \left| \frac{\alpha}{\sqrt{2}} \left(\psi_1 + 0 \right) \right. \right\rangle \\
&= \frac{\alpha}{\sqrt{2}} \underset{=1}{\langle \psi_1 \left| \psi_1 \right\rangle} \\
&= \frac{\alpha}{\sqrt{2}}
\end{aligned}$$

and

$$\begin{aligned}
\langle \psi_2 \left| x\, \psi_1 \right\rangle &= \left\langle \psi_2 \left| \frac{\alpha}{\sqrt{2}} \left(\sqrt{2}\psi_2 + \psi_0 \right) \right. \right\rangle \\
&= \frac{\alpha}{\sqrt{2}} \left(\sqrt{2} \underset{=1}{\langle \psi_2 \left| \psi_2 \right\rangle} + \underset{=0}{\langle \psi_2 \left| \psi_1 \right\rangle} \right) \\
&= \alpha.
\end{aligned}$$

The peak intensities are proportional to the square of these transition integrals, and to the population of the initial state. Therefore, the 1 to 2 peak intensity is

$$\frac{\alpha^2 \rho_1}{\frac{\alpha^2}{2}\rho_0} = 2\frac{\rho_1}{\rho_0},$$

relative to the 0 to 1 peak intensity. Of course, in the harmonic approximation, these transitions occur at the same frequency. However, high resolution gas phase spectroscopy can easily resolve the small shift in frequency due to anharmonicity. Thus, the 1 to 2 hot appears in a real spectrum with intensity about double the excited state population relative to the ground state.

4.6 Appendix: kinetic energy in spherical coordinates

To get the relative kinetic energy in Cartesian coordinates we need to express the derivative operators, $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$ in terms of derivatives with respect to spherical coordinates. This is done using the chain rule. It requires the derivatives of the spherical coordinates in terms of Cartesian coordinates. Specifically,

$$\begin{aligned}
\frac{\partial\theta}{\partial x} &= \frac{\partial\theta}{\partial\cos\theta}\frac{\partial\cos\theta}{\partial x} \\
&= -\frac{1}{\sin\theta}\frac{\partial}{\partial x}\frac{z}{(x^2+y^2+z^2)^{1/2}} \\
&= \frac{1}{\sin\theta}\frac{xz}{(x^2+y^2+z^2)^{3/2}} \\
&= \frac{1}{\sin\theta}\frac{\cos\phi\sin\theta\cos\theta}{r} \\
&= \frac{\cos\phi\cos\theta}{r},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\theta}{\partial y} &= \frac{\partial\theta}{\partial\cos\theta}\frac{\partial\cos\theta}{\partial y} \\
&= -\frac{1}{\sin\theta}\frac{\partial}{\partial y}\frac{z}{(x^2+y^2+z^2)^{1/2}} \\
&= \frac{1}{\sin\theta}\frac{yz}{(x^2+y^2+z^2)^{3/2}} \\
&= \frac{\sin\phi\cos\theta}{r},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\theta}{\partial z} &= \frac{\partial\theta}{\partial\cos\theta}\frac{\partial\cos\theta}{\partial z} \\
&= -\frac{1}{\sin\theta}\frac{\partial}{\partial z}\frac{z}{(x^2+y^2+z^2)^{1/2}} \\
&= -\frac{1}{\sin\theta}\left[\frac{1}{(x^2+y^2+z^2)^{1/2}} - \frac{z^2}{(x^2+y^2+z^2)^{3/2}}\right] \\
&= -\frac{1}{\sin\theta}\frac{1-\cos^2\theta}{r} \\
&= -\frac{\sin\theta}{r},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\phi}{\partial x} &= \frac{\partial\phi}{\partial\tan\phi}\frac{\partial\tan\phi}{\partial x} \\
&= \frac{1}{\sec^2\phi}\frac{\partial}{\partial x}\frac{y}{x} \\
&= -\cos^2\phi\frac{y}{x^2} \\
&= -\cos^2\phi\frac{\sin\phi}{r\cos^2\phi\sin\theta} \\
&= -\frac{\sin\phi}{r\sin\theta},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\phi}{\partial y} &= \frac{\partial\phi}{\partial\tan\phi}\frac{\partial\tan\phi}{\partial y} \\
&= \frac{1}{\sec^2\phi}\frac{\partial}{\partial y}\frac{y}{x} \\
&= \cos^2\phi\frac{1}{x} \\
&= \cos^2\phi\frac{1}{r\cos\phi\sin\theta} \\
&= \frac{\cos\phi}{r\sin\theta},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\phi}{\partial z} &= \frac{\partial\phi}{\partial\tan\phi}\frac{\partial\tan\phi}{\partial z} \\
&= \frac{1}{\sec^2\phi}\frac{\partial}{\partial z}\frac{y}{x} = 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial r}{\partial x} &= \frac{\partial\left(x^2+y^2+z^2\right)^{1/2}}{\partial x} \\
&= \frac{x}{r} \\
&= \cos\phi\sin\theta
\end{aligned}$$

$$\begin{aligned}\frac{\partial r}{\partial y} &= \frac{y}{r}\\ &= \sin\phi\sin\theta\end{aligned}$$

and

$$\begin{aligned}\frac{\partial r}{\partial z} &= \frac{z}{r}\\ &= \cos\theta.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r}+\frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta}+\frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi}\\ &= \cos\phi\sin\theta\frac{\partial}{\partial r}+\frac{\cos\phi\cos\theta}{r}\frac{\partial}{\partial\theta}-\frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r}+\frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta}+\frac{\partial\phi}{\partial y}\frac{\partial}{\partial\phi}\\ &= \sin\phi\sin\theta\frac{\partial}{\partial r}+\frac{\sin\phi\cos\theta}{r}\frac{\partial}{\partial\theta}+\frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial z} &= \frac{\partial r}{\partial z}\frac{\partial}{\partial r}+\frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta}+\frac{\partial\phi}{\partial z}\frac{\partial}{\partial\phi}\\ &= \cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}.\end{aligned}$$

Note that

$$\begin{aligned}x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z} &= r\cos\phi\sin\theta\left(\cos\phi\sin\theta\frac{\partial}{\partial r}+\frac{\cos\phi\cos\theta}{r}\frac{\partial}{\partial\theta}-\frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\\ &\quad +r\sin\phi\sin\theta\left(\sin\phi\sin\theta\frac{\partial}{\partial r}+\frac{\sin\phi\cos\theta}{r}\frac{\partial}{\partial\theta}+\frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\\ &\quad +r\cos\theta\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\\ &= r\left(\cos^2\phi+\sin^2\phi\right)\sin^2\theta\frac{\partial}{\partial r}+\left(\cos^2\phi+\sin^2\phi\right)\sin\theta\cos\theta\frac{\partial}{\partial\theta}\\ &\quad +r\cos^2\theta\frac{\partial}{\partial r}-\sin\theta\cos\theta\frac{\partial}{\partial\theta}\\ &= r\left(\sin^2\theta+\cos^2\theta\right)\frac{\partial}{\partial r}=r\frac{\partial}{\partial r}.\end{aligned}$$

In terms of momentum operators, this becomes

$$x\hat{p}_x+y\hat{p}_y+z\hat{p}_z=-i\hbar r\frac{\partial}{\partial r}=r\hat{p}_r$$

which gives

$$\begin{aligned}
r^2\hat{p}_r^2 &= (r\hat{p}_r)^2 - r\left[\hat{p}_r, r\right]\hat{p}_r \\
&= (x\hat{p}_x + y\hat{p}_y + z\hat{p}_z)^2 + i\hbar r\hat{p}_r \\
&= x\hat{p}_x x\hat{p}_x + y\hat{p}_y y\hat{p}_y + z\hat{p}_z z\hat{p}_z + 2\left(y\hat{p}_y z\hat{p}_z + z\hat{p}_z x\hat{p}_x + x\hat{p}_x y\hat{p}_y\right) + i\hbar r\hat{p}_r \\
&= x^2\hat{p}_x^2 + y^2\hat{p}_y^2 + z^2\hat{p}_z^2 + x\left[\hat{p}_x, x\right]\hat{p}_x + y\left[\hat{p}_y, y\right]\hat{p}_y + z\left[\hat{p}_z, z\right]\hat{p}_z \\
&\quad + 2\left(y\hat{p}_y z\hat{p}_z + z\hat{p}_z x\hat{p}_x + x\hat{p}_x y\hat{p}_y\right) + i\hbar r\hat{p}_r \\
&= x^2\hat{p}_x^2 + y^2\hat{p}_y^2 + z^2\hat{p}_z^2 - i\hbar\left(x\hat{p}_x + y\hat{p}_y + z\hat{p}_z\right) + 2\left(y\hat{p}_y z\hat{p}_z + z\hat{p}_z x\hat{p}_x + x\hat{p}_x y\hat{p}_y\right) + i\hbar r\hat{p}_r \\
&= x^2\hat{p}_x^2 + y^2\hat{p}_y^2 + z^2\hat{p}_z^2 + 2\left(y\hat{p}_y z\hat{p}_z + z\hat{p}_z x\hat{p}_x + x\hat{p}_x y\hat{p}_y\right).
\end{aligned}$$

Now we can write the kinetic energy operator as the sum of radial and angular kinetic energies. The angular kinetic energy - which still depends on r through a prefactor - takes the form,

$$\hat{H}_{\text{ang}} = \frac{1}{2mr^2}\hat{\mathbf{L}}^2. \tag{4.13}$$

To see this and to find the form of the radial kinetic energy, we evaluate

$$\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2.$$

The first term, expressed in terms of Cartesan relative coordinates and momenta, is given by

$$\begin{aligned}
\hat{L}_x^2 &= (y\hat{p}_z - z\hat{p}_y)^2 \\
&= y^2\hat{p}_z^2 - y\hat{p}_z z\hat{p}_y - z\hat{p}_y y\hat{p}_z + z^2\hat{p}_y^2 \\
&= y^2\hat{p}_z^2 + z^2\hat{p}_y^2 - yz\hat{p}_z\hat{p}_y - y\left[\hat{p}_z, z\right]\hat{p}_y - zy\hat{p}_y\hat{p}_z - z\left[\hat{p}_y, y\right]\hat{p}_z \\
&= y^2\hat{p}_z^2 + z^2\hat{p}_y^2 - 2yz\hat{p}_y\hat{p}_z - (-i\hbar)\, y\hat{p}_y - (-i\hbar)\, z\hat{p}_z \\
&= y^2\hat{p}_z^2 + z^2\hat{p}_y^2 - 2yz\hat{p}_y\hat{p}_z + i\hbar\left(y\hat{p}_y + z\hat{p}_z\right)
\end{aligned}$$

The next term, $\hat{L}_y^2$, (and subsequently $\hat{L}_z^2$) is obtained from this expression by replacing y by z and z by x (and x by y, in the subsequent step). Specifically,

$$\begin{aligned}
\hat{L}_y^2 &= (z\hat{p}_x - x\hat{p}_z)^2 \\
&= z^2\hat{p}_x^2 + x^2\hat{p}_z^2 - 2zx\hat{p}_z\hat{p}_x + i\hbar\left(z\hat{p}_z + x\hat{p}_x\right).
\end{aligned}$$

and

$$\begin{aligned}
\hat{L}_z^2 &= (x\hat{p}_y - y\hat{p}_x)^2 \\
&= x^2\hat{p}_y^2 + y^2\hat{p}_x^2 - 2xy\hat{p}_x\hat{p}_y + i\hbar\left(x\hat{p}_x + y\hat{p}_y\right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\hat{\mathbf{L}}^2 &= \left(x^2 + y^2\right)\hat{p}_z^2 + \left(y^2 + z^2\right)\hat{p}_x^2 + \left(z^2 + x^2\right)\hat{p}_y^2 \\
&\quad -2\left(yz\hat{p}_y\hat{p}_z + zx\hat{p}_z\hat{p}_x + xy\hat{p}_x\hat{p}_y\right) + 2i\hbar\left(x\hat{p}_x + y\hat{p}_y + z\hat{p}_z\right) \\
&= r^2\left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right) \\
&\quad - \left(x^2\hat{p}_x^2 + y^2\hat{p}_y^2 + z^2\hat{p}_z^2 + 2\left(yz\hat{p}_y\hat{p}_z + zx\hat{p}_z\hat{p}_x + xy\hat{p}_x\hat{p}_y\right)\right) + 2i\hbar r\hat{p}_r \\
&= r^2\hat{\mathbf{p}}^2 - r^2\hat{p}_r^2 + 2i\hbar r\hat{p}_r,
\end{aligned}$$

and the relative kinetic energy is given by

$$\begin{aligned}\frac{1}{2m}\hat{\mathbf{p}}^2 &= \frac{1}{2mr^2}\left(\hat{\mathbf{L}}^2 + r^2\hat{p}_r^2 - 2i\hbar r\hat{p}_r\right)\\ &= \frac{1}{2mr^2}\hat{\mathbf{L}}^2 + \frac{1}{2m}\left[\left(-i\hbar\frac{\partial}{\partial r}\right)^2 + \frac{(-i\hbar)^2}{r}\frac{\partial}{\partial r}\right]\\ &= \hat{H}_{\mathrm{ang}} - \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right). \qquad (4.14)\end{aligned}$$

It is the sum of *angular* and *radial* parts.

Chapter 5

Rotations of a diatomic molecule

5.1 Angular Momentum Eigenfunctions

In Chapter 4, we showed that the energy eigenfunctions for the relative motion of a diatomic molecule have the form,

$$\psi\left(r,\theta,\phi\right)=\frac{1}{r}\varphi\left(r\right)Y\left(\theta,\phi\right),$$

where $\varphi\left(r_{\text{eq}}+x\right)$ is a harmonic oscillator energy eigenfunction of x, and $Y\left(\theta,\phi\right)$ is an eigenfunction of $\hat{\mathbf{L}}^2$;

$$\hat{\mathbf{L}}^2 Y\left(\theta,\phi\right)=L^2 Y\left(\theta,\phi\right). \tag{5.1}$$

To make this equation, we need $\hat{\mathbf{L}}^2$ in spherical coordinates. The components of angular momentum are given as follows:

$$\begin{aligned}
\hat{L}_x &= y\hat{p}_z - z\hat{p}_y \\
&= -i\hbar\left[r\sin\theta\sin\phi\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\right. \\
&= \left.-r\cos\theta\left(\sin\phi\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta}+\frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\right] \\
&= -i\hbar\left(\left(-\sin^2\theta\sin\phi-\cos^2\theta\sin\phi\right)\frac{\partial}{\partial\theta}-\frac{\cos\theta\cos\phi}{\sin\theta}\frac{\partial}{\partial\phi}\right) \\
&= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta}-\cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \qquad \sin^2\theta+\cos^2\theta=1
\end{aligned}$$

$$
\begin{aligned}
\hat{L}_y &= z\hat{p}_x - x\hat{p}_z \\
&= -i\hbar\left[r\cos\theta\left(\cos\phi\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\right. \\
&\qquad \left. - r\sin\theta\cos\phi\left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\right] \\
&= -i\hbar\left(\cos^2\theta\cos\phi\frac{\partial}{\partial\theta} + \sin^2\theta\cos\phi\frac{\partial}{\partial\theta} - \frac{\cos\theta\sin\phi}{\sin\theta}\frac{\partial}{\partial\phi}\right) \\
&= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{L}_z &= x\hat{p}_y - y\hat{p}_x \\
&= -i\hbar\left[r\sin\theta\cos\phi\left(\sin\phi\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\right. \\
&\qquad \left. - r\sin\theta\sin\phi\left(\cos\phi\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\right] \\
&= -i\hbar\left(\cos^2\phi + \sin^2\phi\right)\frac{\partial}{\partial\phi} \\
&= -i\hbar\frac{\partial}{\partial\phi}.
\end{aligned}
$$

The z component of angular momentum is simplest because, in spherical coordinates, rotation about the z axis corresponds to simply advancing ϕ. Spherical coordinates could have been defined differently, if we wished for $\hat{L}_x$ or $\hat{L}_y$ to have the simplest form.

We actually need the square angular momentum components. $\hat{L}_z^2$ is easy;

$$
\hat{L}_z^2 = -\hbar^2\frac{\partial^2}{\partial\phi^2}.
$$

For the other two components, we only need $\hat{L}_x^2 + \hat{L}_y^2$. So, we write

$$
\begin{aligned}
\hat{L}_x^2 &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)\left(-i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)\right) \\
&= -\hbar^2\left(\sin^2\phi\frac{\partial^2}{\partial\theta^2} + \sin\phi\cos\phi\frac{\partial}{\partial\theta}\cot\theta\frac{\partial}{\partial\phi} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\theta} + \cot^2\theta\cos\phi\frac{\partial}{\partial\phi}\cos\phi\frac{\partial}{\partial\phi}\right) \\
&= -\hbar^2\left(\begin{array}{c} \sin^2\phi\tfrac{\partial^2}{\partial\theta^2} + \sin\phi\cos\phi\tfrac{\partial}{\partial\theta}\cot\theta\tfrac{\partial}{\partial\phi} + \cot\theta\cos\phi\left(\cos\phi\tfrac{\partial}{\partial\theta} + \sin\phi\tfrac{\partial^2}{\partial\phi\partial\theta}\right) \\ + \cot^2\theta\cos\phi\tfrac{\partial}{\partial\phi}\cos\phi\tfrac{\partial}{\partial\phi} \end{array}\right)
\end{aligned}
$$

$$\begin{aligned}
\hat{L}_y^2 &= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta}-\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\left(-i\hbar\left(\cos\phi\frac{\partial}{\partial\theta}-\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\right)\\
&= -\hbar^2\left(\begin{array}{c}\cos^2\phi\frac{\partial^2}{\partial\theta^2}-\sin\phi\cos\phi\frac{\partial}{\partial\theta}\cot\theta\frac{\partial}{\partial\phi}-\cot\theta\sin\phi\frac{\partial}{\partial\phi}\cos\phi\frac{\partial}{\partial\theta}\\ +\cot^2\theta\sin\phi\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi}\end{array}\right)\\
&= -\hbar^2\left(\begin{array}{c}\cos^2\phi\frac{\partial^2}{\partial\theta^2}-\sin\phi\cos\phi\frac{\partial}{\partial\theta}\cot\theta\frac{\partial}{\partial\phi}-\cot\theta\sin\phi\left(-\sin\phi\frac{\partial}{\partial\theta}+\cos\phi\frac{\partial^2}{\partial\phi\partial\theta}\right)\\ +\cot^2\theta\sin\phi\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi}\end{array}\right)
\end{aligned}$$

and note that when these expressions are added, the second and fourth terms (with inner parentheses expanded) in the two expressions cancel. The remaining terms take the form,

$$\begin{aligned}
\hat{L}_x^2+\hat{L}_y^2 &= -\hbar^2\left[\left(\sin^2\phi+\cos^2\phi\right)\frac{\partial^2}{\partial\theta^2}+\cot\theta\left(\sin^2\phi+\cos^2\phi\right)\frac{\partial}{\partial\theta}+\cot^2\theta\left(\sin^2\phi+\cos^2\phi\right)\frac{\partial^2}{\partial\phi^2}\right]\\
&= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\cot^2\theta\frac{\partial^2}{\partial\phi^2}\right)\\
&= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)
\end{aligned}$$

In principle, we would like to find the simultaneous eigenfunctions of all three components of $\hat{\mathbf{L}}$. However, this is not possible because the components of $\hat{\mathbf{L}}$ do not commute. From $[x,\hat{p}_x]=i\hbar$ (and similarly for y and z), and components of $\hat{\mathbf{L}}$ expressed in terms of Cartesian coordinates and momenta (see Sec. 2.4.2),

$$\left[\hat{L}_x,\hat{L}_y\right]=i\hbar\hat{L}_z,$$

$$\left[\hat{L}_y,\hat{L}_z\right]=i\hbar\hat{L}_x$$

and

$$\left[\hat{L}_z,\hat{L}_x\right]=i\hbar\hat{L}_y.$$

However, each component of $\hat{\mathbf{L}}$ commutes with $\hat{\mathbf{L}}^2$. This can be shown using the above commutators (also note that, in spherical coordinates, $\hat{L}_z$ clearly commutes with $\hat{L}_x^2+\hat{L}_y^2$) - see Example 5.1. Thus, it is possible to find functions of θ and ϕ that are simultaneous eigenfunctions of $\hat{\mathbf{L}}^2$ and one component of $\hat{\mathbf{L}}$. We choose that component to be $\hat{L}_z$, as spherical coordinates are already convenient for $\hat{L}_z$. Specifically, we seek

$$Y(\theta,\phi)=\Theta(\theta)\,\Phi(\phi)$$

such that

$$\hat{\mathbf{L}}^2Y(\theta,\phi)=L^2Y(\theta,\phi)$$

and

$$\hat{L}_zY(\theta,\phi)=L_zY(\theta,\phi).$$

The latter equation is simplified by dividing by $\Theta(\theta)$. Since $\hat{L}_z$ affects only functions of ϕ, $\Theta(\theta)$ cancels out on both sides. The equation takes the form,

$$\hat{L}_z \Phi(\phi) = L_z \Phi(\phi).$$

This equation is easy to solve. Explicitly,

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

or

$$\frac{\partial}{\partial\phi} \Phi(\phi) = \frac{iL_z}{\hbar} \Phi(\phi)$$

which has the solution,

$$\Phi(\phi) = A \exp\left(\frac{iL_z\phi}{\hbar}\right).$$

The azimuthal angle, ϕ, varies from 0 to 2π. For values beyond 2π, the same points are repeated - ϕ is equivalent to $\phi - 2\pi$. Consequently, only wavefunctions that are periodic in ϕ, with period 2π, correspond to physically admissable wavefunctions. Applying this periodic boundary condition, we must have

$$\Phi(\phi + 2\pi) = \Phi(\phi)$$

or

$$\begin{aligned} A \exp\left(\frac{iL_z\phi}{\hbar} + \frac{2\pi i L_z}{\hbar}\right) &= A \exp\left(\frac{iL_z\phi}{\hbar}\right) \\ \exp\left(\frac{2\pi i L_z}{\hbar}\right) &= 1. \end{aligned}$$

This condition is satisfied if

$$\frac{L_z}{\hbar} = m_\ell, \qquad \text{where } m_\ell \text{ is any integer}$$

where m_ℓ is the L_z quantum number (the orbital magnetic quantum number when $\hat{\mathbf{L}}$ is the orbital angular momentum of an electron).

The eigenfunctions of $\hat{L}_z$ are the functions,

$$\begin{aligned} \Phi_{m_\ell}(\phi) &= A \exp(i m_\ell \phi) \qquad (5.2) \\ &= \frac{1}{\sqrt{2\pi}} \exp(i m_\ell \phi) \end{aligned}$$

labeled by integer quantum number, m_ℓ. The corresponding eigenvalue is

$$L_z = m_\ell \hbar. \qquad (5.3)$$

Here we see that $\hbar$ is the quantum of angular momentum - L_z changes only in increments of $\hbar$. In the second line of Eq. 5.2 , the normalization constant is made explicit. It follows from setting $\int_0^{2\pi} \left|\Phi_{m_\ell}(\phi)\right|^2 d\phi = 1$ - see the Appendix, Sec. 5.5.

Equation 5.1 has the following explicit form, in spherical coordinates:

$$\begin{aligned}
\left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right) Y(\theta,\phi) &= L^2 Y(\theta,\phi) \\
-\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial\phi^2}\right)\Theta(\theta)\Phi(\phi) &= L^2\Theta(\theta)\Phi(\phi) \\
-\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta} + \left(\frac{1}{\sin^2\theta} + 1\right)\frac{\partial^2}{\partial\phi^2}\right)\Theta(\theta)\Phi_{m_\ell}(\phi) &= L^2\Theta(\theta)\Phi_{m_\ell}(\phi) \\
-\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta} + \left(\frac{1}{\sin^2\theta} + 1\right)\left(-m_\ell^2\right)\right)\Theta(\theta)\Phi_{m_\ell}(\phi) &= L^2\Theta(\theta)\Phi_{m_\ell}(\phi) \\
-\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta} - m_\ell^2\left(\frac{1}{\sin^2\theta} + 1\right)\right)\Theta(\theta) &= L^2\Theta(\theta).
\end{aligned}$$

In the last step, we divide out $\Phi_{m_\ell}(\phi)$ to get an equation for $\Theta(\theta)$. This equation depends solely on θ, and completes the separation of the six dimensional TISE of two particles in three dimensional space into six TISE's in one dimension. The twist here is that the TISE for $\Theta(\theta)$ depends on m_ℓ, the quantum number associated with the L_z eigenvalue equation for $\Phi(\phi)$. We will solve the above θ equation using ladder operators for the quantum number m_ℓ.

Let

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y$$

and

$$\hat{L}_- = \hat{L}_x - i\hat{L}_y.$$

Consider the commutators of $\hat{L}_+$ and $\hat{L}_-$ with $\hat{L}_z$,

$$\begin{aligned}
\left[\hat{L}_z, \hat{L}_+\right] &= \left[\hat{L}_z, \hat{L}_x\right] + i\left[\hat{L}_z, \hat{L}_y\right] \\
&= i\hbar\hat{L}_y + i\left(-i\hbar\hat{L}_x\right) \\
&= \hbar\left(\hat{L}_x + i\hat{L}_y\right) \\
&= \hbar\hat{L}_+
\end{aligned}$$

and

$$\begin{aligned}
\left[\hat{L}_z, \hat{L}_-\right] &= \left[\hat{L}_z, \hat{L}_x\right] - i\left[\hat{L}_z, \hat{L}_y\right] \\
&= i\hbar\hat{L}_y - i\left(-i\hbar\hat{L}_x\right) \\
&= -\hbar\left(\hat{L}_x - i\hat{L}_y\right) \\
&= -\hbar\hat{L}_-.
\end{aligned}$$

From the first commutator, we get

$$\begin{aligned}\hat{L}_z\hat{L}_+\Theta(\theta)\Phi_{m_\ell}(\phi) &= \left[\hat{L}_z,\hat{L}_+\right]\Theta(\theta)\Phi_{m_\ell}(\phi)+\hat{L}_+\hat{L}_z\Theta(\theta)\Phi_{m_\ell}(\phi)\\ &= \hbar\hat{L}_+\Theta(\theta)\Phi_{m_\ell}(\phi)+\hat{L}_+m_\ell\hbar\Theta(\theta)\Phi_{m_\ell}(\phi)\\ &= (m_\ell+1)\hbar\hat{L}_+\Theta(\theta)\Phi_{m_\ell}(\phi).\end{aligned}$$

This means that $\hat{L}_+\Theta(\theta)\Phi_{m_\ell}(\phi)$ is an eigenfunction of $\hat{L}_z$ associated with quantum number, $m_\ell+1$. $\hat{L}_+$ is a raising operator. Specifically, $\hat{L}_+\Theta(\theta)\Phi_{m_\ell}(\phi)$ is proportional to $\Phi_{m_\ell+1}(\phi)$. Similarly,

$$\begin{aligned}\hat{L}_z\hat{L}_-\Theta(\theta)\Phi_{m_\ell}(\phi) &= \left[\hat{L}_z,\hat{L}_-\right]\Theta(\theta)\Phi_{m_\ell}(\phi)+\hat{L}_-\hat{L}_z\Theta(\theta)\Phi_{m_\ell}(\phi)\\ &= -\hbar\hat{L}_-\Theta(\theta)\Phi_{m_\ell}(\phi)+\hat{L}_-m_\ell\hbar\Theta(\theta)\Phi_{m_\ell}(\phi)\\ &= (m_\ell-1)\hbar\hat{L}_-\Theta(\theta)\Phi_{m_\ell}(\phi).\end{aligned}$$

$\hat{L}_-$ is a lowering operator which lowers the m_ℓ quantum number by 1.

The trick to finding the other quantum number - the quantum number associated with $\hat{\mathbf{L}}^2$ - is to realize that L_z^2 cannot exceed L^2. Since L^2 is fixed, there must be a maximum and minimum value for m_ℓ, which we denote by $\pm\ell$. Specifically,

$$-\ell \le m_\ell \le \ell.$$

ℓ is the remaining quantum number. As such, we label Θ by ℓ and m_ℓ (it depends on both of these quantum numbers), and write the simultaneous eigenfunctions of $\hat{\mathbf{L}}^2$ and L_z as

$$Y_{\ell,m_\ell}(\theta,\phi)=\Theta_{\ell,m_\ell}(\theta)\Phi_{m_\ell}(\phi),$$

$\ell=0,1,2,\ldots$ and $m_\ell=-\ell,\ldots,\ell$. The ℓ quantum number is related to the total angular momentum.

In spherical coordinates, the raising and lowering operators take the form,

$$\begin{aligned}\hat{L}_+ &= \hat{L}_x+i\hat{L}_y\\ &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta}-\cot\theta\cos\phi\frac{\partial}{\partial\phi}+i\left(\cos\phi\frac{\partial}{\partial\theta}-\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\right)\\ &= -i\hbar\left(i(\cos\phi+i\sin\phi)\frac{\partial}{\partial\theta}-\cot\theta(\cos\phi+i\sin\phi)\frac{\partial}{\partial\phi}\right)\\ &= \hbar\exp(i\phi)\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)\end{aligned}$$

and

$$\begin{aligned}
\hat{L}_- &= \hat{L}_x - i\hat{L}_y \\
&= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi} - i\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\right) \\
&= -i\hbar\left(i\left(-\cos\phi + i\sin\phi\right)\frac{\partial}{\partial\theta} - \cot\theta\left(\cos\phi - i\sin\phi\right)\frac{\partial}{\partial\phi}\right) \\
&= \hbar\exp\left(-i\phi\right)\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right).
\end{aligned}$$

Since there is a top state - the state for which $m_\ell = \ell$ - it must be the solution to the equation,

$$\hat{L}_+\Theta_{\ell,\ell}\left(\theta\right)\Phi_\ell\left(\phi\right) = 0. \tag{5.4}$$

Otherwise, $\hat{L}_+$ would further raise m_ℓ to $\ell+1$, and no such state exists. In spherical coordinates, Eq. 5.4 takes the form,

$$\hbar\exp\left(i\phi\right)\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\Theta_{\ell,\ell}\left(\theta\right)\Phi_\ell\left(\phi\right) = 0$$

or

$$\begin{aligned}
\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\Theta_{\ell,\ell}\left(\theta\right)\Phi_\ell\left(\phi\right) &= 0 \\
\left(\frac{\partial}{\partial\theta} + i\cot\theta\left(i\ell\right)\right)\Theta_{\ell,\ell}\left(\theta\right)\Phi_\ell\left(\phi\right) &= 0 \\
\left(\frac{d}{d\theta} - \ell\cot\theta\right)\Theta_{\ell,\ell}\left(\theta\right) &= 0. \qquad \text{divide by } \Phi_\ell\left(\phi\right)
\end{aligned}$$

This equation is solved as follows:

$$\begin{aligned}
\frac{d\Theta}{d\theta} &= \ell\cot\theta\Theta \\
\frac{d\Theta}{\Theta} &= \ell\cot\theta d\theta \\
\int\frac{d\Theta}{\Theta} &= \ell\int\cot\theta d\theta \\
\ln\Theta &= \ell\int\frac{\cos\theta}{\sin\theta}d\theta \\
&= \ell\int\frac{1}{u}du \qquad u = \sin\theta \\
&= \ell\ln\left(\sin\theta\right) + \text{constant} \\
&= \ln\left(\sin^\ell\theta\right) + \text{constant}
\end{aligned}$$

or

$$\Theta\left(\theta\right) = \Theta_{\ell,\ell}\left(\theta\right) = B_\ell\sin^\ell\theta.$$

In the Appendix, Sec. 5.5, the normalization condition for $\Theta(\theta)$ is shown to be

$$\begin{aligned}\int_0^\pi \Theta^2(\theta)\sin\theta d\theta &= 1\\ B_\ell^2\int_0^\pi \sin^{2\ell+1}\theta d\theta &= 1\end{aligned}$$

or

$$B_\ell = I_\ell^{-1/2},$$

where

$$\begin{aligned}I_\ell &= \int_0^\pi \sin^{2\ell+1}\theta d\theta = \int_0^\pi \sin^{2\ell}\theta\sin\theta d\theta\\ &= -\int_0^\pi \left(\sin^2\theta\right)^\ell d\cos\theta\\ &= \int_{-1}^1 \left(1-x^2\right)^\ell dx\\ &= \frac{\ell(\ell-1)\cdots 1}{\left(\ell+\frac{1}{2}\right)\left(\ell-\frac{1}{2}\right)\cdots\frac{3}{2}\times\frac{1}{2}}\\ &= \frac{2^{\ell+1}\ell!}{(2\ell+1)(2\ell-1)\cdots 3\times 1}\\ &= \frac{2^{\ell+1}\ell!}{(2\ell+1)!!}\\ &= 2,\frac{4}{3},\frac{16}{15},\frac{32}{35},\qquad \text{for } \ell = 0,1,2,3,\ldots\end{aligned}$$

Once we have the top state, we can obtain any other state (associated with the same value of ℓ) by applying the lowering operator, $\hat{L}_-$. However, we still need to express the eigenvalues of $\hat{\mathbf{L}}^2$ in terms of quantum number, ℓ. To that end, note that

$$\begin{aligned}\hat{L}_-\hat{L}_+ &= \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right)\\ &= \hat{L}_x^2 + \hat{L}_y^2 + i\hat{L}_x\hat{L}_y - i\hat{L}_y\hat{L}_x\\ &= \hat{\mathbf{L}}^2 - \hat{L}_z^2 + i\left[\hat{L}_x,\hat{L}_y\right]\\ &= \hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z.\end{aligned}$$

Therefore,

$$\hat{\mathbf{L}}^2 = \hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z$$

and

$$\begin{aligned}\hat{\mathbf{L}}^2 Y_{\ell,\ell}(\theta,\phi) &= \left(\hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z\right)Y_{\ell,\ell}(\theta,\phi)\\ &= \hat{L}_-\hat{L}_+ Y_{\ell,\ell}(\theta,\phi) + \hat{L}_z^2 Y_{\ell,\ell}(\theta,\phi) + \hbar\hat{L}_z Y_{\ell,\ell}(\theta,\phi)\\ &= \left(0 + (\ell\hbar)^2 + \hbar(\ell\hbar)\right)Y_{\ell,\ell}(\theta,\phi)\\ &= \ell(\ell+1)\hbar^2 Y_{\ell,\ell}(\theta,\phi).\end{aligned}$$

The eigenvalues of $\hat{\mathbf{L}}^2$ are $\ell(\ell+1)\hbar^2$ where $\ell = 0, 1, 2, \ldots$. There are $2\ell + 1$ states with this eigenvalue. Since $\left[\hat{L}_-, \hat{\mathbf{L}}^2\right] = 0$, $\hat{L}_-$ lowers the m_ℓ quantum number without affecting ℓ. Altogether,

$$\hat{\mathbf{L}}^2 Y_{\ell,m_\ell}(\theta,\phi) = \ell(\ell+1)\hbar^2 Y_{\ell,m_\ell}(\theta,\phi), \qquad \ell = 0, 1, 2, \ldots \tag{5.5}$$

and

$$\hat{L}_z Y_{\ell,m_\ell}(\theta,\phi) = m_\ell \hbar Y_{\ell,m_\ell}(\theta,\phi) \qquad -\ell \le m_\ell \le \ell. \tag{5.6}$$

The eigenvalues, $L^2 = \ell(\ell+1)\hbar^2$, are *degenerate* when $\ell > 0$. The $Y_{\ell,m_\ell}(\theta,\phi)$ are called the *spherical harmonics*. They are products of complex exponential functions (in ϕ) and (associated) Legendre polynomials (in θ).

We now get explicit expressions for the actions of $\hat{L}_+$ and $\hat{L}_-$. Specifically, we determine the coefficients in the following equations:

$$\hat{L}_- Y_{\ell,m_\ell} = C_{-;\ell,m_\ell-1} Y_{\ell,m_\ell-1}$$

and

$$\hat{L}_+ Y_{\ell,m_\ell} = C_{+;\ell,m_\ell} Y_{\ell,m_\ell+1}.$$

The lowering operator is the Hermitian conjugate of the raising operator - check;

$$\hat{L}_- = \hat{L}_+^\dagger.$$

Consequently,

$$\begin{aligned}
\left\langle \hat{L}_- Y_{\ell,m_{\ell+1}} \right| \left. Y_{\ell,m_\ell} \right\rangle &= \langle C_{-;\ell,m_\ell} Y_{\ell,m_\ell} | Y_{\ell,m_\ell} \rangle \\
&= C^*_{-;\ell,m_\ell} \langle Y_{\ell,m_\ell} | Y_{\ell,m_\ell} \rangle \\
&= C^*_{-;\ell,m_\ell} \qquad Y_{\ell,m_\ell} \text{ is normalized} \\
&= \left\langle Y_{\ell,m_{\ell+1}} \left| \hat{L}_-^\dagger Y_{\ell,m_\ell} \right.\right\rangle \\
&= \left\langle Y_{\ell,m_{\ell+1}} \left| \hat{L}_+ Y_{\ell,m_\ell} \right.\right\rangle \\
&= \langle Y_{\ell,m_\ell+1} | C_{+;\ell,m_\ell} Y_{\ell,m_\ell+1} \rangle \\
&= C_{+;\ell,m_\ell}. \qquad Y_{\ell,m_\ell+1} \text{ is normalized}
\end{aligned}$$

Also,

$$\begin{aligned}
\left\langle \hat{L}_+ Y_{\ell,m_\ell} \left| \hat{L}_+ Y_{\ell,m_\ell} \right.\right\rangle &= \langle C_{+;\ell,m_\ell} Y_{\ell,m_\ell+1} | C_{+;\ell,m_\ell} Y_{\ell,m_\ell+1} \rangle \\
&= \left|C_{+;\ell,m_\ell}\right|^2 \langle Y_{\ell,m_\ell+1} | Y_{\ell,m_\ell+1} \rangle = \left|C_{+;\ell,m_\ell}\right|^2 \qquad Y_{\ell,m_\ell+1} \text{ is normalized} \\
&= \left\langle Y_{\ell,m_\ell} \left| \hat{L}_- \hat{L}_+ Y_{\ell,m_\ell} \right.\right\rangle \qquad \hat{L}_+^\dagger = \hat{L}_- \\
&= \left\langle Y_{\ell,m_\ell} \left| \left(\hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar \hat{L}_z\right) Y_{\ell,m_\ell} \right.\right\rangle \\
&= \left\langle Y_{\ell,m_\ell} \left| \left(\ell(\ell+1)\hbar^2 - m_\ell^2 \hbar^2 - m_\ell \hbar^2\right) Y_{\ell,m_\ell} \right.\right\rangle \\
&= \left[\ell(\ell+1) - m_\ell(m_\ell+1)\right]\hbar^2 \langle Y_{\ell,m_\ell} | Y_{\ell,m_\ell} \rangle \\
&= \left[\ell(\ell+1) - m_\ell(m_\ell+1)\right]\hbar^2. \qquad Y_{\ell,m_\ell} \text{ is normalized}
\end{aligned}$$

Thus,

$$|C_{+;\ell,m_\ell}| = \left[\ell(\ell+1) - m_\ell(m_\ell+1)\right]^{1/2}\hbar$$

and

$$C^*_{-;\ell,m_\ell} = C_{+;\ell,m_\ell}.$$

We let $C_{+;\ell,m_\ell} = C_{\ell,m_\ell}$ be real, but with an additional $(-1)^\ell$ factor - i.e.,

$$C_{\ell,m_\ell} = (-1)^\ell\left[\ell(\ell+1) - m_\ell(m_\ell+1)\right]^{1/2}\hbar.$$

$C_{-;\ell,m_\ell} = C_{\ell,m_\ell}$ because C_{ℓ,m_ℓ} is real.

$$\begin{aligned}\hat{L}_-Y_{\ell,m_\ell} &= C_{\ell,m_\ell-1}Y_{\ell,m_\ell-1}\\ &= (-1)^\ell\left[\ell(\ell+1) - (m_\ell-1)m_\ell\right]^{1/2}\hbar Y_{\ell,m_\ell-1}\end{aligned}$$

and

$$\begin{aligned}\hat{L}_+Y_{\ell,m_\ell} &= C_{\ell,m_\ell}Y_{\ell,m_\ell+1}\\ &= (-1)^\ell\left[\ell(\ell+1) - m_\ell(m_\ell+1)\right]^{1/2}\hbar Y_{\ell,m_\ell+1}.\end{aligned}$$

The s state ($\ell = 0$) is the top state and bottom state of its ladder of states;

$$\begin{aligned}Y_{0,0}(\theta,\phi) &= \Theta_{0,0}(\theta)\,\Phi_0(\phi) = (2\pi I_0)^{-1/2}\\ &= (4\pi)^{-1/2}.\end{aligned}$$

It is simply a constant. Since $|Y_{0,0}|^2 = (4\pi)^{-1}$ is constant, this is an isotropic state with no prefered direction.

The p states ($\ell = 1$) are obtained as follows: The top state is

$$\begin{aligned}Y_{1,1}(\theta,\phi) &= \Theta_{1,1}(\theta)\,\Phi_1(\phi) = (2\pi I_1)^{-1/2}\sin\theta\exp(i\phi)\\ &= \left(\frac{8\pi}{3}\right)^{-1/2}\sin\theta\exp(i\phi)\end{aligned}$$

Here, $|Y_{1,1}|^2 = (8\pi/3)^{-1}\sin^2\theta$ is largest for $\theta = \pi/2$; i.e., in the $x\ y$ plane. There are nodes at the poles - $\theta = 0$ or π. Applying the lowering operator gives $Y_{1,0}$. Specifically,

$$\begin{aligned}C_{1,0}Y_{1,0} &= \hat{L}_-Y_{1,1}\\ (-1)^1\left[1(1+1) - 0(0+1)\right]^{1/2}\hbar Y_{1,0} &= \hbar\exp(-i\phi)\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)(2\pi I_1)^{-1/2}\sin\theta\exp(i\phi)\\ (2)^{1/2}Y_{1,0} &= -(2\pi I_1)^{-1/2}\exp(-i\phi)\left(-\frac{\partial}{\partial\theta}\sin\theta + (i\cos\theta)(i)\right)\exp(i\phi)\end{aligned}$$

or

$$\begin{aligned}Y_{1,0} &= (4\pi I_1)^{-1/2}\exp(-i\phi)(\cos\theta + \cos\theta)\exp(i\phi)\\ &= (4\pi I_1)^{-1/2}\,2\cos\theta\\ &= \left(\frac{4\pi}{3}\right)^{-1/2}\cos\theta\end{aligned}$$

The introduction of the $(-1)^{\ell}$ factor in the definition of $Y_{1,0}$ makes it a positive function. This state is concentrated at the poles - i.e., along the z axis - with a node in the $x\ y$ plane. The bottom state is given by another application of the lowering operator;

$$\begin{aligned} C_{1,-1} Y_{1,-1} &= \hat{L}_{-} Y_{1,0} \\ (-1)^{1}\left[1(1+1)-(-1)(-1+1)\right]^{1/2} \hbar Y_{1,-1} &= \hbar \exp(-i\phi)\left(-\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)(4\pi I_1)^{-1/2} 2\cos\theta \\ (2)^{1/2} Y_{1,-1} &= -(4\pi I_1)^{-1/2} 2\exp(-i\phi)\left(-\frac{\partial}{\partial\theta}\cos\theta+0\right) \end{aligned}$$

or

$$\begin{aligned} Y_{1,-1}(\theta,\phi) &= -(2\pi I_1)^{-1/2}\sin\theta\exp(-i\phi) \\ &= -\left(\frac{8\pi}{3}\right)^{-1/2}\sin\theta\exp(-i\phi) \end{aligned}$$

This state is concentrated in the $x\ y$ plane, with the same probability density as $Y_{1,1}$ - i.e., $|Y_{1,-1}|^2 = |Y_{1,1}|^2$. The bottom state has the the same $\Theta(\theta)$ function as the top state, except for the sign. In general,

$$Y_{\ell,-m_\ell}(\theta,\phi) = (-1)^{m_\ell} Y^{*}_{\ell,m_\ell}(\theta,\phi).$$

It is therefore sufficient to determine the spherical harmonics for $0 \le m_\ell \le \ell$.

It is possible to simplify the computation of the spherical harmonics by defining the lowering operator for the $\Theta_{\ell,m_\ell}(\theta)$ functions,

$$\hat{b}_{-;\ \ell,m_\ell}\Theta_{\ell,m_\ell}(\theta) = \Theta_{\ell,m_\ell-1}(\theta).$$

Specifically, since

$$(-1)^{\ell}\left[\ell(\ell+1)-(m_\ell-1)m_\ell\right]^{1/2}\hbar Y_{\ell,m_\ell-1} = \hbar\exp(-i\phi)\left(-\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)Y_{\ell,m_\ell},$$

$$\begin{aligned} &\Theta_{\ell,m_\ell-1}(\theta)(2\pi)^{-1/2}\exp(i(m_\ell-1)\phi) \\ = \quad &(-1)^{\ell}\left[\ell(\ell+1)-(m_\ell-1)m_\ell\right]^{-1/2}\left(-\frac{\partial}{\partial\theta}-m_\ell\cot\theta\right)\Theta_{\ell,m_\ell}(\theta)(2\pi)^{-1/2}\exp(i(m_\ell-1)\phi) \end{aligned}$$

and

$$\hat{b}_{-;\ \ell,m_\ell} = (-1)^{\ell+1}\left[\ell(\ell+1)-(m_\ell-1)m_\ell\right]^{-1/2}\left(\frac{d}{d\theta}+m_\ell\cot\theta\right).$$

The raising operator for $\Theta_{\ell,m_\ell}(\theta)$ is similarly defined;

$$\hat{b}_{+;\ \ell,m_\ell}\Theta_{\ell,m_\ell}(\theta) = \Theta_{\ell,m_\ell+1}(\theta),$$

$$\hat{b}_{+;\ \ell,m_\ell} = (-1)^{\ell+1}\left[\ell(\ell+1)-m_\ell(m_\ell+1)\right]^{-1/2}\left(\frac{d}{d\theta}-m_\ell\cot\theta\right).$$

Also,

$$\Theta_{\ell,-m_\ell}(\theta) = (-1)^{m_\ell}\Theta_{\ell,m_\ell}(\theta).$$

Example 5.1. Determine the five $\ell = 2$ spherical harmonics, $Y_{2,m_\ell}(\theta,\phi)$, $m_\ell = -2,\ldots,2$.

Solution 5.1. Start with the top Θ_{2,m_ℓ} function.

$$\begin{aligned}\Theta_{2,2}(\theta) &= I_2^{-1/2}\sin^2\theta \\ &= \left(\frac{16}{15}\right)^{-1/2}\sin^2\theta.\end{aligned}$$

Apply the lowering operator - we use the operator specific to the Θ_{ℓ,m_ℓ} to simplify the calculation - to get $\Theta_{2,1}$.

$$\begin{aligned}\Theta_{2,1}(\theta) &= \hat{b}_{-;\,2,2}\Theta_{2,2}(\theta) \\ &= (-1)^3\left[2\times 3 - 1\times 2\right]^{-1/2}\left(\frac{d}{d\theta} + 2\cot\theta\right)\left(\frac{16}{15}\right)^{-1/2}\sin^2\theta \\ &= -\frac{1}{2}\left(\frac{16}{15}\right)^{-1/2}(2\cos\theta\sin\theta + 2\cos\theta\sin\theta) \\ &= -2\left(\frac{16}{15}\right)^{-1/2}\sin\theta\cos\theta \\ &= -\left(\frac{4}{15}\right)^{-1/2}\sin\theta\cos\theta.\end{aligned}$$

Another application of the lowering operator gives,

$$\begin{aligned}\Theta_{2,0}(\theta) &= \hat{b}_{-;\,2,1}\Theta_{2,1}(\theta) \\ &= \left[2\times 3 - 0\times 1\right]^{-1/2}\left(\frac{d}{d\theta} + \cot\theta\right)\left(\frac{4}{15}\right)^{-1/2}\sin\theta\cos\theta \\ &= \left(\frac{6\times 4}{15}\right)^{-1/2}\left(\cos^2\theta - \sin^2\theta + \cos^2\theta\right) \\ &= \left(\frac{8}{5}\right)^{-1/2}\left(3\cos^2\theta - 1\right).\end{aligned}$$

Consequently,

$$Y_{2,2}(\theta,\phi) = \left(\frac{32\pi}{15}\right)^{-1/2}\sin^2\theta\exp(2i\phi),$$

$$Y_{2,1}(\theta,\phi) = -\left(\frac{8\pi}{15}\right)^{-1/2}\sin\theta\cos\theta\exp(i\phi),$$

$$Y_{2,0}(\theta,\phi) = \left(\frac{16\pi}{5}\right)^{-1/2}\left(3\cos^2\theta - 1\right),$$

$$Y_{2,-1}(\theta,\phi) = \left(\frac{8\pi}{15}\right)^{-1/2}\sin\theta\cos\theta\exp(-i\phi)$$

and

$$Y_{2,-2}(\theta,\phi) = \left(\frac{32\pi}{15}\right)^{-1/2} \sin^2\theta \exp(-2i\phi).$$

5.2 The rigid rotor

If a diatomic molecule is in its vibrational ground state (most diatomics at ordinary temperature), it is essentially a *rigid rotor* with Hamiltonian,

$$H_{\text{rr}} = \frac{1}{2mr_{\text{eq}}^2}\hat{\mathbf{L}}^2 = \frac{1}{2I_{\text{eq}}}\hat{\mathbf{L}}^2,$$

where $I_{\text{eq}} = mr_{\text{eq}}^2$ is the equilibrium moment of inertia of the molecule. This is the term left out in Chapter 4, in order to focus on the vibrational motion. In general, if the vibrational quantum number is not too large, the relative Hamiltonian is the sum of vibrational and rotational Hamiltonians - the energy eigenvalues are sums of vibrational and rotational energies. Otherwise, the x dependence of the moment of inertia cannot be neglected. Under ordinary conditions, this separation of vibration and rotation is a good approximation, and the rotational energy eigenvalues are the energy eigenvalues of the rigid rotor,

$$E_\ell = \frac{\ell(\ell+1)\hbar^2}{2I_{\text{eq}}}, \qquad \ell = 0, 1, 2, \ldots.$$

It is customary to use the letter J as the total angular momentum quantum when describing molecular rotation. Quantum number ℓ is reserved for electron orbital angular momentum. Thus, we write

$$E_{\text{rot, }J} = \frac{J(J+1)\hbar^2}{2I_{\text{eq}}}, \qquad J = 0, 1, 2, \ldots$$

for the molecular rotational energy. The associated rotational eigenfunctions are written as

$$Y_{J,M_J}(\theta,\phi), \qquad -J \le M_J \le J.$$

The angular momentum operators are also denoted by the letter J.

Expectation values of $\hat{J}_x$ and $\hat{J}_y$, for these states, can be determined by expressing $\hat{J}_x$ and $\hat{J}_y$ in terms of $\hat{J}_+$ and $\hat{J}_-$, and then using the known action of $\hat{J}_+$ and $\hat{J}_-$ on the Y_{J,M_J}. Since

$$\begin{aligned}
\hat{J}_x Y_{J,M_J} &= \frac{\hat{J}_+ + \hat{J}_-}{2} Y_{J,M_J} \\
&= \frac{(-1)^J \hbar}{2}\left([J(J+1) - M_J(M_J+1)]^{1/2} Y_{J,M_J+1} + [J(J+1) - M_J(M_J-1)]^{1/2} Y_{J,M_J-1}\right)
\end{aligned}$$

is a linear combination of Y_{J,M_J+1} and Y_{J,M_J-1}, functions orthogonal to Y_{J,M_J},

$$\left\langle Y_{J,M_J} \middle| \hat{J}_x Y_{J,M_J} \right\rangle = 0.$$

Non-zero expectation values are obtained only for $\hat{J}_x^2$ or $\hat{J}_y^2$ which have $\hat{J}_-\hat{J}_+$ and $\hat{J}_+\hat{J}_-$ terms which produce Y_{J,M_J}. For example, consider the expectation value of $\hat{J}_x$ for the top state, $Y_{J,J}$.

$$\begin{aligned} \left\langle Y_{J,J} \left| \hat{J}_x^2 Y_{J,J} \right. \right\rangle &= \frac{\hbar^2}{4}\left[J\left(J+1\right)-J\left(J-1\right)\right]\left\langle Y_{J,J}\left|Y_{J,J}\right.\right\rangle \qquad \text{arises from } \hat{J}_+\hat{J}_- \\ &= \frac{\hbar^2}{2}J. \end{aligned}$$

This state has the maximum possible z component of angular momentum, yet still has x and y components. This is a manifestion of the angular momentum commutator relationships. If the z component of angular is known, the x and y components cannot be known - and hence cannot be zero. This means that the angular momentum vector cannot point directly along the z axis. The $2J+1$ states, Y_{J,M_J} correspond to angular momentum vectors of length, $\sqrt{J\left(J+1\right)}\hbar$, with component $M_J\hbar$ along z. The vector is otherwise uncertain. The possible orientations of the angular momentum vector, for each state, maps out a cone (or a disk, if $M_J=0$). The discrete set of orientations are given by setting

$$\cos\theta_J = \frac{M_J}{\sqrt{J\left(J+1\right)}}.$$

This discretization of orientations is called *space quantization.* Figure 5.1 shows the angular momentum cones for $J=2$.

5.3 Rotational Spectroscopy

Dipole transitions between rotational levels are associated with changes in quantum number J. Different M_J states with the same J value have the same energy. Transitions between these states requires a magnetic field to split these levels. Transitions between different J states occurs for states connected by a non-zero transition matrix element. Suppose the oscillating electric field of incoming light is along the z axis (this is how we could define the z axis). In this case, the z component of the dipole moment of the molecule couples to the incoming light. This component is evidently $\mu_z\left(r,\theta\right)=\mu\left(r\right)\cos\theta$. The $\mu\left(r\right)$ factor in this expression leads to *vibrational transitions* - treated in Sec. 4.5. The $\cos\theta$ factor leads to *rotational transitions.* However, transitions occur only if $\mu_{\text{eq}}=\mu\left(r_{\text{eq}}\right)$ is non-zero - i.e., the molecule must have a *permanent dipole moment.* Homonuclear diatomics consequently exhibit no rotational spectra.

For rotational transitions, the matrix elements of $\mu_{\text{eq}}\cos\theta$ determine the probability of transition. To determine the selection rule for rotational transitions, it is sufficient to consider the matrix elements of the form,

$$\left\langle Y_{J',0}\left|\cos\theta Y_{J,0}\right.\right\rangle.$$

The $M_J=0$ state, $Y_{J,0}$, is a J th order polynomial in $\cos\theta$. Multiplying by $\cos\theta$ gives a $J+1$ th order polynomial which can be written as a combination of

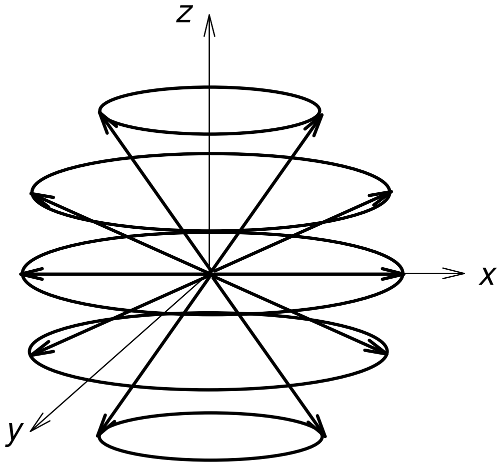


Figure 5.1: Space quanitzation. The possible angular momentum vector orientations associated with the Y_{2,M_J} states, $M_J = -2, -1, 0, 1, 2$ (from bottom to top).

$Y_{J+1,0}$ and $Y_{J-1,0}$. Consequently, the rotational transition selection rules are

$$\Delta J = \pm 1. \tag{5.7}$$

and

$$\Delta M_J = 0. \tag{5.8}$$

5.3.1 Microwave absorption spectroscopy

The change in energy associated with a $\Delta J = +1$ transition - corresponding to absorption of light - is given by

$$\begin{aligned} \Delta E_{\mathrm{rot},\ J} &= E_{\mathrm{rot},\ J+1} - E_{\mathrm{rot},\ J} = \frac{(J+1)(J+2)\hbar^2}{2I_{\mathrm{eq}}} - \frac{J(J+1)\hbar^2}{2I_{\mathrm{eq}}} \\ &= 2(J+1)\frac{\hbar^2}{2I_{\mathrm{eq}}} \\ &= 2(J+1)B_{\mathrm{rot}}, \end{aligned} \tag{5.9}$$

where

$$B_{\mathrm{rot}} = \frac{\hbar^2}{2I_{\mathrm{eq}}}$$

is called the *rotational constant*. Unlike the vibrational energy levels, the rotational levels are not equally spaced. Consequently, even though transitions only

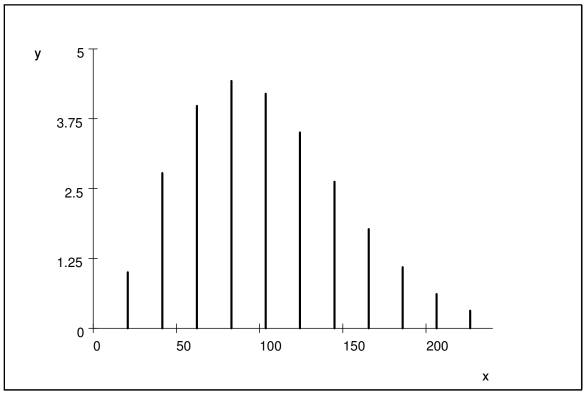


Figure 5.2: An ideal microwave absorption spectrum - i.e., the rotational spectrum - of HCl at 400 K. The horizontal axis is wavenumbers in cm^{-1}. The vertical axis is peak intensity relative to the $J = 0$ peak.

occur between neighboring levels, there is a series of rotational transition frequencies - each frequency associated with a different initial J value. Equation 5.9 shows that the frequencies, $\nu_{\mathrm{rot},\ J} = \Delta E_{\mathrm{rot},\ J}/h$, are equally spaced. The intensity of each peak is determined by the population of diatomic molecules initially at level J. Since there are $2J+1$ states (the different M_J values) with energy, $E_{\mathrm{rot},\ J}$, the peak intensities are described (normalized to the $J = 0$ peak intensity) by

$$\begin{aligned}\rho_{\mathrm{rot},\ J}(T) &= (2J+1)\exp\left(-\frac{E_{\mathrm{rot},\ J}}{k_{\mathrm{B}}T}\right)\\ &= (2J+1)\exp\left(-\frac{J(J+1)\,B_{\mathrm{rot}}}{k_{\mathrm{B}}T}\right).\end{aligned}$$

This is the *Boltzmann distribution* of rotational state populations.

Above we see the rotational spectrum of HCl at 400 K, with wavenumbers on the horizontal axis (in cm^{-1}). The transition peak intensities follow the shape of $\rho_{\mathrm{rot},\ J}(T)$. The *transition frequency* associated with ground state, $J = 0$, is

$$\nu_{\mathrm{rot},\ 0} = \frac{\Delta E_{\mathrm{rot},\ 0}}{h} = \frac{2B_{\mathrm{rot}}}{h}.$$

Rotational frequencies are usually reported as *wavenumbers* in cm^{-1};

$$\begin{aligned}\bar{\nu}_{\mathrm{rot},\ 0} &= \frac{\nu_{\mathrm{rot},\ 0}}{c} = \frac{2B_{\mathrm{rot}}}{hc}\\ &= 2\bar{B}_{\mathrm{rot}},\end{aligned}$$

where $\bar{B}_{\mathrm{rot}}$ is the rotational constant in wavenumbers. The wavenumber spacings between successive peaks is

$$\begin{aligned}\Delta\bar{\nu}_{\text{rot}} &= \bar{\nu}_{\text{rot},\ J+1} - \bar{\nu}_{\text{rot},\ J} \\ &= \frac{\Delta E_{\text{rot},\ J+1}}{hc} - \frac{\Delta E_{\text{rot},\ J}}{hc} \\ &= 2\left(J+2\right)\bar{B}_{\text{rot}} - 2\left(J+1\right)\bar{B}_{\text{rot}} \\ &= 2\bar{B}_{\text{rot}}.\end{aligned}$$

Thus, the rotational transitions occur at $\bar{\nu}_{\text{rot}} = 2\bar{B}_{\text{rot}},\ 4\bar{B}_{\text{rot}},\ 6\bar{B}_{\text{rot}}, \ldots$, with the Boltzmann distribution providing a peak intensity envelope.

There is an overall exponential decay of peak intensities (of the form, $\exp\left(-\alpha J^2\right)$) for large increasing J. However, the $2J+1$ degeneracy factor increases (with increasing J) the number of possible transtions, giving rise to a maximum peak intensity for $J > 0$. Setting $d\rho_{\text{rot},\ J}/dJ = 0$, and solving for J (choose nearest integer) gives the J level with the highest population - the level for which the increase in degeneracy is balanced by the exponential decrease due to increasing energy.

$$\begin{aligned}\frac{d\rho_{\text{rot},\ J}\left(T\right)}{dJ} &= 2\exp\left(-\frac{J\left(J+1\right)B_{\text{rot}}}{k_{\text{B}}T}\right) - \left(2J+1\right)^2\frac{B_{\text{rot}}}{k_{\text{B}}T}\exp\left(-\frac{J\left(J+1\right)B_{\text{rot}}}{k_{\text{B}}T}\right) \\ &= \left(2-\left(2J+1\right)^2\frac{B_{\text{rot}}}{k_{\text{B}}T}\right)\exp\left(-\frac{J\left(J+1\right)B_{\text{rot}}}{k_{\text{B}}T}\right) = 0.\end{aligned}$$

Because the exponential function is always positive, we can divide by the exponential factor to get

$$2-\left(2J+1\right)^2\frac{B_{\text{rot}}}{k_{\text{B}}T} = 0$$

or

$$\begin{aligned}J_{\max} &= \frac{1}{2}\left(\frac{2k_{\text{B}}T}{B_{\text{rot}}}\right)^{1/2} - \frac{1}{2} \\ &= \left(\frac{k_{\text{B}}T}{2B_{\text{rot}}}\right)^{1/2} - \frac{1}{2}.\end{aligned}$$

Microwave absorption has an important application in cooking - the *microwave oven.* A microwave oven produces microwave radiation tuned to the absorption band of water in the liquid phase. The water molecules, in food, make J to $J+1$ transitions - they rotate faster. The increased rotational energy gets distributed to other degrees of freedom by collisions with other molecules. The food gets hot.

5.3.2 The classical rigid rotor

It is instructive to consider the classical rigid rotor model of a diatomic molecule. If the diatomic molecule has angular momentum L, then the rotational frequency is given by (recall that $L = I_{\text{eq}}\omega$, where $\omega = 2\pi\nu_{\text{rot}}$ is the angular velocity - or angular frequency)

$$\nu_{\text{rot}} = \frac{1}{2\pi}\frac{L}{I_{\text{eq}}}.$$

In classical mechanics, the rotational frequency varies continuously with angular momentum. In quantum mechanics angular momentum is quantized. If we associate the J th level of the quantum system with a classical angular momentum, $L = (J+1)\,\hbar$, $J = 0, 1, \ldots$, then the allowed rotational frequencies are

$$\begin{aligned} \nu_{\text{rot, } J} &= \frac{1}{2\pi}\frac{(J+1)\,\hbar}{I_{\text{eq}}} \\ &= \frac{1}{h}\frac{(J+1)\,\hbar^2}{I_{\text{eq}}} \\ &= \frac{2\,(J+1)\,B_{\text{rot}}}{h}. \end{aligned}$$

These are the quantum transition frequencies. Thus, rotational transitions occur when the incident microwave frequency matches quantized values of the classical rotational frequency of the diatomic molecule. Successive J value peaks correspond to progressively larger angular momentum values - in increments of $\hbar$. This correspondence between quantum transition frequencies and classical rotational frequencies requires identifying angular momentum, $L = (J+1)\,\hbar$, with the J th transition (from level J to $J+1$). This angular momentum is the root-mean-square of the angular momenta associated with the J th and $(J+1)$ th quantum states - $\sqrt{J\,(J+1)}\hbar$ and $\sqrt{(J+1)\,(J+2)}\hbar$, respectively. Specifically,

$$\begin{aligned} \sqrt{\frac{J\,(J+1)\,\hbar^2 + (J+1)\,(J+2)\,\hbar^2}{2}} &= \sqrt{\frac{(J+1)\,(2J+2)}{2}}\hbar \\ &= (J+1)\,\hbar = L. \end{aligned}$$

The J th and $(J+1)$ th quantum states are initial and final states of the quantum transition.

5.4 Ro-vibrational spectroscopy

The infrared spectra of diatomic molecules do not correspond to a single peak at the vibrational frequency. The reason is that the vibrational quantum number selection rule, $\Delta\upsilon = \pm 1$, is accompanied by the rotational quantum number selection rule, $\Delta J = \pm 1$. For example, in an infrared absorption spectrum ($\Delta\upsilon = +1$), the change in energy of the molecule is the change in vibrational energy plus the change in rotational energy (a much smaller value). The change in rotational energy cannot be zero because $\Delta J = 0$ is not allowed. Only $\Delta J = +1$ and $\Delta J = -1$ transitions are possible.

In the case of $\Delta J = +1$, the peaks of the rotational absorption spectrum (these lines correspond to $\Delta J = +1$) are seen again - this time shifted by the vibrational frequency into the infrared range. These peaks constitute the *R branch* of the *ro-vibrational spectrum* - the *gas phase infrared absorption spectrum.* The associated wavenumbers are

$$\begin{aligned}\bar{\nu}_{\mathrm{R},J} &= \bar{\nu}_{\mathrm{vib}} + \frac{E_{\mathrm{rot},\ J+1} - E_{\mathrm{rot},\ J}}{hc} \\ &= \bar{\nu}_{\mathrm{vib}} + \left[(J+1)(J+2) - J(J+1)\right]\bar{B}_{\mathrm{rot}} \\ &= \bar{\nu}_{\mathrm{vib}} + 2(J+1)\bar{B}_{\mathrm{rot}}, \qquad J = 0, 1, 2, \ldots.\end{aligned}$$

The case of $\Delta J = -1$ is not seen in the rotational absorption spectrum, but it appears in the vibrational absorption spectrum because the sum of the change in vibrational and rotational energy is still positive in this case. The associated transition frequencies are below the vibrational absorption frequency because rotational energy decreases in the transition. The transition wavenumbers in this case are

$$\begin{aligned}\bar{\nu}_{\mathrm{P},J} &= \bar{\nu}_{\mathrm{vib}} + \frac{E_{\mathrm{rot},\ J-1} - E_{\mathrm{rot},\ J}}{hc} \\ &= \bar{\nu}_{\mathrm{vib}} + \left[(J-1)J - J(J+1)\right]\bar{B}_{\mathrm{rot}} \\ &= \bar{\nu}_{\mathrm{vib}} - 2J\bar{B}_{\mathrm{rot}}, \qquad J = 1, 2, \ldots.\end{aligned}$$

Note that there is no $J = 0$ transition because that is the lowest J state - it cannot be lowered to $J - 1$. This series of peaks, below $\bar{\nu}_{\mathrm{vib}}$ is called the *P branch* of the *ro-vibrational spectrum.*

The *Q branch* - with $\Delta J = 0$ - is forbidden for diatomic molecules, and not seen. However, the Q branch is seen for triatomic and larger molecules. The selection rules are $\Delta J = 0$, or ± 1, beyond diatomics.

The idealized infrared absorption spectrum of $^{1}\mathrm{H}^{35}\mathrm{Cl}$ is depicted below.The horizontal axis gives the wavenumbers in units of cm^{-1}. Note the asymmetry, about $\bar{\nu}_{\mathrm{vib}} = 2991\ \mathrm{cm}^{-1}$, resulting from the different initial states in the two series - the P branch starts with $J = 1$, while the R branch starts with $J = 0$.

In real ro-vibrational spectra, there are deviations from the above idealized spectrum. One deviation is due to centrifugal distortion. For large J states, the angular momentum is large enough to shift the molecular equilibrium bond distance - the molecule is stretched. This shifts the moment of inertia to larger values. The spacing between peaks becomes J dependent, decreasing with increasing J. The P and R branches would appear to be compressed near the outside edges, compared to the idealized spectrum shown above.

5.4.1 Raman Spectroscopy

Raman spectroscopy has different selection rules. The transition matrix element for a Raman transition - inelastic scattering of a UV frequency photon - is a matrix element of the molecule polarizability which is proportional to $\cos^2\theta$.

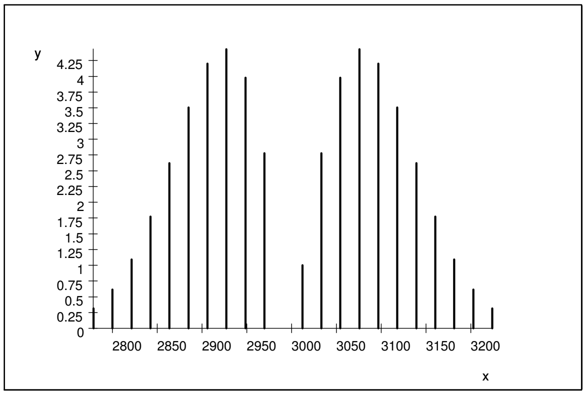


Figure 5.3: The idealized infrared absorption spectrum of $^{1}H^{35}Cl$. Relative peak intensity (relative to the R $J = 0$ peak) is plotted versus wavenumbers in cm^{-1}.

The selection rule is $\Delta J = 0$ or ± 2. The O, Q and S branches ($\Delta J = -2, 0$ and 2, respectively) are observed. Moreover, there are many vibrational modes that are not *infrared active* - electric dipole does not change - that are *Raman active*. The other Raman selection rule is that polarizability must change upon vibration. The polarizability often changes when the electric dipole does not. In fact, for *centrosymmetric* molecules - molecules with a center of symmetry - Raman active modes are always infrared inactive, and vice versa - they are mutually exclusive.

Example 5.2. Consider the superposition state,

$$\begin{aligned} Y_x &= \frac{1}{2^{1/2}} \left(Y_{1,1} - Y_{1,-1} \right) \\ &= \left(\frac{4\pi}{3} \right)^{-1/2} \sin\theta \cos(\phi) \, . \end{aligned}$$

(a) What is the expectation value of $\hat{J}_z$ for a rigid rotor in this state?

(b) Show that p_x is an eigenfunction of $\hat{J}^2$. What is the associated eigenvalue?

(c) For which θ and ϕ is the probability density, $|Y_x|^2$, largest? In terms of Cartesian axes, which directions are mostly likely?

Solution 5.2.

(a) The $\hat{J}_z$ expectation is

$$\begin{aligned}
\left\langle Y_x \middle| \hat{J}_z Y_x \right\rangle &= \left\langle \frac{1}{2^{1/2}} \left(Y_{1,1} - Y_{1,-1}\right) \middle| \hat{J}_z \frac{1}{2^{1/2}} \left(Y_{1,1} - Y_{1,-1}\right) \right\rangle \\
&= \frac{1}{2} \left\langle Y_{1,1} - Y_{1,-1} \middle| \hat{J}_z Y_{1,1} - \hat{J}_z Y_{1,-1} \right\rangle \\
&= \frac{1}{2} \left\langle Y_{1,1} - Y_{1,-1} \middle| \hbar Y_{1,1} - (-\hbar) Y_{1,-1} \right\rangle \\
&= \frac{\hbar}{2} \left(\left\langle Y_{1,1} \middle| Y_{1,1} + Y_{1,-1} \right\rangle - \left\langle Y_{1,-1} \middle| Y_{1,1} + Y_{1,-1} \right\rangle \right) \\
&= \frac{\hbar}{2} \left(\underbrace{\left\langle Y_{1,1} \middle| Y_{1,1} \right\rangle}_{=1} + \underbrace{\left\langle Y_{1,1} \middle| Y_{1,-1} \right\rangle}_{=0} - \underbrace{\left\langle Y_{1,-1} \middle| Y_{1,1} \right\rangle}_{=0} - \underbrace{\left\langle Y_{1,-1} \middle| Y_{1,-1} \right\rangle}_{=1} \right) \\
&= 0.
\end{aligned}$$

(b) Apply $\hat{J}^2$ to Y_x;

$$\begin{aligned}
\hat{J}^2 Y_x &= \hat{J}^2 \frac{1}{2^{1/2}} \left(Y_{1,1} - Y_{1,-1}\right) \\
&= \frac{1}{2^{1/2}} \left(\hat{J}^2 Y_{1,1} - \hat{J}^2 Y_{1,-1} \right) \\
&= \frac{1}{2^{1/2}} \left(2\hbar^2 Y_{1,1} - 2\hbar^2 Y_{1,-1} \right) \\
&= 2\hbar^2 \frac{1}{2^{1/2}} \left(Y_{1,1} - Y_{1,-1}\right) \\
&= 2\hbar^2 Y_x.
\end{aligned}$$

p_x is an eigenfunction of $\hat{J}^2$, with eigenvalue $2\hbar^2$.

(c) The probability density is

$$\left| Y_x (\theta, \phi) \right|^2 = \left(\frac{4\pi}{3} \right)^{-1} \sin^2 \theta \cos^2 (\phi).$$

This function is a maximum when $\theta = \pi/2$ (there is only one maximum of $\sin^2 \theta$ in the range, $0 \le \theta \le \pi$) and $\phi = 0$ or π (here the range is $0 \le \phi < 2\pi$). $\theta = \pi/2$ corresponds to the $x\,y$ plane, while $\phi = 0$ or π correspond to the positive and negative x axis, respectively. The density is concentrated in the directions of the positive and negative x axis.

Example 5.3. Consider the superposition state,

$$\begin{aligned}
Y &= 2^{-1/2} \left(Y_{1,1} - Y_{0,0}\right) \\
&= (8\pi)^{-1/2} \left(\left(\frac{3}{2} \right)^{1/2} \sin \theta \exp (i\phi) - 1 \right)
\end{aligned}$$

(a) What is the expectation value of $\hat{J}_z$ for a rigid rotor in this state?

(b) Suppose Y is the state of a rigid rotor at time zero. Determine the expectation of $\cos\phi$ at the times, $t = 0$, $\frac{\hbar}{4B_{\text{rot}}}$ and $\frac{\hbar}{2B_{\text{rot}}}$. Show that this expectation value is periodic in time. What is the period?

Solution 5.3.

(a) The expectation value of $\hat{J}_z$ is

$$\begin{aligned}
& \langle Y \left| \hat{J}_z Y \right\rangle \\
= \; & 2^{-1} \langle Y_{1,1} - Y_{0,0} \left| \hat{J}_z \left(Y_{1,1} - Y_{0,0} \right) \right\rangle \\
= \; & 2^{-1} \langle Y_{1,1} - Y_{0,0} \left| \hat{J}_z Y_{1,1} - \hat{J}_z Y_{0,0} \right\rangle \\
= \; & 2^{-1} \langle Y_{1,1} - Y_{0,0} \left| \hbar Y_{1,1} - 0 \times Y_{0,0} \right\rangle \\
= \; & 2^{-1} \hbar \left(\underbrace{\langle Y_{1,1} | Y_{1,1} \rangle}_{=1} - \underbrace{\langle Y_{0,0} | Y_{1,1} \rangle}_{=0} \right) \\
= \; & \frac{\hbar}{2}.
\end{aligned}$$

(b) The time dependence of the state is given by solving the TDSE. Since the state is already written as a sum of energy eigenstates of the rigid rotor, the time dependence is given by applying time dependent phase factors to the coefficients of the energy eigenstates - here, $Y_{1,1}$ and $Y_{0,0}$. The phase of these factors is $-E_{\text{rot},\, J} t/\hbar$. Specifically, at time t we have

$$\begin{aligned}
Y_t &= 2^{-1/2} \left(\exp\left(-i \frac{E_{\text{rot},\,1} t}{\hbar} \right) Y_{1,1} - \exp\left(-i \frac{E_{\text{rot},\,0} t}{\hbar} \right) Y_{0,0} \right) \\
&= 2^{-1/2} \left(\exp\left(-i \frac{2B_{\text{rot}} t}{\hbar} \right) Y_{1,1} - Y_{0,0} \right),
\end{aligned}$$

since $E_{\text{rot},\,1} = 2B_{\text{rot}}$ and $E_{\text{rot},\,0} = 0$. We write this equation in the form,

$$\begin{aligned}
Y_t &= 2^{-1/2} \left(\exp(-i\tau) Y_{1,1} - Y_{0,0} \right) \\
&= (8\pi)^{-1/2} \left(\left(\frac{3}{2} \right)^{1/2} \sin\theta \exp\left(i (\phi - \tau) \right) - 1 \right)
\end{aligned}$$

and evaluate the expectation of $\cos\phi$.

$$\begin{aligned}
& \langle Y_t \left| \cos\phi Y_t \right\rangle \\
= \; & \int_0^{2\pi} \int_0^{\pi} \left| Y_t(\theta, \phi) \right|^2 \cos\phi \sin\theta \, d\theta d\phi \\
= \; & (8\pi)^{-1} \int_0^{2\pi} \int_0^{\pi} \left(\frac{3}{2} \sin^2\theta - 2 \left(\frac{3}{2} \right)^{1/2} \sin\theta \cos(\phi - \tau) + 1 \right) \cos\phi \sin\theta \, d\theta d\phi.
\end{aligned}$$

There are three terms inside the integrand. We integrate them separately. The last term is easily evaluated. It gives

$$\begin{aligned}
& (8\pi)^{-1}\int_0^{2\pi}\int_0^{\pi}\cos\phi\sin\theta\,d\theta d\phi \\
= {} & (8\pi)^{-1}\underbrace{\int_0^{2\pi}\cos\phi\,d\phi}_{=0}\int_0^{\pi}\sin\theta\,d\theta \\
= {} & 0.
\end{aligned}$$

$\cos\phi$ integrates to zero over the interval 0 to 2π, as it is periodic on this interval. Alternatively, just do the the integral - the antiderivative, $\sin\phi$, is also periodic on the interval. The next term to consider is

$$\begin{aligned}
& (8\pi)^{-1}\int_0^{2\pi}\int_0^{\pi}\frac{3}{2}\sin^2\theta\cos\phi\sin\theta\,d\theta d\phi \\
= {} & \frac{3}{16\pi}\underbrace{\int_0^{2\pi}\cos\phi\,d\phi}_{=0}\int_0^{\pi}\sin^3\theta\,d\theta \\
= {} & 0.
\end{aligned}$$

It is also zero. Thus,

$$\begin{aligned}
& \langle Y_t\,|\cos\phi Y_t\rangle \\
= {} & -(8\pi)^{-1}2\left(\frac{3}{2}\right)^{1/2}\int_0^{2\pi}\int_0^{\pi}\sin\theta\cos(\phi-\tau)\cos\phi\sin\theta\,d\theta d\phi \\
= {} & -(4\pi)^{-1}\left(\frac{3}{2}\right)^{1/2}\int_0^{2\pi}\cos(\phi-\tau)\cos\phi\,d\phi\int_0^{\pi}\sin^2\theta\,d\theta.
\end{aligned}$$

The θ integral is

$$\begin{aligned}
\int_0^{\pi}\sin^2\theta\,d\theta &= \frac{1}{2}\int_0^{\pi}(1-\cos 2\theta)\,d\theta \\
&= \frac{\pi}{2}.
\end{aligned}$$

The ϕ integral depends upon τ. There are three cases: (i) for $t=0$, $\tau=0$; (ii) for $t=\frac{h}{4B_{\text{rot}}}$,

$$\tau=\frac{2B_{\text{rot}}t}{\hbar}=\pi$$

and (iii) for $t=\frac{h}{2B_{\text{rot}}}$, $\tau=2\pi$. Since cosine is periodic with period 2π, the expectation of $\cos\phi$ is periodic in τ with the same period. The expectation is the same for $t=0$ and $t=\frac{h}{2B_{\text{rot}}}$, the period of the oscillation in time. For these times,

$$\begin{aligned}\int_0^{2\pi} \cos(\phi-\tau)\cos\phi\, d\phi &= \int_0^{2\pi} \cos^2\phi\, d\phi \\ &= \frac{1}{2}\int_0^{2\pi} (1+\cos 2\theta)\ d\theta \\ &= \pi,\end{aligned}$$

and

$$\begin{aligned}\langle Y_t\ |\cos\phi Y_t\rangle &= -(4\pi)^{-1}\left(\frac{3}{2}\right)^{1/2}\times\pi\times\frac{\pi}{2} \\ &= -\frac{\pi}{8}\left(\frac{3}{2}\right)^{1/2}.\end{aligned}$$

The second time, $t=\frac{h}{4B_{\mathrm{rot}}}$, is half of the period of the expectation. In this case,

$$\begin{aligned}\int_0^{2\pi} \cos(\phi-\tau)\cos\phi\, d\phi &= \int_0^{2\pi} \cos(\phi-\pi)\cos\phi\, d\phi \\ &= -\int_0^{2\pi} \cos^2\theta\, d\theta \\ &= -\pi,\end{aligned}$$

and

$$\langle Y_t\ |\cos\phi Y_t\rangle = \frac{\pi}{8}\left(\frac{3}{2}\right)^{1/2}.$$

The expectation of $\cos\phi$ is periodic. It oscillates between $\pm\frac{\pi}{8}\left(\frac{3}{2}\right)^{1/2}$ with period, $\frac{h}{2B_{\mathrm{rot}}}=2\pi\frac{I_{\mathrm{eq}}}{\hbar}$, the period of a classical rigid rotor with angular momentum $\hbar$.

Example 5.4. Determine the moment of inertia of ${}^1\mathrm{H}^{35}\mathrm{Cl}$ molecules in the gas phase from the spectrum shown in Fig. 5.3. Use your result to get the equilibrium bond distance. You can use $m_1 = 1$ amu and $m_2 = 35$ amu for the masses of the two atoms. Analysis of high resolution data requires highly precise masses.

Solution 5.4. From the figure, we can see that from 2800 to 3200 cm^{-1} there are slightly less than 19 spacings (including the un-allowed Q peak). The spacing is estimated to be about

$$\begin{aligned}&\frac{405}{19}\ \mathrm{cm}^{-1} \\ =\ &21.3\ \mathrm{cm}^{-1}.\end{aligned}$$

Setting this equal to $2\bar{B}_{\text{rot}}$ gives

$$2\bar{B}_{\text{rot}} = \frac{\hbar^2}{hcI_{\text{eq}}} = 21.3\ \text{cm}^{-1},$$

or

$$\begin{aligned} I_{\text{eq}} &= \frac{h}{4\pi^2 c \times 21.3\ \text{cm}^{-1}} \\ &= \frac{6.63 \times 10^{-34}\ \text{kg m}^2\ \text{s}^{-1}}{4\pi^2 \times 3.00 \times 10^{10}\ \text{cm s}^{-1} \times 21.3\ \text{cm}^{-1}} \\ &= 2.63 \times 10^{-47}\ \text{kg m}^2. \end{aligned}$$

The moment of inertia is just the product of reduced mass and equilibrium distance squared. The reduced mass is

$$\begin{aligned} m &= \frac{m_1 m_2}{m_1 + m_2} = \frac{35}{36}\ \text{amu} \\ &= 0.972\ \text{amu} \times \frac{1}{6.02 \times 10^{23}\ \text{amu g}^{-1}} \\ &= 1.61 \times 10^{-24}\ \text{g} \\ &= 1.61 \times 10^{-27}\ \text{kg}. \end{aligned}$$

Thus,

$$\begin{aligned} r_{\text{eq}} &= \left(\frac{I_{\text{eq}}}{m}\right)^{1/2} \\ &= \left(\frac{2.63 \times 10^{-47}\ \text{kg m}^2}{1.61 \times 10^{-27}\ \text{kg}}\right)^{1/2} \\ &= 1.28 \times 10^{-10}\ \text{m} \\ &= 128\ \text{pm}. \end{aligned}$$

Example 5.5. Match each of the spectra in list a, with selection rules from list b.

(a) Spectra:
   i. Microwave absorption spectrum of a diatomic molecule
   ii. Infrared absorption spectrum of a diatomic molecule
   iii. Infrared absorption spectrum of a triatomic molecule
   iv. Raman scattering spectrum of a triatomic molecule
   v. Optical absorption spectrum of a quantum dot modeled as a particle in a box

(b) Selection rules:
   i. $\Delta v = 1$

  ii. $\Delta J = 0, \pm 1$
  iii. $\Delta J = 1$
  iv. $\Delta J = 0, \pm 2$
  v. $\Delta J = \pm 1$
  vi. $\Delta n = 1, 3, 5, \ldots$

Solution 5.5. Spectra:

(a)
  i. Microwave absorption spectrum of a diatomic molecule: $\Delta J = 1$
  ii. Infrared absorption spectrum of a diatomic molecule: $\Delta v = 1$ and $\Delta J = \pm 1$
  iii. Infrared absorption spectrum of a triatomic molecule: $\Delta v = 1$ and $\Delta J = 0, \pm 1$
  iv. Raman scattering spectrum of a triatomic molecule: $\Delta v = 1$ and $\Delta J = 0, \pm 2$
  v. Optical absorption spectrum of a quantum dot modeled as a particle in a box: $\Delta n = 1, 3, 5, \ldots$

5.5 Appendix: the volume element in spherical coordinates

Wavefunctions, $\psi(x, y, z)$, are normalized by requiring that

$$\langle \psi | \psi \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi(x, y, z)|^2 \, dxdydz = 1.$$

When the wavefunction is expressed in terms of spherical coordinates, the three dimensional integral over Cartesian coordinates, x, y and z, must be written as a three dimensional integral over r, θ and ϕ. Because the transformation between Cartestian and spherical coordinates is not linear, the volume element,

$$\begin{aligned} dV &= dxdydz \\ &= r^2 \sin\theta drd\theta d\phi \end{aligned} \quad (5.10)$$

has r and θ dependence, when expressed in terms of $drd\theta d\phi$.

To prove the above formula for the volume element, consider infinitesimal changes in r, θ and ϕ. Each of these changes is associated with an infinitesimal vector. For example, dr, is associated with the vector,

$$\begin{pmatrix} \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial r} \end{pmatrix} dr.$$

The three vectors define a parallelepiped in three dimensions - they determine three edges emanating from a common vertex. The volume of this parallelepiped is the volume element in spherical coordinates. This volume is the dot product of one edge vector with the cross product of the other two. This result can be expressed as a 3×3 determinant, with each vector providing a column;

$$
\begin{aligned}
dV &= \begin{vmatrix} \frac{\partial x}{\partial r}dr & \frac{\partial x}{\partial \theta}d\theta & \frac{\partial x}{\partial \phi}d\phi \\ \frac{\partial y}{\partial r}dr & \frac{\partial y}{\partial \theta}d\theta & \frac{\partial y}{\partial \phi}d\phi \\ \frac{\partial z}{\partial r}dr & \frac{\partial z}{\partial \theta}d\theta & \frac{\partial z}{\partial \phi}d\phi \end{vmatrix} \\
&= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} drd\theta d\phi \qquad \text{a property of determinants} \\
&= \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} drd\theta d\phi \\
&= r^2 \left(\cos\theta \begin{vmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \end{vmatrix} \right. \\
&= \left. +\sin\theta \begin{vmatrix} \sin\theta\cos\phi & -\sin\theta\sin\phi \\ \sin\theta\sin\phi & \sin\theta\cos\phi \end{vmatrix} \right) drd\theta d\phi \\
&= r^2 \left(\cos^2\theta\sin\theta\left(\cos^2\phi + \sin^2\phi\right) + \sin^3\theta\left(\cos^2\phi + \sin^2\phi\right)\right) drd\theta d\phi \\
&= r^2 \sin\theta\left(\cos^2\theta + \sin^2\theta\right) drd\theta d\phi \\
&= r^2 \sin\theta \, drd\theta d\phi.
\end{aligned}
$$

With this volume element, we can express the normalization condition for a wavefunction expressed in terms of spherical coordinates. Specifically,

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \left|\psi\left(r,\theta,\phi\right)\right|^2 r^2 \sin\theta \, drd\theta d\phi = 1.$$

The ranges of integration covers all of space. The radial coordinate extends from 0 to ∞. The angle θ varies from 0 to π - from the positive z axis to the negative z axis. The angle in the x y plane varies from 0 to 2π.

If the wavefunction is a product of functions of just r, just θ and just ϕ, then the factors can be separately normalized. Specifically, if

$$\psi\left(r,\theta,\phi\right) = \psi_r\left(r\right) Y\left(\theta,\phi\right),$$

then

$$\left\langle \psi_r \middle| \psi_r \right\rangle = \int_0^{\infty} \left|\psi_r\left(r\right)\right|^2 r^2 \, dr = 1$$

and

$$\left\langle Y \middle| Y \right\rangle = \int_0^{2\pi}\int_0^{\pi} \left|Y\left(\theta,\phi\right)\right|^2 \sin\theta \, d\theta d\phi = 1.$$

If

$$Y\left(\theta,\phi\right) = \Theta\left(\theta\right)\Phi\left(\phi\right),$$

then

$$\langle \Theta \,| \Theta \rangle = \int_0^{\pi} \left| \Theta \left(\theta \right) \right|^2 \sin \theta \, d\theta = 1.$$

and

$$\langle \Phi \,| \Phi \rangle = \int_0^{2\pi} \left| \Phi \left(\phi \right) \right|^2 d\phi = 1.$$

The inner product $\langle \cdot \,| \cdot \rangle$ is an integral defined by the context. It is an integral, with appropriate weight function (i.e., such as r^2 or $\sin \theta$), over all admissable values of the coordinates of the wavefunctions in the inner product.

Chapter 6

Perturbations

There are few systems for which the Schrödinger equation can be solved exactly. Most systems must be solved using numerical methods. One approach, *perturbation theory*, is very useful when the system of interest is close to a system which has already been solved. In this approach, the system which is already solved is the *zero order system*, and the system is said to be *perturbed* from of the zero order system. The Hamiltonian of the perturbed system is the zero order Hamiltonian, $\hat{H}^{(0)}$, plus a perturbation, $\hat{H}^{(1)}$. If the perturbation is not too large, then the energy eigenvalues and eigenfunctions of the perturbed system can be written as the corresponding zero order values plus corrections.that can be computed.

6.1 Time independent perturbation theory

Let

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}^{(1)},$$

where λ is a convenient perturbation scaling parameter. In the end, we set it to 1. However, in the meantime, keeping λ free lets us keep track of the order of correction terms. Suppose $\left\{\psi_n^{(0)}\right\}$ are the zero order eigenfunctions;

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}. \tag{6.1}$$

The eigenvalues and eigenfunctions of $\hat{H}$ can be written as power series in the scaling parameter, λ:

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \tag{6.2}$$

and

$$\psi_n = \psi_n^{(0)} + \lambda \psi_n^{(1)} + \dots. \tag{6.3}$$

Equations for the correction terms, $E_n^{(1)}$, $E_n^{(2)}$, $\psi_n^{(1)}$ etc., are determined by substituting the power series expressions into the perturbed system Schrödinger equation,

$$\hat{H}\psi_n = E_n\psi_n,$$

to give

$$\left(\hat{H}^{(0)} + \lambda\hat{H}^{(1)}\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \ldots\right) = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots\right)\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \ldots\right).$$

The left side of this equation has the form,

$$\begin{aligned}
&\hat{H}^{(0)}\left(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \ldots\right) + \hat{H}^{(1)}\left(\lambda\psi_n^{(0)} + \lambda^2\psi_n^{(1)} + \ldots\right)\\
&\hat{H}^{(0)}\psi_n^{(0)} + \lambda\left[\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}^{(1)}\psi_n^{(0)}\right] + \lambda^2\left[\hat{H}^{(0)}\psi_n^{(2)} + \hat{H}^{(1)}\psi_n^{(1)}\right] + \ldots.
\end{aligned}$$

The right side has the form,

$$\begin{aligned}
& E_n^{(0)}\psi_n^{(0)} + \lambda E_n^{(1)}\psi_n^{(0)} + \lambda^2 E_n^{(2)}\psi_n^{(0)} + \ldots + \lambda E_n^{(0)}\psi_n^{(1)} + \lambda^2 E_n^{(1)}\psi_n^{(1)} + \ldots + \lambda^2 E_n^{(0)}\psi_n^{(2)}\\
= \;& E_n^{(0)}\psi_n^{(0)} + \lambda\left[E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}\right] + \lambda^2\left[E_n^{(2)}\psi_n^{(0)} + E_n^{(1)}\psi_n^{(1)} + E_n^{(0)}\psi_n^{(2)}\right] + \ldots.
\end{aligned}$$

Equating coefficients of like powers of λ on either side of the equation gives the following relations:

Coefficients of λ^0:

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}.$$

This is just the unperturbed Schrödinger equation.

Coefficients of λ^1:

$$\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}^{(1)}\psi_n^{(0)} = E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}. \tag{6.4}$$

Restrict $\psi_n^{(1)}$ to be orthogonal to $\psi_n^{(0)}$. This removes redundancy in Eq. 6.3 . Specifically, any component of $\psi_n^{(0)}$ included in $\psi_n^{(1)}$ should already be considered included within the first term. We can assume that $\psi_n^{(0)}$ is normalized. In this case, the ψ_n of Eq. 6.3 is not normalized. This is not a problem; ψ_n can always be normalized after λ is set to 1. Now, take the inner product of both sides of the above equation, with $\psi_n^{(0)}$, to get

$$\begin{aligned}
\left\langle\psi_n^{(0)}\middle|\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}^{(1)}\psi_n^{(0)}\right\rangle &= \left\langle\psi_n^{(0)}\middle|E_n^{(1)}\psi_n^{(0)} + E_n^{(0)}\psi_n^{(1)}\right\rangle\\
\left\langle\psi_n^{(0)}\middle|\hat{H}^{(0)}\psi_n^{(1)}\right\rangle + \left\langle\psi_n^{(0)}\middle|\hat{H}^{(1)}\psi_n^{(0)}\right\rangle &= E_n^{(1)}\underset{=1}{\left\langle\psi_n^{(0)}\middle|\psi_n^{(0)}\right\rangle} + E_n^{(0)}\underset{=0}{\left\langle\psi_n^{(0)}\middle|\psi_n^{(1)}\right\rangle}\\
\left\langle\hat{H}^{(0)}\psi_n^{(0)}\middle|\psi_n^{(1)}\right\rangle + \left\langle\psi_n^{(0)}\middle|\hat{H}^{(1)}\psi_n^{(0)}\right\rangle &= E_n^{(1)} \qquad \hat{H}^{(0)} \text{ is Hermitian}\\
\left\langle E_n^{(0)}\psi_n^{(0)}\middle|\psi_n^{(1)}\right\rangle + \left\langle\psi_n^{(0)}\middle|\hat{H}^{(1)}\psi_n^{(0)}\right\rangle &= E_n^{(1)}\\
E_n^{(0)}\underset{=0}{\left\langle\psi_n^{(0)}\middle|\psi_n^{(1)}\right\rangle} + \left\langle\psi_n^{(0)}\middle|\hat{H}^{(1)}\psi_n^{(0)}\right\rangle &= E_n^{(1)} \qquad E_n^{(0)} \text{ is real}
\end{aligned}$$

Specifically, the first order correction to the n th energy eigenvalue is the expectation value (in the n th zeroth order energy eigenstate) of the perturbation:

$$E_n^{(1)} = \left\langle \psi_n^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle . \tag{6.5}$$

The n th energy eigenvalue, estimated to first order in the perturbation, is just the expectation value of the perturbed Hamiltonian,

$$\begin{aligned} E_n &\cong E_n^{(0)} + E_n^{(1)} \\ &= \left\langle \psi_n^{(0)} \middle| \hat{H}^{(0)} \psi_n^{(0)} \right\rangle + \left\langle \psi_n^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle \\ &= \left\langle \psi_n^{(0)} \middle| \hat{H} \psi_n^{(0)} \right\rangle . \end{aligned}$$

We are done with Eq. 6.4. So far we have only considered the $\psi_n^{(0)}$ component of the equation. Now, we consider all other components. Specifically, take the inner product of both sides of Eq. 5 with $\psi_{n'}^{(0)}$, for all $n' \neq n$:

$$\begin{aligned} \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(0)} \psi_n^{(1)} + \hat{H}^{(1)} \psi_n^{(0)} \right\rangle &= \left\langle \psi_{n'}^{(0)} \middle| E_n^{(1)} \psi_n^{(0)} + E_n^{(0)} \psi_n^{(1)} \right\rangle \\ \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(0)} \psi_n^{(1)} \right\rangle + \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle &= E_n^{(1)} \underbrace{\left\langle \psi_{n'}^{(0)} \middle| \psi_n^{(0)} \right\rangle}_{=0} + E_n^{(0)} \underbrace{\left\langle \psi_{n'}^{(0)} \middle| \psi_n^{(1)} \right\rangle}_{=c_{n'}^{(1)}} \\ \left\langle \hat{H}^{(0)} \psi_{n'}^{(0)} \middle| \psi_n^{(1)} \right\rangle + \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle &= E_n^{(0)} c_{n'}^{(1)} \\ \left\langle E_{n'}^{(0)} \psi_{n'}^{(0)} \middle| \psi_n^{(1)} \right\rangle + \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle &= E_n^{(0)} c_{n'}^{(1)} \\ \left(E_n^{(0)} - E_{n'}^{(0)} \right) c_{n'}^{(1)} &= \left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle , \end{aligned}$$

or

$$c_{n'}^{(1)} = \frac{\left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_{n'}^{(0)}} .$$

The $c_{n'}^{(1)}$ are the expansion coefficients of $\psi_n^{(1)}$, expanded in terms of the $\psi_{n'}^{(0)}$, $n' \neq n$. Specifically,

$$\begin{aligned} \psi_n^{(1)} &= \sum_{n' \neq n} c_{n'}^{(1)} \psi_{n'}^{(0)} \\ &= \sum_{n' \neq n} \frac{\left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_{n'}^{(0)}} \psi_{n'}^{(0)} \\ &= - \sum_{n' \neq n} \frac{\left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle}{E_{n'}^{(0)} - E_n^{(0)}} \psi_{n'}^{(0)} \end{aligned} \tag{6.6}$$

Here, we see that the $n' \neq n$ contributions that are largest are those for which $E_{n'}^{(0)}$ is close to $E_n^{(0)}$ (and $\left\langle \psi_{n'}^{(0)} \left| \hat{H}^{(1)} \psi_n^{(0)} \right. \right\rangle$ is not zero). In fact, the formalism fails when there is degeneracy - i.e., when there are $n' \neq n$ values for $E_{n'}^{(0)} = E_n^{(0)}$ and the denominator of the $c_{n'}^{(1)}$ expression vanishes.

Degenerate case: Suppose there is a set of zero order eigenstates, $\left\{ \psi_k^{(0)} \right\}$, all associated the same zero order energy eigenvalue, $E_n^{(0)}$. In this case, we construct the matrix of $\hat{H}^{(1)}$ in the space spanned by the $\psi_k^{(0)}$. The k', k th element of this matrix is just

$$\left\langle \psi_{k'}^{(0)} \left| \hat{H}^{(1)} \psi_k^{(0)} \right. \right\rangle .$$

The eigenvalues of this perturbation matrix determine the first order shifts in the zero order eigenvalue, $E_n^{(0)}$. The perturbation can produce g distinct shifts for an eigenvalue with g-fold degeneracy, and thereby completely lift the degeneracy - splitting the degenerate level into distinct levels. In other cases, the degeneracy is partially lifted by the perturbation. The eigenvectors of the perturbation matrix determine the combinations of $\psi_k^{(0)}$ associated with the distinct first order shifts of the zero order energy - call these $\psi_j^{(0)}$. These combinations can then be used in the general perturbation described above and below. The denominator problem is removed because $\psi_j^{(0)}$ is an eigenfunction of the perturbation (in the $\left\{ \psi_k^{(0)} \right\}$ basis), and consequently, $\left\langle \psi_{j'}^{(0)} \left| \hat{H}^{(1)} \psi_j^{(0)} \right. \right\rangle = 0$, $j' \neq j$. Only non-zero denominators have non-zero coefficients.

Higher powers of λ: Equating coefficients of higher powers of λ give the remaining equations of perturbation theory. Here, we simply give the equation for the second order corrections to the energy eigenvalues:

$$E_n^{(2)} = -\sum_{n' \neq n} \frac{\left| \left\langle \psi_{n'}^{(0)} \left| \hat{H}^{(1)} \psi_n^{(0)} \right. \right\rangle \right|^2}{E_{n'}^{(0)} - E_n^{(0)}} \tag{6.7}$$

Each term in this sum is a correction to the n th energy level, associated with another level. The correction is the square of the perturbation matrix element connecting the two levels, divided by the energy difference between the levels. The sign of the difference is such that other levels - connected to the n th level by the perturbation - effectively *repel* the n th level. The shift is greatest coming from nearby energy levels which are well-connected to n th level - i.e., the perurbation matrix element is large. Thus, the next higher energy connected to the n th level shifts the n th level to lower energy, while the next lowest level connected to the n

th level shifts it to higher energy. Since there are no states with lower energy than the ground state, the second order correction to the ground state energy is always negative.

Example 6.1. Consider a particle in a one dimensional box - from $x = 0$ to L - subject to a perturbation,

$$\hat{H}^{(1)} = \begin{cases} \varepsilon, & 0 < x < L/2 \\ 0, & \text{otherwise} \end{cases},$$

$\varepsilon > 0$.

(a) Determine the energy levels of this particle, at the level of first order perturbation theory. Is the first order correction positive or negative? Does this make sense?

(b) What are the transition frequencies for the $n = 1$ to 2 and 2 to 3 transitions?

(c) Determine the first order correction to $\psi_1^{(0)}$ associated with the first excited state, $\psi_2^{(0)}$. This is the largest first order correction to $\psi_1^{(0)}$. Determine the associated correction to the ground state expectation of x. Does the sign of the correction make sense?

Solution 6.1.

(a) The zero order energy levels of the particle in a one dimensional box are

$$E_n^{(0)} = \frac{\hbar^2\pi^2 n^2}{2mL^2}.$$

The first order correction to the n th energy level is

$$\begin{aligned} \left\langle \psi_n^{(0)} \middle| \hat{H}^{(1)} \psi_n^{(0)} \right\rangle &= \frac{2}{L}\int_0^L \varepsilon\theta_{[0,L/2]}(x)\sin^2\left(\frac{n\pi x}{L}\right) dx \\ &= \frac{2\varepsilon}{L}\int_0^{L/2} \sin^2\left(\frac{n\pi x}{L}\right) dx \\ &= \frac{\varepsilon}{L}\int_0^{L/2}\left(1 - \cos\left(\frac{2n\pi x}{L}\right)\right) dx \\ &= \frac{\varepsilon}{L}\left(\frac{L}{2} - \left[\frac{L}{2n\pi}\sin\left(\frac{2n\pi x}{L}\right)\right]_0^{L/2}\right) \\ &= \frac{\varepsilon}{2}\left(1 - \frac{1}{n\pi}\left[\underbrace{\sin(n\pi)}_{=0} - 0\right]\right) \\ &= \frac{\varepsilon}{2}. \end{aligned}$$

Therefore, the energy levels corrected to first order are

$$E_n \cong \frac{\hbar^2\pi^2 n^2}{2mL^2} + \frac{\varepsilon}{2}.$$

The corrections are all positive - largest for $n = 1$ and smallest for $n = 2$. A positive correction makes sense because the perturbation is a positive potential energy. Also, the correction is $\varepsilon/2$, as the perturbation extends over only half the width of the box, and - in the zero order states - the particle is equally likely to be on the right side of the box (where the perturbation is) or the left side of the box.

(b) To first order, the first three energy levels are

$$E_1 \cong \frac{\hbar^2\pi^2}{2mL^2} + \frac{\varepsilon}{2},$$

$$E_2 \cong 4\frac{\hbar^2\pi^2}{2mL^2} + \frac{\varepsilon}{2}$$

and

$$E_3 \cong 9\frac{\hbar^2\pi^2}{2mL^2} + \frac{\varepsilon}{2}.$$

Therefore,

$$\nu_{1 \text{ to } 2} = \frac{\Delta E_{1 \text{ to } 2}}{h} = \frac{1}{h}3\frac{\hbar^2\pi^2}{2mL^2}$$

and

$$\nu_{2 \text{ to } 3} = \frac{\Delta E_{2 \text{ to } 3}}{h} = \frac{1}{h}5\frac{\hbar^2\pi^2}{2mL^2}.$$

The transition frequencies are not shifted by the perturbation (to first order) because all level are shifted by the same amount.

(c) The first order correction to the ground state is

$$\psi_1^{(1)} = -\sum_{n'>1} \frac{\left\langle \psi_{n'}^{(0)} \middle| \hat{H}^{(1)}\psi_1^{(0)} \right\rangle}{E_{n'}^{(0)} - E_1^{(0)}} \psi_{n'}^{(0)}.$$

Including only the first excited state contribution, which is the largest contribution, gives

$$\psi_1^{(1)} \cong -\frac{\left\langle \psi_2^{(0)} \middle| \hat{H}^{(1)}\psi_1^{(0)} \right\rangle}{E_2^{(0)} - E_1^{(0)}} \psi_2^{(0)}.$$

Since

$$\begin{aligned}
\left\langle \psi_2^{(0)} \middle| \hat{H}^{(1)}\psi_1^{(0)} \right\rangle &= \frac{2\varepsilon}{L}\int_0^{L/2} \sin\left(\frac{2\pi x}{L}\right)\sin\left(\frac{\pi x}{L}\right)\,dx \\
&= \frac{\varepsilon}{L}\int_0^{L/2}\left(\cos\left(\frac{\pi x}{L}\right) - \cos\left(\frac{3\pi x}{L}\right)\right)dx \\
&= \frac{\varepsilon}{L}\left(\left[\frac{L}{\pi}\sin\left(\frac{\pi x}{L}\right)\right]_0^{L/2} - \left[\frac{L}{3\pi}\sin\left(\frac{3\pi x}{L}\right)\right]_0^{L/2}\right)dx \\
&= \frac{\varepsilon}{\pi}\left(1 + \frac{1}{3}\right) = \frac{4\varepsilon}{3\pi},
\end{aligned}$$

$$\psi_1^{(1)} \cong -\left(\frac{2}{L}\right)^{1/2} \frac{4\varepsilon}{3\pi} \frac{\sin\left(\frac{2\pi x}{L}\right)}{3\frac{\hbar^2\pi^2}{2mL^2}},$$

and

$$\begin{aligned}\psi_1 &\cong \left(\frac{2}{L}\right)^{1/2}\left(\sin\left(\frac{\pi x}{L}\right) - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\sin\left(\frac{2\pi x}{L}\right)\right)\\ &= \psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)}.\end{aligned}$$

This state is not normalized. However, normalized, it takes the form,

$$\begin{aligned}\psi_1 &\cong \left(1 + \frac{64m^2L^4\varepsilon^2}{81\hbar^4\pi^6}\right)^{-1/2}\left(\psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)}\right)\\ &\cong \left(\psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)}\right)\left(1 - \frac{32m^2L^4\varepsilon^2}{81\hbar^4\pi^6}\right)\\ &\cong \psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)} \qquad \text{to first order in } \varepsilon\end{aligned}$$

The expectation of x for this state is

$$\begin{aligned}\langle\psi_1\,|x\psi_1\rangle &= \left\langle \psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)}\,\middle|\,x\left(\psi_1^{(0)} - \frac{8mL^2\varepsilon}{9\hbar^2\pi^3}\psi_2^{(0)}\right)\right\rangle\\ &\cong \left\langle\psi_1^{(0)}\,\middle|\,x\psi_1^{(0)}\right\rangle - \frac{16mL^2\varepsilon}{9\hbar^2\pi^3}\left\langle\psi_1^{(0)}\,\middle|\,x\psi_2^{(0)}\right\rangle \qquad \text{to first order in } \varepsilon\\ &= \frac{L}{2} - \frac{16mL^2\varepsilon}{9\hbar^2\pi^3}\left(-\frac{16L}{9\pi^2}\right) \qquad \text{see Eq. 3.9 and 3.10}\\ &= \frac{L}{2}\left(1 + \frac{512mL^2\varepsilon}{81\hbar^2\pi^5}\right).\end{aligned}$$

Here, we see that the perturbation leads to a positive correction to the ground state x expectation. The mean position of the particle in shifted to the right of the center of the box. This is because the perturbation is positive, and it is localized to the left side of the box. The perturbation repels the particle to the right side of the box. Had ε been negative, the expectation of x would be shifted to the left - the perturbation would attract the particle, rather than repel.

Example 6.2. Consider an electron in a one dimensional box of width, L, subject to the perturbation,

$$H^{(1)} = \mathcal{E}ex.$$

This is the interaction energy of the electron in a box, subject to electric field $\mathcal{E}$. This is the *Stark effect* for an electron in a box.

(a) What is the first order correction to the energy levels due to the electric field?

(b) What is the largest first order correction to the ground state due to the electric field? Does the sign of the correction make sense? Where is the electron most likely to be found, when in the ground state?

Solution 6.2.

(a) The first order correction to the n th energy level is

$$\begin{aligned} E_n^{(1)} &= \left\langle \psi_n^{(0)} \middle| H^{(1)} \psi_n^{(0)} \right\rangle \\ &= \mathcal{E} e \left\langle \psi_n^{(0)} \middle| x \psi_n^{(0)} \right\rangle \\ &= \mathcal{E} e \frac{L}{2}. \end{aligned}$$

The first order correction reflects the expectation value of the position, x. We found, in chapter 3, that the position expectation is $L/2$ for all states. The first order Stark effect for a electron in a box is just the classical electric potential energy of the electron is the center of the box. This correction is not zero because the electric potential is zero at $x = 0$, rather than the center of the box. The electric potential energy of an electron in the box is positive because the force on a negatively charged electron is to the left when the electric field is positive - the left side of the box has the lowest electric potential energy.

(b) The first order correction to the ground state is a sum over excited states with coefficients proportional to the matrix element, $\left\langle \psi_n^{(0)} \middle| x \psi_1^{(0)} \right\rangle$, and inversely proportional to the zero order energy difference, $E_n^{(0)} - E_1^{(0)}$. In Chapter 3, we evaluated this matrix element and found that only the even n elements are nonzero. By far, the largest matrix elements are for $\Delta n = \pm 1$. Thus, the largest - and only significant - first order correction to $\psi_1^{(0)}$ is from $n = 2$;

$$\begin{aligned} \psi_1 &\cong \psi_1^{(0)} - \frac{\left\langle \psi_2^{(0)} \middle| H^{(1)} \psi_1^{(0)} \right\rangle}{E_2^{(0)} - E_1^{(0)}} \\ &= \psi_1^{(0)} - \frac{\mathcal{E} e \left\langle \psi_2^{(0)} \middle| x \psi_1^{(0)} \right\rangle}{\frac{4\hbar^2\pi^2}{2m_e L^2} - \frac{\hbar^2\pi^2}{2m_e L^2}} \psi_2^{(0)} \\ &= \psi_1^{(0)} + \frac{\mathcal{E} e \frac{16L}{9\pi^2}}{\frac{4\hbar^2\pi^2}{2m_e L^2} - \frac{\hbar^2\pi^2}{2m_e L^2}} \psi_2^{(0)} \qquad \text{see Eq. 3.10} \\ &= \psi_1^{(0)} + \frac{32 \mathcal{E} e m_e L^3}{27 \hbar^2 \pi^4} \psi_2^{(0)}. \end{aligned}$$

> The sign of the correction is positive. Thus, $\psi_1^{(0)}$ and $\psi_2^{(0)}$ interfere constructively for $x < L/2$, and destructively for $x > L/2$. The position probability distribution of the corrected ground state is skewed to the left side of the box where the electric potential energy of the electron is lowest.

6.2 Anharmonicity

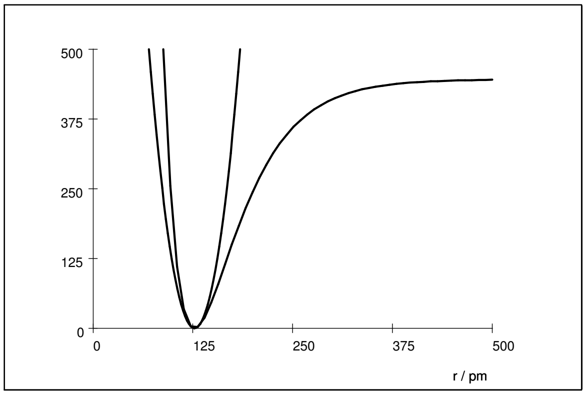
In Chapter 4, using a Taylor expansion of the vibrational potential, we obtained the harmonic oscillator as a model of molecular vibration, by only keeping terms up to x^2. The latter series has only one term, since we took $V(r_{\mathrm{eq}})$ to be zero (sets the energy reference value), and $dV(r)/dr = 0$ at r_{eq} where it has a minimum. The higher order terms in the Taylor series are generally not zero. They can be included using perturbation theory beause they constitute a perturbation to the harmonic oscillator Hamiltonian, and the harmonic oscillator system is completely solved.

Suppose the vibrational potential energy is approximated to fourth order in deviations from the equilibrium geometry - i.e., as a Taylor series in x up to the x^4 term. The vibrational Hamiltonian takes the form,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{k}{2}x^2 - \gamma x^3 + \delta x^4,$$

where the coefficient of x^3 is made explicitly negative ($\gamma > 0$) as the potential well typically becomes wider than the harmonic well for large positive x values.

Figure 6.1 shows a typical vibrational potential energy, together with the harmonic (quadratic) approximation. The realistic diatomic vibrational potential well is wider (than the harmonic well) for positive x, narrower for negative x, and it reaches an asymptotic value - the *equilibrium dissociation energy*, $D_{\mathrm{eq}} = D_0 + \hbar\omega/2$. The bond dissociation energy, D_0, is D_{eq} minus the zero point energy.



The Morse potential (solid line) for ^{1}H^{35}Cl. This is a more realistic model of a diatomic vibration potential than the harmonic potential (dashed line). The well depth here is $D_e = 446$ kJ mol^{-1}. The displacement is given in units of $\left(\frac{2D_e}{m\omega^2}\right)^{1/2} = 54$ pm for ^{1}H^{35}Cl.

Here, we consider the effect of anharmonicity at third and fourth order. These terms determine a perturbation of the harmonic oscillator Hamiltonian - the zero order Hamiltonian. To first order, the energy shift of the υ th state is given by

$$\begin{aligned} E_\upsilon^{(1)} &= \left\langle \psi_\upsilon^{(0)} \left| \hat{H}^{(1)} \psi_\upsilon^{(0)} \right. \right\rangle \\ &= \left\langle \psi_\upsilon^{(0)} \left| \left(-\gamma x^3 + \delta x^4\right) \psi_\upsilon^{(0)} \right. \right\rangle \\ &= -\gamma \left\langle \psi_\upsilon^{(0)} \left| x^3 \psi_\upsilon^{(0)} \right. \right\rangle + \delta \left\langle \psi_\upsilon^{(0)} \left| x^4 \psi_\upsilon^{(0)} \right. \right\rangle . \end{aligned}$$

We know that x can be written as a lowering operator plus a raising operator. x^3 consists of the eight possible combinations of raising or lowering three times, in succession. Since it is impossible to return to quantum number υ with an odd number of raising or lowerings, the matrix element $\left\langle \psi_\upsilon^{(0)} \left| x^3 \psi_\upsilon^{(0)} \right. \right\rangle = 0$ and no first order shift results from the third order anharmonicity. Third order anharmonicity does produce a second order shift in energy. This result is rationalized by noting that, while the third order anharmonicity widens the well on the right, it narrows the well on the left. The two effects partly balance, making the effect on energy more subtle. In contrast, the fourth order anharmonicity produces a first order correction to energy. This is because x^4 includes terms with two raising and two lowering operators, taking $\psi_\upsilon^{(0)}$ back to $\psi_\upsilon^{(0)}$ in order to give non-zero contributions to the $\psi_\upsilon^{(0)}$ expectation value of x^4. There are six such terms. The expectation value is positive for all υ. So the sign of the shift in energy is controlled by δ. Specifically, if $\delta < 0$, then the energy shift is negative. This corresponds to a perturbation that widens the well on both the left and right. If $\delta > 0$, then the energy shift is positive. This corresponds to a perturbation that narrows the well.

$$\begin{aligned}
E_\upsilon^{(1)} &= \delta \left\langle \psi_\upsilon^{(0)} \middle| x^4 \psi_\upsilon^{(0)} \right\rangle \\
&= \alpha^4 \delta \left\langle \psi_\upsilon^{(0)} \middle| y^4 \psi_\upsilon^{(0)} \right\rangle \\
&= \frac{\alpha^4 \delta}{4} \left\langle \psi_\upsilon^{(0)} \middle| \left(\hat{a} + \hat{a}^\dagger\right)^4 \psi_\upsilon^{(0)} \right\rangle \\
&= \frac{\alpha^4 \delta}{4} \Big((\upsilon+1)(\upsilon+2) + (\upsilon+1)^2 + (\upsilon+1)\upsilon + \upsilon(\upsilon+1) + \upsilon^2 + \upsilon(\upsilon-1) \Big) \\
&= \frac{\alpha^4 \delta}{4} \left(6\upsilon^2 + 6\upsilon + 3\right).
\end{aligned}$$

As realistic diatomic potentials widen faster than the associated harmonic approximation, the energy level spacings get smaller as υ increases. Moreover, there is always a dissociation threshold energy, D_{eq}, where the bond breaks. Consequently, the number of bound states is finite - with energies only up to D_{eq}. For the Morse potential, the energy levels are given - in wavenumbers - by

$$\bar{E}_{\mathrm{vib},\ \upsilon} = \frac{E_{\mathrm{vib},\ \upsilon}}{hc} = \left(\upsilon + \frac{1}{2}\right) \bar{\nu}_{\mathrm{vib}} - \left(\upsilon + \frac{1}{2}\right)^2 x_{\mathrm{eq}} \bar{\nu}_{\mathrm{vib}},$$

for

$$\upsilon = 1, 2, \ldots, \upsilon_{\max} = \left\lfloor \frac{1}{2x_{\mathrm{eq}}} \right\rfloor,$$

where

$$x_{\mathrm{eq}} = \frac{\bar{\nu}_{\mathrm{vib}}}{4\bar{D}_{\mathrm{eq}}}$$

is called the *anharmonicity constant*. For example, $x_{\mathrm{eq}} = 0.020$ for $^1\mathrm{H}^{35}\mathrm{Cl}$, and the Morse model predicts 25 bound vibrational states. Because the spacings are decreasing, hot bands appear at lower frequency than the fundamental transition, $\upsilon = 0$ to 1. In fact, the shift to lower frequency allows extraction of the anharmonicity constant from spectral data.

Anharmonicity also generally results in a shift in the mean bond distance - i.e., expectation of r - it increases with increasing υ. This increases the moment of inertia, decreasing the spacing between levels. Hot bands thus have narrower spacings between lines.

6.3 Rotational Zeeman and Stark effects

Here we consider a diatomic molecule in the gas phase. The temperature is such that excited vibrational states have negligible excitation. The diatomic molecule is essentially a rigid rotor. Two perturbations are considered; the interaction with a constant magnetic field - resulting in the rotational *Zeeman effect* - and the interaction with a constant electric field - resulting in the rotational *Stark effect*.

The unperturbed system here is the rigid rotor, with energy eigenstates given by the spherical harmonics, Y_{J,M_J}, with associated eigenvalues, $B_{\mathrm{rot}}J\,(J+1)$; i.e.,

$$\begin{aligned}\hat{H}^{(0)}Y_{J,M_J} &= \frac{1}{2I_{\mathrm{eq}}}\hat{J}^2 Y_{J,M_J}\\ &= B_{\mathrm{rot}}J\,(J+1)\,Y_{J,M_J}.\end{aligned}$$

The $J>0$ energy levels are degenerate. In general, degenerate perturbation theory is required for these level. First, we consider the Zeeman effect.

When a diatomic molecule rotates, the nuclear charges and the electron charges rotate. The resulting positive and negative current loops have magnetic moments. Because of the breakdown in the separation of motion of the electrons and the nuclei, and because of diamagnetic currents in the electron density induced by the magnetic field, there is a net rotational magnetic dipole moment,

$$\hat{\boldsymbol{\mu}}_m = g_{\mathrm{rot}}\mu_{m,N}\frac{\hat{\mathbf{J}}}{\hbar},$$

where

$$\mu_{m,N} = \frac{e\hbar}{2m_p}$$

is the *nuclear magneton* - the natural atomic unit of magnetic moment for nuclei. It corresponds to the rotation of two protons about their center of mass with an angular momentum of $\hbar$. The rotational g-factor of the molecule, g_{rot}, is just the magnetic dipole moment of the molecule, due to rotation, expressed in nuclear magnetons. If the molecule is closed shell - i.e., it has no unpaired electrons - there is no electron spin magnetic dipole moment. The electron spin magnetic moment would otherwise be the dominant component of total molecular magnetic moment. There can also be nuclear spin magnetic moments - see Sec. XX. Here, we consider only the interaction of the applied magnetic field, $\mathcal{B}$, with the rotational magnetic dipole - i.e., the rotational Zeeman effect,

$$\begin{aligned}\hat{H}^{(1)} &= -\hat{\boldsymbol{\mu}}_m\cdot\mathcal{B} = -\hat{\mu}_{m,z}\mathcal{B}_z\\ &= -\mathcal{B}_z g_{\mathrm{rot}}\mu_{m,N}\frac{\hat{J}_z}{\hbar},\end{aligned}$$

where we choose the z axis to be the direction of the electric field, $\mathcal{B}$.

The rigid rotor Hamiltonian is just

$$\hat{H}^{(0)} = \hat{J}^2/2I_{\mathrm{eq}}.$$

Since $\hat{J}^2$ and $\hat{J}_z$ commute, $\hat{H}^{(0)}$ and $\hat{H}^{(1)}$ commute, and have a common set of eigenfunctions. These are just the spherical harmonics. In this case, first order perturbation theory is exact. The degeneracy of the $J>0$ energy levels of the rigid rotor can be ignored because all off-diagonal matrix elements connecting distinct M_J states are zero;

$$\left\langle Y_{J,M_J}\left|\hat{H}^{(1)}Y_{J,M_J'}\right.\right\rangle = 0 \qquad \text{if } M_J' \neq M_J.$$

The energy eigenvalues are shifted by the expectation value of the Zeeman interaction,

$$\begin{aligned}
E^{(1)}_{\mathrm{rot},J,M_J} &= \left\langle Y_{J,M_J} \left| \hat{H}^{(1)} Y_{J,M_J} \right. \right\rangle \\
&= -\frac{\mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N}}{\hbar} \left\langle Y_{J,M_J} \left| \hat{J}_z Y_{J,M_J} \right. \right\rangle \\
&= -\frac{\mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N}}{\hbar} \left\langle Y_{J,M_J} \left| M_J \hbar Y_{J,M_J} \right. \right\rangle \qquad Y_{J,M_J} \text{ is an eigenfunction of } \hat{J}_z \\
&= -\mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N} M_J \underbrace{\left\langle Y_{J,M_J} \left| Y_{J,M_J} \right. \right\rangle}_{=1} \\
&= -\mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N} M_J .
\end{aligned}$$

Altogether,

$$\begin{aligned}
E_{\mathrm{rot},J,M_J} &= E^{(0)}_{\mathrm{rot},J,M_J} + E^{(1)}_{\mathrm{rot},J,M_J} \\
&= \frac{\hbar^2}{2I_{\mathrm{eq}}} J\left(J+1\right) - \mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N} M_J .
\end{aligned}$$

The magnetic field lifts the degeneracy associated with the M_J quantum number. For each J, there are $2J+1$ levels with equal spacing, $\mathcal{B}_z g_{\mathrm{rot}} \mu_{m,N}$. The lowest of these levels is the $M_J = J$ level. The highest is the $M_J = -J$ level. Transitions between these states are induced by radio frequency pulses such that the associated oscillating magnetic field is in the x y plane - suppose it is along x. The Zeeman interaction of the molecular magnetic dipole with this oscillating field introduces the operator $\hat{J}_x = \left(\hat{J}_+ + \hat{J}_-\right)/2$, which produces magnetic dipole transitions with the selection rule,

$$\Delta M_J = \pm 1.$$

The interaction of the rigid rotor with an electric field - the Stark effect - is given by

$$\begin{aligned}
\hat{H}^{(1)} &= -\hat{\boldsymbol{\mu}}_e \cdot \boldsymbol{\mathcal{E}} = -\hat{\mu}_{e,z} \mathcal{E}_z \\
&= -\mathcal{E}_z \mu_{e,\mathrm{eq}} \cos\theta ,
\end{aligned}$$

where the electric field, $\mathcal{E}$, is along the z axis. This effect requires that the molecule have a permanent electric dipole moment, $\mu_{e,\mathrm{eq}}$. Since this operator affects only functions of θ, it cannot connect distinct M_J states. Degenerate perturbation is again unnecessary for the $J > 0$ levels. First consider the first order correction of the 0, 0 level. Specifically,

$$
\begin{aligned}
E_{0,0}^{(1)} &= \left\langle \psi_{0,0}^{(0)} \left| \hat{H}^{(1)} \psi_{0,0}^{(0)} \right.\right\rangle \\
&= \left\langle Y_{0,0} \left| \mathcal{E}_z \mu_{e,\mathrm{eq}} \cos\theta Y_{0,0} \right.\right\rangle \\
&= \mathcal{E}_z \mu_{e,\mathrm{eq}} \left\langle \Theta_{0,0} \left| \cos\theta \Theta_{0,0} \right.\right\rangle \underbrace{\left\langle \Phi_0 \left| \Phi_0 \right.\right\rangle}_{=1} \\
&= \mathcal{E}_z \mu_{e,\mathrm{eq}} \int_0^\pi \cos\theta \left(\frac{1}{\sqrt{2}}\right)^2 \sin\theta \, d\theta \qquad \Theta_{0,0} = \frac{1}{\sqrt{2}} \\
&= \frac{\mathcal{E}_z \mu_{e,\mathrm{eq}}}{4} \int_0^\pi \sin 2\theta \, d\theta \\
&= \frac{\mathcal{E}_z \mu_{e,\mathrm{eq}}}{8} \left[-\cos 2\theta\right]_0^\pi \\
&= 0.
\end{aligned}
$$

There is no first order energy correction for the 0, 0 level. In fact, all first order energy corrections are zero. This is proven by noting that the Y_{J,M_J} are eigenfunctions of reflection through the $x\ y$ plane. This follows because the rigid rotor Hamiltonian commutes with the reflection operator. Since applying the reflection operator twice is just the identity operator, the eigenvalues of the reflection operator are $+1$ and -1; i.e., Y_{J,M_J} are either *even* or *odd* with respect to reflection through the $x\ y$ plane. The sign of the reflection eigenvalue is called the *parity* of Y_{J,M_J}. In terms of spherical coordinates, reflection through the $x\ y$ plane corresponds to the the transformation $\theta \to \pi - \theta$. Since $\Theta_{J,M_J}(\theta)$ has well defined parity under, $\theta \to \pi - \theta$, its square is even under this transformation. The remaining factor in the integrand of the θ integral, that contributes to $E_{J,M_J}^{(1)}$, is $\cos\theta \sin\theta$. This function is odd (it reverses sign) under $\theta \to \pi - \theta$. Thus, the integrand of the θ integral is odd, and the θ integral is zero - the 0 to $\pi/2$ portion of the integral cancels the $\pi/2$ to π portion.

The Stark effect is second order - i.e., the leading correction to the energy eigenvalues is the second order correction. Here, we evaluate the second order energy correction for the ground and first excited states. We also determine the first order corrected eigenstates - note that a second order perturbation such as the Stark effect still has first order corrections to the eigenstates.

For the Stark effect in the z direction, being considered here, the energy levels remain independent of M_J - they retain their degeneracy. This is because all matrix elements of the perturbation connecting distinct M_J states are zero. The second order corrected rotational Stark effect ground state energy is (see

Eq. 6.7)

$$\begin{aligned}
E_{\text{rot},0} &\cong E^{(0)}_{\text{rot},0} + E^{(2)}_{\text{rot},0} \\
&= 0 - \sum_{J>0} \frac{\left|\langle Y_{J,0} \left| \hat{H}^{(1)} Y_{0,0} \right\rangle\right|^2}{E_{\text{rot},\ J} - E_{\text{rot},\ 0}} \\
&= -\mathcal{E}_z^2 \mu_{e,\text{eq}}^2 \sum_{J>0} \frac{\left|\langle Y_{J,0} \left|\cos\theta Y_{0,0}\right\rangle\right|^2}{B_{\text{rot}} J\left(J+1\right) - 0} \\
&= -\frac{\mathcal{E}_z^2 \mu_{e,\text{eq}}^2}{B_{\text{rot}}} \sum_{J>0} \frac{\left|\langle Y_{J,0} \left|\cos\theta Y_{0,0}\right\rangle\right|^2}{J\left(J+1\right)}.
\end{aligned}$$

Since

$$Y_{0,0}\left(\theta,\phi\right) = \left(4\pi\right)^{-1/2}$$

and

$$Y_{1,0}\left(\theta,\phi\right) = \left(\frac{4\pi}{3}\right)^{-1/2} \cos\theta,$$

$$\begin{aligned}
\cos\theta Y_{0,0}\left(\theta,\phi\right) &= \left(4\pi\right)^{-1/2} \cos\theta \\
&= 3^{-1/2} Y_{1,0}\left(\theta,\phi\right).
\end{aligned}$$

Thus, only the $J, M_J = 1, 0$ term contributes to the sum, and

$$\begin{aligned}
E_{0,0} &\cong -\frac{\mathcal{E}_z^2 \mu_{e,\text{eq}}^2}{B_{\text{rot}}} \frac{\left|\langle Y_{1,0} \left|\cos\theta Y_{0,0}\right\rangle\right|^2}{1\left(1+1\right)} \\
&= -\frac{\mathcal{E}_z^2 \mu_{e,\text{eq}}^2}{2B_{\text{rot}}} \left|\langle Y_{1,0} \left| 3^{-1/2} Y_{1,0}\left(\theta,\phi\right)\right\rangle\right|^2 \\
&= -\frac{\mathcal{E}_z^2 \mu_{e,\text{eq}}^2}{6B_{\text{rot}}}.
\end{aligned}$$

The ground state energy is lowered in proportion to the square of the electric field.

The associated first order corrected eigenfunction is (see Eq. 6.6)

$$\begin{aligned}
\psi_{0,0}\left(\theta,\phi\right) &\cong Y_{0,0}\left(\theta,\phi\right) - \frac{\langle Y_{1,0} \left| \hat{H}^{(1)} Y_{0,0} \right\rangle}{E_{\text{rot},\ 1} - E_{\text{rot},\ 0}} Y_{1,0}\left(\theta,\phi\right) \\
&= Y_{0,0}\left(\theta,\phi\right) + \frac{\mathcal{E}_z \mu_{e,\text{eq}}}{2B_{\text{rot}}} \langle Y_{1,0} \left|\cos\theta Y_{0,0}\right\rangle Y_{1,0}\left(\theta,\phi\right) \\
&= Y_{0,0}\left(\theta,\phi\right) + \frac{\mathcal{E}_z \mu_{e,\text{eq}}}{3^{1/2} 2B_{\text{rot}}} Y_{1,0}\left(\theta,\phi\right).
\end{aligned}$$

The ground state eigenstate is perturbed in proportion to the ratio of the electric dipole interaction energy and the energy spacing between the first two levels.

It acquires a $Y_{1,0}(\theta,\phi)$ component. $Y_{0,0}$ is everywhere positive, while $Y_{1,0}$ is positive along the positive z axis and negative along the negative z axis. Consequently, the two components of $\psi_{0,0}$ interfere constructively along the positive z axis, and destructively along the negative z axis. The ground state probability distribution for the orientation of the molecule is shifted towards alignment with the electric field.

Since

$$Y_{2,0}(\theta,\phi)=\left(\frac{16\pi}{5}\right)^{-1/2}\left(3\cos^2\theta-1\right),$$

$$\begin{aligned}
&\cos\theta Y_{1,0}(\theta,\phi)\\
= & \left(\frac{4\pi}{3}\right)^{-1/2}\cos^2\theta\\
= & \frac{1}{3}\left(\frac{5}{12}\right)^{-1/2}\left(\frac{16\pi}{5}\right)^{-1/2}\left(3\cos^2\theta-1\right)+\frac{1}{3}\left(\frac{5}{12}\right)^{-1/2}\left(\frac{16\pi}{5}\right)^{-1/2}\\
= & \left(\frac{4}{15}\right)^{1/2}Y_{2,0}(\theta,\phi)+3^{1/2}Y_{0,0}(\theta,\phi)
\end{aligned}$$

is a linear combination of $Y_{2,0}$ and $Y_{0,0}$, there are two contributions to the second order energy and first order eigenstate corrections for the $1,0$ state - matrix elements with any other Y_{J,M_J} are zero.

$$\begin{aligned}
E_{1,0}^{(2)} &\cong -\frac{\mathcal{E}_z^2\mu_{e,\mathrm{eq}}^2}{B_{\mathrm{rot}}}\left(\frac{\left|\langle Y_{0,0}\left|\cos\theta Y_{1,0}\right\rangle\right|^2}{0-1(1+1)}+\frac{\left|\langle Y_{2,0}\left|\cos\theta Y_{1,0}\right\rangle\right|^2}{2(2+1)-1(1+1)}\right)\\
&= -\frac{\mathcal{E}_z^2\mu_{e,\mathrm{eq}}^2}{B_{\mathrm{rot}}}\left(-\frac{\left|3^{1/2}\right|^2}{2}+\frac{\left|\left(\frac{4}{15}\right)^{1/2}\right|^2}{4}\right)\\
&= \frac{\mathcal{E}_z^2\mu_{e,\mathrm{eq}}^2}{B_{\mathrm{rot}}}\left(\frac{3}{2}-\frac{1}{15}\right)=\frac{43\mathcal{E}_z^2\mu_{e,\mathrm{eq}}^2}{30B_{\mathrm{rot}}}.
\end{aligned}$$

The $1,0$ state is shifted to higher energy because of the coupling to $0,0$. The $0,0$ level is lowered, while the $1,0$ level is raised.

The $1,0$ excited state is

$$\begin{aligned}
&\psi_{1,0}(\theta,\phi)\\
\cong\ & Y_{1,0}(\theta,\phi)-\frac{\left\langle Y_{0,0}\left|\hat{H}^{(1)}Y_{1,0}\right.\right\rangle}{E_{\mathrm{rot},\ 0}-E_{\mathrm{rot},\ 1}}Y_{0,0}(\theta,\phi)-\frac{\left\langle Y_{2,0}\left|\hat{H}^{(1)}Y_{1,0}\right.\right\rangle}{E_{\mathrm{rot},\ 2}-E_{\mathrm{rot},\ 1}}Y_{2,0}(\theta,\phi)\\
=\ & Y_{1,0}(\theta,\phi)-\frac{\mathcal{E}_z\mu_{e,\mathrm{eq}}}{B_{\mathrm{rot}}}\left(\frac{\langle Y_{0,0}\left|\cos\theta Y_{1,0}\right\rangle Y_{0,0}(\theta,\phi)}{2}-\frac{\langle Y_{2,0}\left|\cos\theta Y_{1,0}\right\rangle Y_{2,0}(\theta,\phi)}{4}\right)\\
=\ & Y_{1,0}(\theta,\phi)-\frac{\mathcal{E}_z\mu_{e,\mathrm{eq}}}{2B_{\mathrm{rot}}}\left(3^{1/2}Y_{0,0}(\theta,\phi)-15^{-1/2}Y_{2,0}(\theta,\phi)\right).
\end{aligned}$$

The state is shifted away from alignment with the electric field.

For the $1,1$ excited state, since

$$Y_{1,1}(\theta,\phi) = -\left(\frac{8\pi}{3}\right)^{-1/2} \sin\theta \exp(i\phi)$$

and

$$Y_{2,1}(\theta,\phi) = -\left(\frac{8\pi}{15}\right)^{-1/2} \sin\theta\cos\theta \exp(i\phi),$$

we have

$$\cos\theta Y_{1,1}(\theta,\phi) = \frac{1}{5^{1/2}} Y_{2,1}(\theta,\phi)$$

and

$$\begin{aligned} E_{1,1}^{(2)} &\cong -\frac{\mathcal{E}_z^2 \mu_{e,\mathrm{eq}}^2}{B_{\mathrm{rot}}} \frac{\left|\langle Y_{2,1} \left|\cos\theta Y_{1,1}\right\rangle\right|^2}{2(2+1) - 1(1+1)} \\ &= -\frac{\mathcal{E}_z^2 \mu_{e,\mathrm{eq}}^2}{B_{\mathrm{rot}}} \left(\frac{\left|\frac{1}{5^{1/2}}\right|^2}{4}\right) \\ &= -\frac{\mathcal{E}_z^2 \mu_{e,\mathrm{eq}}^2}{20 B_{\mathrm{rot}}}. \end{aligned}$$

The $1,1$ level is lowered in energy - by a much smaller amount than $0,0$. The $1,-1$ level is lowered by the same amount. Here we see that the electric field partially lefts the M_J degeneracy. In general, the energy depends upon J and $|M_J|$. Levels with $M_J \neq 0$ are doubly degenerate.

There is a general recursion formula that gives all of the spherical harmonic matrix elements of $\cos\theta$. It shows that $\cos\theta$ only connects J states with $\Delta J = \pm 1$. This is the origin of the selection rule for rotational spectroscopy - see Sec. 5.3 - that arises when the electric field is oscillating. Had we considered an electric field in the x or y directions, the perturbation would be proportional to $\sin\theta\cos\phi$ or $\sin\theta\sin\phi$. The $\sin\theta$ operator connects only states with $\Delta J = \pm 1$, just like $\cos\theta$. The $\cos\phi$ or $\sin\phi$ operator connects states with $\Delta M_J = \pm 1$, since the trigonometric functions are plus and minus combinations of $\exp(i\phi)$ and $\exp(-i\phi)$. These operators act as raising and lowering operators for Φ_{M_J} functions. This is the origin of the $\Delta M_J = 0, \pm 1$ selection rule. It accounts for electric fields in all directions, since incident light can have have different angle of incidence and polarization - i.e., different electric field directions.

6.4 Time Dependent Perturbation Theory

Perturbation theory can also be adapted to the solution of the *time dependent Schrödinger equation.* This is how we accommodate an oscillating time dependent electric field and explain light induced transitions.

Suppose that a system is governed by a Hamiltonian, $\hat{H}^{(0)}$, with know eigenvalues and eigenfunctions (possibly generated numerically - the system need not be solvable analytically). If the system is initially (i.e., at $t = 0$) in state $\psi_n(x)$, an eigenstate of $\hat{H}^{(0)}$, then the solution to the time dependent Schrödinger equation is just

$$\Psi^{(0)}(x,t) = \exp\left(\frac{-iE_n t}{\hbar}\right)\psi_n(x),$$

where E_n is the energy eigenvalue associated with $\psi_n(x)$, and x is the system coordinate - or set of coordinates. Here, we see that the wavefunction changes with time only via a phase factor. In this case, observable properties of the system do not change with time. Eigenstates of the system Hamiltonian are called *stationary states.*

Now consider a time dependent perturbation, $\hat{H}^{(1)}(t)$, of the above system. The perturbation is applied (turned on) at $t = 0$, when the system state is just $\psi_n(x)$. As time progresses, the perturbation can cause the state of the system to spread over other eigenstates of the unperturbed system. The spreading is made explicit by expressing the time dependent wavefunction as a sum over unperturbed system eigenstates,

$$\Psi(x,t) = \sum_{n'} c_{n'}(t)\,\psi_{n'}(x),$$

with time dependent expansion coefficients. This expression is completely general. It does not depend on the perturbation being small. The set of $\psi_{n'}(x)$ eigenstates provide a complete basis - any wavefunction can be expanded in terms of these basis functions.

Substituting the above equation into the time dependent Schrödinger equation gives

$$\begin{aligned}
i\hbar\frac{\partial\Psi(x,t)}{\partial t} &= \hat{H}(t)\,\Psi(x,t) \\
&= \left(\hat{H}^{(0)} + \hat{H}^{(1)}(t)\right)\Psi(x,t) \\
i\hbar\frac{\partial}{\partial t}\sum_{n'} c_{n'}(t)\,\psi_{n'}(x) &= \left(\hat{H}^{(0)} + \hat{H}^{(1)}(t)\right)\sum_{n'} c_{n'}(t)\,\psi_{n'}(x) \\
i\hbar\sum_{n'}\frac{\partial c_{n'}(t)}{\partial t}\psi_{n'}(x) &= \sum_{n'} c_{n'}(t)\left(\hat{H}^{(0)} + \hat{H}^{(1)}(t)\right)\psi_{n'}(x) \\
&= \sum_{n'} c_{n'}(t)\left(E_{n'} + \hat{H}^{(1)}(t)\right)\psi_{n'}(x).
\end{aligned}$$

Take the inner product of both sides of the last line with $\psi_n(x)$. The result is

$$\begin{aligned}
i\hbar\sum_{n'}\frac{\partial c_{n'}(t)}{\partial t}\underset{\substack{=1 \text{ if } n'=n\\ =0 \text{ otherwise}}}{\langle\psi_n|\psi_{n'}\rangle} &= \sum_{n'} c_{n'}(t)\left\langle\psi_n\left|\left(E_{n'}+\hat{H}^{(1)}(t)\right)\psi_{n'}\right.\right\rangle \\
i\hbar\frac{\partial c_n(t)}{\partial t} &= \sum_{n'} c_{n'}(t)\left(E_{n'}\langle\psi_n|\psi_{n'}\rangle+\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle\right) \\
&= E_n c_n(t)+\sum_{n'}\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle c_{n'}(t). \qquad (6.8)
\end{aligned}$$

The TDSE becomes a set of coupled differential equations for the time dependent expansion coefficients. These equations are further illuminated by introducing the coefficients, $b_n(t)$, such that

$$c_n(t)=\exp\left(\frac{-iE_nt}{\hbar}\right)b_n(t).$$

These new coefficients capture any deviation from the unperturbed system time evolution. Substituting the above expression for the coefficients into the right side of Eq. 6.8 gives

$$i\hbar\frac{\partial}{\partial t}\left(\exp\left(\frac{-iE_nt}{\hbar}\right)b_n(t)\right)=i\hbar\left(\left(\frac{-iE_n}{\hbar}\right)b_n(t)+\frac{\partial b_n(t)}{\partial t}\right)\exp\left(\frac{-iE_nt}{\hbar}\right).$$

The left side takes the form,

$$E_n\exp\left(\frac{-iE_nt}{\hbar}\right)b_n(t)+\sum_{n'}\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle\exp\left(\frac{-iE_{n'}t}{\hbar}\right)b_{n'}(t)$$

Multiplying both sides of Eq. 6.8 by $\exp\left(\frac{iE_nt}{\hbar}\right)$ gives

$$\begin{aligned}
i\hbar\left(\left(\frac{-iE_n}{\hbar}\right)b_n(t)+\frac{\partial b_n(t)}{\partial t}\right) &= E_nb_n(t)+\sum_{n'}\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle\exp\left(\frac{i(E_n-E_{n'})t}{\hbar}\right)b_{n'}(t) \\
E_nb_n(t)+i\hbar\frac{\partial b_n(t)}{\partial t} &= E_nb_n(t)+\sum_{n'}\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle\exp(i\omega_{n,n'}t)\,b_{n'}(t),
\end{aligned}$$

or

$$i\hbar\frac{\partial b_n(t)}{\partial t}=\sum_{n'}\left\langle\psi_n\left|\hat{H}^{(1)}(t)\psi_{n'}\right.\right\rangle\exp(i\omega_{n,n'}t)\,b_{n'}(t), \qquad (6.9)$$

where

$$\omega_{n,n'}=\frac{E_n-E_{n'}}{\hbar}=2\pi\frac{\Delta E_{n,n'}}{h}$$

is the angular transition frequency for n' to n transitions. It is $2\pi\times$ the frequency of light with photon energy that matches the change in energy associated with the transition. These equations can be solved numerically for any time dependent perturbation. However, it is instructive to construct a perturbative solution with successive higher order corrections. In the time dependent case, the perturbation order parameter is time itself - the longer the perturbation acts, the greater its effect. With this in mind, we solve Eq. 6.9 iteratively.

Integrating Eq. 6.9 from 0 to t gives

$$\begin{aligned} i\hbar\int_0^t \frac{\partial b_n(t')}{\partial t'}dt' &= \int_0^t \sum_{n'} \left\langle \psi_n \left| \hat{H}^{(1)}(t')\, \psi_{n'} \right. \right\rangle \exp\left(i\omega_{n,n'}t'\right) b_{n'}(t')\, dt' \\ i\hbar\left(b_n(t) - b_n(0)\right) &= \sum_{n'} \int_0^t \left\langle \psi_n \left| \hat{H}^{(1)}(t')\, \psi_{n'} \right. \right\rangle \exp\left(i\omega_{n,n'}t'\right) b_{n'}(t')\, dt', \end{aligned}$$

or

$$b_n(t) = b_n(0) - \frac{i}{\hbar}\sum_{n'} \int_0^t \left\langle \psi_n \left| \hat{H}^{(1)}(t')\, \psi_{n'} \right. \right\rangle \exp\left(i\omega_{n,n'}t'\right) b_{n'}(t')\, dt'.$$

The first term on the right in the above equation constitutes the zero order approximation - wherein the b_n coefficients do not change with time. The first order approximation is obtained if the zero order approximation of $b_n(t')$ is used in the integrand - i.e., let $b_n(t') = b_n(0)$ (note that $t' < t$) The result is

$$\begin{aligned} b_n^{(1)}(t) &= b_n(0) - \frac{i}{\hbar}\sum_{n'} \int_0^t \left\langle \psi_n \left| \hat{H}^{(1)}(t')\, \psi_{n'} \right. \right\rangle \exp\left(i\omega_{n,n'}t'\right) b_{n'}(0)\, dt' \\ &= b_n(0) - \frac{i}{\hbar}\sum_{n'} \int_0^t \left\langle \psi_n \left| \hat{H}^{(1)}(t')\, \psi_{n'} \right. \right\rangle \exp\left(i\omega_{n,n'}t'\right) dt' b_{n'}(0) \qquad (6.10) \end{aligned}$$

The second order approximation results if the first order approximation is substituted into the integral - and so forth.

6.4.1 Electric dipole transitions

We consider the first order approximation to the expansion coefficients in case of incident light - i.e., an electromagnetic field - interacting with the electric dipole, $\hat{\boldsymbol{\mu}}$, of the system. The time dependent perturbation is the interaction energy,

$$\hat{H}^{(1)}(t) = \hat{\boldsymbol{\mu}}_e \cdot \mathcal{E}(t) = \hat{\mu}_{e,z}\mathcal{E}_z(t),$$

where we choose the z axis to be the direction of the time dependent electric field, $\mathcal{E}(t)$, associated with the incident electromagnetic radiation. The electric field oscillates with angular frequency, $\omega = 2\pi\nu$;

$$\begin{aligned} \mathcal{E}_z(t) &= \mathcal{E}_z(0)\cos(\omega t) \\ &= \frac{1}{2}\mathcal{E}_z(0)\left(\exp(i\omega t) + \exp(-i\omega t)\right). \end{aligned}$$

Consider just the second of the two complex exponential terms. Substituting the associated time dependent perturbation into Eq. 6.10 gives

$$\begin{aligned}
b_n^{(1)}(t) &= b_n(0) - \frac{i}{\hbar}\sum_{n'}\int_0^t \left\langle \psi_n \left| \frac{1}{2}\mathcal{E}_z(0)\exp(-i\omega t')\,\hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle \exp(i\omega_{n,n'}t')\,dt' b_{n'}(0) \\
&= b_n(0) - \frac{i}{\hbar}\frac{1}{2}\mathcal{E}_z(0)\sum_{n'}\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle \int_0^t \exp\left(i(\omega_{n,n'}-\omega)t'\right)dt' b_{n'}(0) \\
&= b_n(0) - \frac{i}{\hbar}\frac{1}{2}\mathcal{E}_z(0)\sum_{n'}\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle \left[\frac{\exp\left(i(\omega_{n,n'}-\omega)t'\right)}{i(\omega_{n,n'}-\omega)}\right]_0^t b_{n'}(0) \\
&= b_n(0) - \frac{1}{\hbar}\frac{1}{2}\mathcal{E}_z(0)\sum_{n'}\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle \left[\frac{\exp\left(i(\omega_{n,n'}-\omega)t\right)-1}{\omega_{n,n'}-\omega}\right] b_{n'}(0) \\
&= b_n(0) - \frac{1}{\hbar}\frac{1}{2}\mathcal{E}_z(0)\sum_{n'}\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle F_{n,n'}(t)\, b_{n'}(0)
\end{aligned}$$

The factor in square brackets, $F_{n,n'}(t)$, oscillates with time at high frequency, unless ω is close to $\omega_{n,n'}$. Suppose we tune the frequency offset, $\Delta\omega = (\omega_{n,n'} - \omega)$, close to zero for some n and n', such that all other frequency offsets, $\Delta\omega'$, are much larger. If $\Delta\omega$ is very small, then $\exp(i\Delta\omega t) \cong 1 + i\Delta\omega t$, for t up to some fraction of $\Delta\omega^{-1}$, and

$$\frac{\exp(i\Delta\omega t)-1}{\Delta\omega} \cong \frac{i\Delta\omega t}{\Delta\omega} = it.$$

This term increases with time - it is the *resonant transition* term. The other terms do not increase with time. They oscillate rapidly, and they are much smaller than the resonant term. This includes terms that arise from the $\exp(i\omega t)$ component of the electric field that we left out. for nonresonant terms, $|F_{n,n''}(t)|$ oscillates between 0 and $2/|\omega_{n,n''} - \omega| \cong 2/|\omega_{n,n''} - \omega_{n,n'}|$. Henceforth, we consider only the much larger resonant term in the above equation. Specifically,

$$b_n(t) \cong b_n(0) - \frac{1}{\hbar}\frac{F_{n,n'}(t)}{2}\mathcal{E}_z(0)\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle b_{n'}(0).$$

If the system is initially in state, $\psi_{n'}$ - i.e.,

$$b_{n'}(0) = 1$$

and

$$b_n(0) = 0$$

- then at time t,

$$b_n(t) = -\frac{i}{\hbar}\frac{F_{n,n'}(t)}{2}\mathcal{E}_z(0)\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle \tag{6.11}$$

If $t < 1/\Delta\omega$,

$$b_n(t) = -\frac{t}{2\hbar}\mathcal{E}_z(0)\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'}\right.\right\rangle,$$

and the coefficient of ψ_n grows linearly with rate,

$$r = -\frac{i}{2\hbar}\mathcal{E}_z(0)\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'} \right. \right\rangle,$$

until the time approaches $1/\Delta\omega$ (at and beyond this time, the above linear approximation for the exponential is not valid). The probability that a transition occurs is given by $|b_n(t)|^2$. For $t < 1/\Delta\omega$,

$$|b_n(t)|^2 = \frac{t^2}{4\hbar^2}\mathcal{E}_z^2(0)\left|\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'} \right. \right\rangle\right|^2.$$

$\mathcal{E}_z^2(0)$ is proportional to the energy density of the incident light - i.e., the *light intensity.*

The above analysis applies for near resonant light with a *precise* frequency. In pactice, light sources have a *distribution* of frequencies. In realizable monochromatic light, the energy density is actually spread out over a narrow band of frequencies. The above analysis is adapted to this case by including - on the right side of Eq. **??** - an integral over ω (which easily converts to an integral over $\Delta\omega$), across the associated band of frequencies. After quadratic growth at short times, the transition probability grows linearly with time with rate given by *Fermi's golden rule.* Specifically,

$$|b_n(t)|^2 = \frac{2\pi}{\hbar}\rho(\omega_{n,n'})\left|\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'} \right. \right\rangle\right|^2 t, \tag{6.12}$$

where $\rho(\omega_{n,n'})$ is the density of light energy - per unit frequency - evaluated at the transition frequency. The transition probability grows with a constant rate,

$$\frac{2\pi}{\hbar}\rho(\omega_{n,n'})\left|\left\langle \psi_n \left| \hat{\mu}_{e,z}\psi_{n'} \right. \right\rangle\right|^2.$$

Fermi's golden rule, Eq. 6.12 , is valid only up to first order in time dependent perturbation theory. Specifically, it is valid only as long as the transition probability is small. For low powered lasers, or ordinary ambient light, first order perturbation theory is widely applicable to induced transitions. Transitions with low probability are seen because there are large number of molecules. For high powered lasers, the above coupled differential equations must be solved for the coefficients. Accurate calcuations might also require higher order - quadrupolar, octapolar, etc. - terms, beyond the electric dipole interaction.

The above analysis does not depend upon the sign of $\omega_{n',n}$. It describes *absorption* transitions, when n' labels the lower energy level, and *stimulated* emission when n' labels the higher state.

Chapter 7

Electronic structure of atoms

7.1 The hydrogen atom

The hydrogen atom, or hydrogen-like ion (He^+, Li^{2+},...), is the simplest atom (ion) with only one electron. It is a two body problem just like the diatomic molecule. Like the diatomic molecule, the six degrees of freedom separate into three center of mass coordinates and three relative coordinates. The difference here is that the nucleus is much heavier than the electron (e.g., proton mass is 1836× electron mass). As such the center of mass is very close to the nucleus and the center of mass motion looks very much like the motion of the nucleus. At the same time, the reduced mass is very close to the electron mass, and the relative motion is almost the motion of the electron. It is thus convenient to think of the relative motion as the motion of the electron - though accurate calculations require using the correct reduced masss. Just as in the diatomic molecule case, the relative motion further separates into rotational and radial motion. The radial TISE takes the form,

$$\left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+V\left(r\right)+\frac{1}{2mr^2}\hat{\mathbf{L}}^2\right)R\left(r\right)=ER\left(r\right),$$

where

$$\hat{\mathbf{L}}^2 Y_{\ell,m_\ell}\left(\theta,\phi\right)=\ell\left(\ell+1\right)\hbar^2 Y_{\ell,m_\ell}\left(\theta,\phi\right),$$

and

$$V\left(r\right)=-\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r}$$

is the Coulombic attraction between the electron and nucleus - the latter with charge $+Ze$ to include hydrogen-like ions. The complete relative motion wavefunction is the product,

$$R\left(r\right)Y_{\ell,m_\ell}.\left(\theta,\phi\right).$$

As in the treatment of diatomic molecule vibrations, we transform into radial TISE into the TISE of a particle in one dimension. Specifically,

$$R(r) = \frac{1}{r}\varphi(r),$$

gives

$$\left(-\frac{\hbar^2}{2m_e}\frac{d^2}{dr^2} - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} + \frac{\ell(\ell+1)\hbar^2}{2m_e r^2}\right)\varphi(r) = E\varphi(r).$$

Here, we approximate the reduced mass as the mass of an electron. The solutions we get are more useful as starting points in the treatment of many electron atoms and molecules in Chapter 8.

The Coulomb potential does not limit the range of motion to a narrow interval about some equilibrium value. Instead, $V(r)$ is pure attractive right to $r = 0$. As such, radial motion extends over a wide range. The rigid rotor approximation is not relevant here - i.e., the centrifugal potential cannot be treated as independent of r. As a result, we must solve the radial TISE with r^{-1} and r^{-2} potential terms,

$$\left(-\frac{\hbar^2}{2m_e}\frac{d^2}{dr^2} - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} + \frac{\ell(\ell+1)\hbar^2}{2m_e r^2}\right)\varphi(r) = E\varphi(r)$$

or

$$\begin{aligned}\left(-\frac{1}{2}\frac{d^2}{dr^2} - \frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\frac{Z}{r} + \frac{\ell(\ell+1)}{2r^2}\right)\varphi(r) &= \frac{m_e E}{\hbar^2}\varphi(r)\\ \left(-\frac{1}{2}\frac{d^2}{dr^2} - \frac{1}{a_0}\frac{Z}{r} + \frac{\ell(\ell+1)}{2r^2}\right)\varphi(r) &= \frac{m_e E}{\hbar^2}\varphi(r),\end{aligned}$$

where

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 52.918 \text{ pm}$$

is called the *Bohr radius* - it is radius of the ground state orbit in the Bohr model of the hydrogen atom.

The radial TISE is now expressed in terms of scaled distance,

$$u = \frac{r}{a_0},$$

and scaled energy,

$$\epsilon = \frac{E}{E_{\mathrm{h}}},$$

where

$$E_{\mathrm{h}} = \frac{\hbar^2}{m_e a_0^2} = \frac{m_e e^4}{16\pi\epsilon_0\hbar^2} = 27.211386 \text{ eV} = 4.359744 \times 10^{-18} \text{ J}$$

is the called the *Hartree* - it is twice the ionization energy of a hydrogen atom (see below). Specifcally, since

$$\frac{d^2}{dr^2} = \left(\frac{du}{dr}\right)^2 \frac{d^2}{du^2} = \frac{1}{a_0^2}\frac{d^2}{du^2},$$

we get

$$\begin{aligned}\left(-\frac{1}{2}\frac{d^2}{du^2} - \frac{Z}{u} + \frac{\ell(\ell+1)}{2u^2}\right)\varphi(u) &= \frac{m_e a_0^2 E}{\hbar^2}\varphi(u) \\ &= \epsilon\varphi(u).\end{aligned}$$

Note that using scaled distance and energy coordinates, is equivalent to using a_0 and E_h as distance and energy units. Consequently, we revert to the original variables, r and E, with the understanding that they are expressed in these units. The Bohr radius, a_0, and the Hartree are the distance and energy units in the *atomic units* system of units.

In atomic units, the TISE takes the form

$$\left(-\frac{1}{2}\frac{d^2}{dr^2} - \frac{Z}{r} + \frac{\ell(\ell+1)}{2r^2}\right)\varphi(r) = E\varphi(r). \tag{7.1}$$

It is instructive to examine the behavior of the solutions to this equation near the boundary radius values - specifically, $r \to 0$ and $r \to \infty$. Both cases begin by re-writing the above equation in the form,

$$\frac{d^2}{dr^2}\varphi(r) = 2\left(-\frac{Z}{r} + \frac{\ell(\ell+1)}{2r^2} - E\right)\varphi(r). \tag{7.2}$$

Small radius: For small r there are two possibilities: $\ell > 0$ and $\ell = 0$.

(a) If $\ell > 0$, the centrifugal potential term is much larger than the other two terms inside the parentheses. Consequently, the solution to the above equation must be close to the solution of

$$\frac{d^2}{dr^2}\varphi(r) = \frac{\ell(\ell+1)}{r^2}\varphi(r).$$

This equation is clearly solved by $Cr^{\ell+1}$. There is another solution, $C'r^{-\ell}$, but it blows up as $r \to 0$. Since $R = \varphi/r$ also blows up in this case, this solution is inadmissable. Consequently,

$$\varphi(r) \cong Cr^{\ell+1} \qquad r \text{ close to } 0$$

and

$$R(r) \cong Cr^{\ell}. \qquad r \text{ close to } 0.$$

For $\ell > 0$, the electron is excluded from the nucleus - the radial wavefunction goes to zero as $r \to 0$. The decay to zero is in accord with a power law, r^ℓ. The exclusion of the electron from the nucleus is stronger with increasing ℓ, consistent with the increasing centrifugal repulsion.

(b) If $\ell = 0$, there is no centrifugal repulsion - angular momentum is zero. In this case, $R(r)$ extends right to $r = 0$. It is a constant, $C > 0$, in the vicinity of the nucleus. This is consistent with the above radial wavefunction, with $\ell = 0$, even though the analysis only applies to $\ell > 0$. Below, we solve the radial TISE for $\ell = 0$ and verify this result.

Large radius: For large r, the Coulomb and centrifugal potentials approach zero. All that is left is the equation of a free particle in one dimension; i.e., the TISE reduces to

$$\frac{d^2}{dr^2}\varphi(r) = -2E\varphi(r).$$

Since the energy of bound states must be negative ($E = 0$ corresponds to the ionization threshold), the solutions to this equation are

$$A\exp\left(\sqrt{-2E}r\right)$$

and

$$B\exp\left(-\sqrt{-2E}r\right).$$

The first of these solutions blows up as $r \to \infty$, and is inadmissable. The remaining solution, exponential decay as $r \to \infty$, applies equally to $R(r)$, at large r. The actual solutions consist of polynomials (of successively higher order) multiplying the above exponential function. The rate of exponential decay is given by $\sqrt{-2E}$. This rate decreases as the bound state energy increases toward zero. Consequently, higher energy orbitals extend to larger r values.

It is possible to construct raising and lowering operators for the hydrogen atom. Since there is a distinct radial TISE for each ℓ value, there is a ladder of states - and raising and lowering operators - for each ℓ value. We will investigate the $\ell = 0$ ladder of states without constructing raising and lowering operators. The $\ell = 0$ ladder has the lowest energy state - the ground state - since it has no angular energy. The lowest energy state has the longest possible wavelength. We know that $\varphi(r) = Cr$, for small r, and that it decays exponentially for large r. The simplest function satisfying these requirements is

$$\varphi(r) = Cr\exp\left(-\frac{r}{a}\right),$$

where a characterizes the relative width of the radial distribution, $|\varphi(r)|^2$. Here, we let the equations give us the relationship between a and E. Substituting the above expression into the radial TISE, for $\ell = 0$, gives

$$\begin{aligned}
\frac{d^2}{dr^2} Cr \exp\left(-\frac{r}{a}\right) &= 2\left(-\frac{Z}{r} - E\right) Cr \exp\left(-\frac{r}{a}\right) \qquad C \text{ cancels} \\
\frac{d}{dr}\left(\exp\left(-\frac{r}{a}\right) - \frac{r}{a}\exp\left(-\frac{r}{a}\right)\right) &= 2\left(-Z - Er\right)\exp\left(-\frac{r}{a}\right) \\
\left(-\frac{1}{a} - \frac{1}{a} + \frac{r}{a^2}\right)\exp\left(-\frac{r}{a}\right) &= 2\left(-Z - Er\right)\exp\left(-\frac{r}{a}\right) \qquad \exp\left(-\frac{r}{a}\right) \text{ cancels} \\
-\frac{2}{a} + \frac{r}{a^2} &= -2Z - 2Er.
\end{aligned}$$

Both the constant and the coefficient of r must be equal - this is an equation of functions. Therefore,

$$a = \frac{1}{Z}$$

and

$$\frac{1}{a^2} = -2E$$

or

$$E = -\frac{1}{2a^2} = -\frac{Z^2}{2}.$$

For the hydrogen atom, $Z = 1$, and the ground state energy is

$$E_{\text{H, g.s.}} = E_1 = -\frac{1}{2} = -13.605693 \text{ eV}.$$

This is the ionization of hydrogen (hydrogen in its ground state). The associated radial wavefunction is

$$R_{1,0}\left(r\right) = C \exp\left(-Zr\right).$$

The coefficient C is determined by normalization. The normalization condition for $R_{1,0}\left(r\right)$ is

$$\int_0^\infty \left|R_{1,0}\left(r\right)\right|^2 r^2 dr = 1,$$

which gives

$$\begin{aligned}
\left|C\right|^2 &= \frac{1}{\int_0^\infty r^2 \exp\left(-2Zr\right)\, dr} \\
&= \frac{1}{\left(2Z\right)^{-3} \int_0^\infty \rho^2 \exp\left(-\rho\right)\, d\rho} \\
&= \frac{8Z^3}{2} = 4Z^3
\end{aligned}$$

and

$$R_{1,0}\left(r\right) = 2Z^{3/2} \exp\left(-Zr\right).$$

The next state in the $\ell = 0$ ladder is found by replacing Cr by $C_1 r + C_2 r^2$ to get the next higher order polynomial prefactor of the exponential decay function.

Equating the distinct functional forms, on the left and right of the equation resulting from the radial TISE, gives the next energy level. On the left side, we have

$$\begin{aligned}
& \frac{d^2}{dr^2}\left(C_1 r + C_2 r^2\right)\exp\left(-\frac{r}{a}\right) \\
= & \frac{d}{dr}\left(\left(C_1 + 2C_2 r\right)\exp\left(-\frac{r}{a}\right) - \frac{\left(C_1 r + C_2 r^2\right)}{a}\exp\left(-\frac{r}{a}\right)\right) \\
= & \left(2C_2 - 2\frac{\left(C_1 + 2C_2 r\right)}{a} + \frac{\left(C_1 r + C_2 r^2\right)}{a^2}\right)\exp\left(-\frac{r}{a}\right) \\
= & \left(2\left(C_2 - \frac{C_1}{a}\right) - \left(\frac{4C_2}{a} - \frac{C_1}{a^2}\right)r + \frac{C_2}{a^2}r^2\right)\exp\left(-\frac{r}{a}\right).
\end{aligned}$$

On the right side,

$$\begin{aligned}
& 2\left(-\frac{Z}{r} - E\right)\left(C_1 r + C_2 r^2\right)\exp\left(-\frac{r}{a}\right) \\
= & -2\left(C_1 Z + \left(C_2 Z + C_1 E\right)r + C_2 E r^2\right)\exp\left(-\frac{r}{a}\right).
\end{aligned}$$

Equating these functions gives three equations:

$$C_2 - \frac{C_1}{a} = -C_1 Z,$$

$$\frac{4C_2}{a} - \frac{C_1}{a^2} = 2\left(C_2 Z + C_1 E\right)$$

and

$$\frac{C_2}{a^2} = -2C_2 E$$

or

$$E = -\frac{1}{2a^2}.$$

The last equation of the three equations gives the energy in terms of a. The second last now takes the form,

$$\begin{aligned}
\frac{4C_2}{a} - \frac{C_1}{a^2} &= 2C_2 Z - \frac{C_1}{a^2} \\
4C_2 a - C_1 &= 2C_2 Z a^2 - C_1 \\
2 &= Za.
\end{aligned}$$

So,

$$a = \frac{2}{Z}$$

and

$$\begin{aligned}
E_2 &= -\frac{1}{2a^2} \\
&= -\frac{Z^2}{2 \times 2^2} = -\frac{Z^2}{8}.
\end{aligned}$$

The first of the three equations gives C_2 in terms of C_1:

$$\begin{aligned} C_2 &= C_1\left(\frac{1}{a} - Z\right) \\ &= -\frac{Z}{2}C_1. \end{aligned}$$

Therefore,

$$R_{2,0}(r) = C_1\left(1 - \frac{Z}{2}r\right)\exp\left(-\frac{Z}{2}r\right).$$

Normalization determines C_1.

$$\begin{aligned} |C_1|^2 &= \frac{1}{\int_0^\infty \left(1 - \frac{Z}{2}r\right)^2 r^2 \exp(-Zr)\,dr} \\ &= \frac{1}{Z^{-3}\int_0^\infty \left(\rho^2 - \rho^3 + \frac{1}{4}\rho^4\right)\exp(-\rho)\,d\rho} \\ &= \frac{Z^3}{2! - 3! + \frac{1}{4}4!} = \frac{Z^3}{2}, \end{aligned}$$

and

$$R_{2,0}(r) = 2^{-1/2}Z^{3/2}\left(1 - \frac{Z}{2}r\right)\exp\left(-\frac{Z}{2}r\right).$$

The integrals needed to get the normalization follow from the standard integral,

$$\int_0^\infty \rho^k \exp(-\rho)\,d\rho = (-1)^k \left.\frac{d^k}{d\lambda^k}I(\lambda)\right|_{\lambda=1},$$

where

$$\begin{aligned} I(\lambda) &= \int_0^\infty \exp(-\lambda\rho)\,d\rho \\ &= \lambda^{-1}\int_0^\infty \exp(-u)\,du \\ &= \lambda^{-1}. \end{aligned}$$

Therefore,

$$\int_0^\infty \rho^k \exp(-\rho)\,d\rho = 1 \times 2 \times \ldots \times k = k!. \tag{7.3}$$

Continuing the above procedure for $n = 3, 4, \ldots$ determines the entire $\ell = 0$ ladder of states, and the associated energy levels. The procedure for the n th state requires equating coefficients of an n th order polynomial on either side of an equation that results from the radial TISE. To get the n th energy level, it is sufficient to consider only the coefficients of r^n and r^{n-1}. Specically, the radial TISE for the n th state can be written

$$\begin{aligned} &\frac{d^2}{dr^2}\left(P_{n-2}(r) + C_{n-1}r^{n-1} + C_n r^n\right)\exp\left(-\frac{r}{a}\right) \\ &= 2\left(-\frac{Z}{r} - E\right)\left(P_{n-2}(r) + C_{n-1}r^{n-1} + C_n r^n\right)\exp\left(-\frac{r}{a}\right), \end{aligned}$$

where $P_{n-2}(r)$ is a polynomial of order $n-2$. After the derivatives are evaluated on the left of this equation, the coefficient of r^n on the left is C_n/a^2. The corresponding coefficient on the right is $-2EC_n$. Therefore,

$$-2E = \frac{1}{a^2},$$

and

$$E = -\frac{1}{2a^2}.$$

This is true for all the states. a is determined from equating the coefficients of r^{n-1}. This gives

$$\begin{aligned} \frac{C_{n-1}}{a^2} - 2\frac{nC_n}{a} &= -2EC_{n-1} - 2ZC_n \\ \frac{C_{n-1}}{a^2} - 2\frac{nC_n}{a} &= \frac{C_{n-1}}{a^2} - 2ZC_n \\ \frac{n}{a} &= Z \end{aligned}$$

or

$$a = \frac{n}{Z},$$

and

$$E_n = -\frac{Z^2}{2n^2} = -E_{\mathrm{r}}\frac{Z^2}{n^2},$$

where $E_{\mathrm{r}} = E_{\mathrm{h}}/2$ is the Rydberg energy. These are all the energy levels of the hydrogen atom (or hydrogen-like ions). Because of the spherical symmetry of the hydrogen atom, the energy level spectrum is highly degenerate. The ladder of states for each successive ℓ value has energy levels that completely overlap with the $\ell = 0$ energy levels - except that the ℓ ladder begins with $n = \ell + 1$. Thus, the energy levels of the $\ell = 0$ ladder begin with $n = 1$, while the energy levels of the $\ell = 1$ ladder begin with $n = 2$, and so on. Alternatively, we say that for $n = 1$, there is only an state (s means $\ell = 0$) - this is the 1s orbital. For $n = 2$, there are s and p states (p means $\ell = 1$). There is only one 2s state. But, there are three 2p states ($m_\ell = 1, 0$ and -1). For $n = 3$ there are s, p and d states (d means $\ell = 2$). The various ladders are well known to chemistry students. The energy levels of hydrogen, with associated orbitals indicated by short horizontal lines.

To see how every ℓ ladder of states shares the same energy levels, we consider the radial TISE for $\ell > 0$ ladder of states. The lowest energy eigenstate must have the form,

$$\varphi(r) = Cr^{\ell+1}\exp\left(-\frac{r}{a}\right).$$

The k th state in the ladder, $k = 1, 2, \ldots$, has the form,

$$\begin{aligned} \varphi_k(r) &= \left(C_1 r^{\ell+1} + \ldots + C_{k-1}r^{\ell+k-1} + C_k r^{\ell+k}\right)\exp\left(-\frac{r}{a}\right) \\ &= \left(P_{k-3}(r)\, r^{\ell+1} + C_{k-1}r^{\ell+k-1} + C_k r^{\ell+k}\right)\exp\left(-\frac{r}{a}\right), \end{aligned}$$

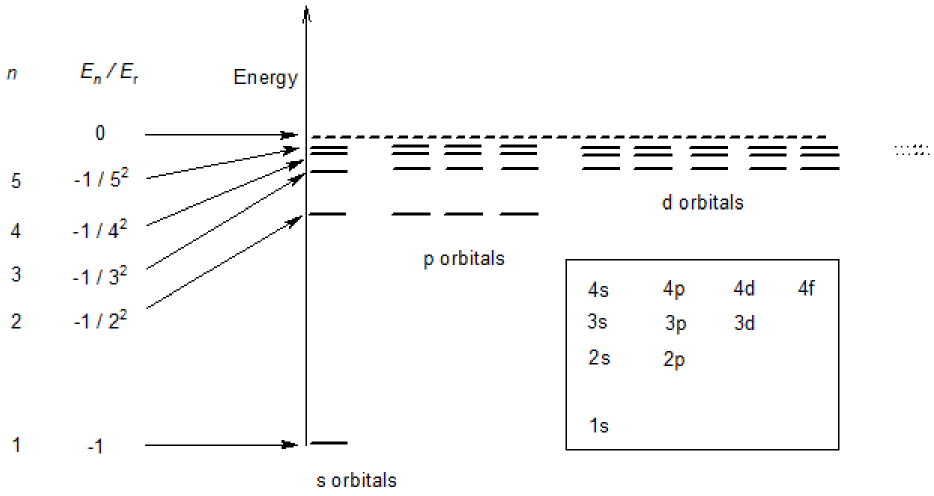


Figure 7.1: The energy levels of hydrogen, with associated orbitals indicated. The inset shows the subshell structure of each shell, up to $n = 4$.

where $P_{k-3}(r)$ is a polynomial of order $k-3$. Substituing the above form into Eq. 7.2 gives

$$\begin{aligned} &\frac{d^2}{dr^2}\left(P_{k-3}(r)\, r^{\ell+1} + C_{k-1} r^{\ell+k-1} + C_k r^{\ell+k}\right)\exp\left(-\frac{r}{a}\right) \\ = \ & 2\left(-\frac{Z}{r} + \frac{\ell(\ell+1)}{2r^2} - E\right)\left(P_{k-3}(r)\, r^{\ell+1} + C_{k-1} r^{\ell+k-1} + C_k r^{\ell+k}\right)\exp\left(-\frac{r}{a}\right). \end{aligned}$$

After the derivatives are evaluated, equating the coefficients of $r^{\ell+k}$ gives

$$\frac{C_k}{a^2} = -2EC_k$$

which gives the same relationship,

$$E = -\frac{1}{2a^2},$$

seen for $\ell = 0$. Equating the coefficients of $r^{\ell+k-1}$ gives

$$\begin{aligned} \frac{C_{k-1}}{a^2} - 2(\ell+k)\frac{C_k}{a} &= -2ZC_k - 2EC_{k-1} \\ \frac{C_{k-1}}{a^2} - 2(\ell+k)\frac{C_k}{a} &= -2ZC_k + \frac{C_{k-1}}{a^2} \end{aligned}$$

which leads to

$$\begin{aligned} a &= \frac{\ell+k}{Z}, \qquad k = 1, 2, \ldots \\ &= \frac{n}{Z} \qquad\quad n = \ell+1, \ell+2, \ldots \end{aligned}$$

If we label the states by $n = \ell + k$, instead of k, the $\ell > 0$ energy levels are the same as those of $\ell = 0$, except that they start at $n = \ell + 1$. This means the largest ℓ consistent with *principle quantum number*, n, is $n - 1$. The centrifugal potential term does not contribute to the determination of E and a above because it lowers the power of r by 2. It contributes only to the coefficients of the polynomial - determined by equating the lower order coefficients.

Energy is labeled by a single quantum number, n, which varies from 1 to ∞. For each n - i.e., for each energy level - the angular momentum quantum number varies from $\ell = 0$ to $\ell = n-1$. For each n and ℓ values, m_ℓ varies from $m_\ell = -\ell$ to $m_\ell = \ell$. The wavefunctions that represent these states have the form,

$$\begin{aligned} R_{n,\ell}(r)\,Y_{\ell,m_\ell}(\theta,\phi) &= P_{n,\ell}\left(\frac{2Zr}{n}\right)\left(\frac{2Zr}{n}\right)^{\ell}\exp\left(-\frac{Zr}{n}\right)Y_{\ell,m_\ell}(\theta,\phi) \\ m_\ell &= -\ell,\ldots,\ell \\ \ell &= 0,\ldots,n-1. \end{aligned}$$

Here, $P_{n,\ell}(\rho_n)$ is a $n-\ell$ th order polynomial. $\left\{P_{n,\ell}(\rho_n)\,\rho_n^{\ell}\exp\left(-\frac{\rho_n}{2}\right)\right\}_{n=\ell+1,\ell+2,\ldots}$ is the ladder of radial eigenfunctions for the angular momentum, ℓ, radial TISE. $\rho_n = 2Zr/n$.

The angular functions are the spherical harmonics - the same functions that represent the rotations of diatomic molecules. The radial functions are indexed by $n-\ell$ since the first radial state of the hydrogen atom for angular momentum quantum number, ℓ, has the same energy as the $n = \ell + 1$ th radial state for zero angular momentum arising from the radial TISE It is instructive to derive the familiar shapes of p and d orbitals from the spherical harmonics with $\ell = 1$ and 2.

The p orbitals, $Y_{1,m_\ell}(\theta,\phi)$, are given as follows:

m_ℓ	$Y_{1,m_\ell}(\theta,\phi)$	$rY_{1,m_\ell}(\theta,\phi)$
1	$\mathrm{p}_1 = (2\pi I_1)^{-1/2}\sin\theta\exp(i\phi)$	$(2\pi I_1)^{-1/2}(x+iy)$
0	$\mathrm{p}_0 = \ \mathrm{p}_z = (4\pi I_1)^{-1/2}\,2\cos\theta$	$(4\pi I_1)^{-1/2}\,2z$
-1	$\mathrm{p}_{-1} = -(2\pi I_1)^{-1/2}\sin\theta\exp(-i\phi)$	$-(2\pi I_1)^{-1/2}(x-iy)$

Real p orbitals are obtained from plus and minus combinations of the p_1 and p_{-1} orbitals:

$$\mathrm{p}_x = \frac{1}{\sqrt{2}}\left(\mathrm{p}_1 - \mathrm{p}_{-1}\right) = (4\pi I_1)^{-1/2}\,2x$$

and

$$\mathrm{p}_y = \frac{1}{i\sqrt{2}}\left(\mathrm{p}_1 + \mathrm{p}_{-1}\right) = (4\pi I_1)^{-1/2}\,2y.$$

The $1/\sqrt{2}$ factors are to ensure normalized orbitals. Dividing the sum of any two orthogonal, normalized orbitals by $\sqrt{2}$ gives a normalized orbital.

The d orbitals, $Y_{2,m_\ell}(\theta,\phi)$, are given as follows:

m_ℓ	$Y_{2,m_\ell}(\theta,\phi)$	$r^2 Y_{2,m_\ell}(\theta,\phi)$
2	$\mathrm{d}_2 = (2\pi I_2)^{-1/2}\sin^2\theta\exp(2i\phi)$	$(2\pi I_2)^{-1/2}(x+iy)^2$ $= (2\pi I_2)^{-1/2}\left(x^2-y^2+2ixy\right)$
1	$\mathrm{d}_1 = -(2\pi I_2)^{-1/2}\, 2\sin\theta\cos\theta\exp(i\phi)$	$-(2\pi I_2)^{-1/2}\, 2(x+iy)z$
0	$\mathrm{d}_0 = \mathrm{d}_{z^2} = (3\pi I_2)^{-1/2}\left(3\cos^2\theta-1\right)$	$(3\pi I_2)^{-1/2}\left(3z^2-r^2\right)$
−1	$\mathrm{d}_{-1} = (2\pi I_2)^{-1/2}\, 2\sin\theta\cos\theta\exp(-i\phi)$	$(2\pi I_2)^{-1/2}\, 2(x-iy)z$
−2	$\mathrm{d}_{-2} = (2\pi I_2)^{-1/2}\sin\theta\exp(-2i\phi)$	$(2\pi I_2)^{-1/2}(x-iy)^2$ $= (2\pi I_2)^{-1/2}\left(x^2-y^2-2ixy\right)$

Real d orbitals are obtained from plus and minus combinations of the d_2 and d_{-2} and, separately, the d_1 and d_{-1} orbitals:

$$\mathrm{d}_{x^2-y^2} = \frac{1}{\sqrt{2}}\left(\mathrm{d}_2+\mathrm{d}_{-2}\right) = (4\pi I_1)^{-1/2}\, 2\left(x^2-y^2\right),$$

$$\mathrm{d}_{xy} = \frac{1}{i\sqrt{2}}\left(\mathrm{d}_2-\mathrm{d}_{-2}\right) = (4\pi I_1)^{-1/2}\, 4xy,$$

$$\mathrm{d}_{xz} = \frac{1}{\sqrt{2}}\left(-\mathrm{d}_1+\mathrm{d}_{-1}\right) = (4\pi I_1)^{-1/2}\, 4xz$$

and

$$\mathrm{d}_{yz} = \frac{1}{i\sqrt{2}}\left(-\mathrm{d}_1-\mathrm{d}_{-1}\right) = (4\pi I_1)^{-1/2}\, 4yz.$$

The first few radial wavefunctions of hydrogen-like ions - in atomic units and expressed in terms of $\rho = 2Zr/n$ - are tabulated as follows:

n	$Z^{-3/2}R_{n,0}(r)$	$Z^{-3/2}R_{n,1}(r)$	$Z^{-3/2}R_{n,2}(r)$
3	$243^{-1/2}\left(6-6\rho+\rho^2\right)\exp\left(-\frac{\rho}{2}\right)$	$486^{-1/2}(4-\rho)\rho\exp\left(-\frac{\rho}{2}\right)$	$2430^{-1/2}\rho^2\exp\left(-\frac{\rho}{2}\right)$
2	$8^{-1/2}(2-\rho)\exp\left(-\frac{\rho}{2}\right)$	$24^{-1/2}\rho\exp\left(-\frac{\rho}{2}\right)$	
1	$2\exp\left(-\frac{\rho}{2}\right)$		

The polynomial prefactors of $\exp\left(-\frac{\rho}{2}\right)$ for $\ell = 0$ are the Laguerre polynomials. For $\ell > 0$, they are associated Laguerre polynomials.

Example 7.1. Expectation values of r for hydrogen atom energy eigenstates.

(a) Determine the expectation value of r for a hydrogen atom in its ground state. How does this compare with the most probable radius?

(b) Determine the expectation value of r for a hydrogen atom in the 2s and 2p states.

(c) What is the probability that a 2s electron is found to be within the first radial lobe? Use wolfram alpha to get the required integral.

Solution 7.1.

(a) The expectation of r, for the ground state, 1s, is

$$\begin{aligned}
\langle \psi_{1\mathrm{s}} | r\psi_{1\mathrm{s}} \rangle &= \langle R_{1,0} | rR_{1,0} \rangle \underbrace{\langle Y_{0,0} | Y_{0,0} \rangle}_{=1} \\
&= \int_0^\infty rR_{1,0}^2(r)\, r^2 dr \\
&= 2^{-4} \int_0^\infty \left(2\exp\left(-\frac{\rho}{2}\right)\right)^2 \rho^3 d\rho, \qquad \rho = 2r \\
&= 2^{-2} \int_0^\infty \exp(-\rho)\, \rho^3 d\rho \\
&= 2^{-2} \times 3! = \frac{3}{2}.
\end{aligned}$$

The mean radial position of the electron in the ground state is 1.5 Bohr = 79.377 pm. The most probable radius is given by finding the maximum of

$$R_{1,0}^2(r)\, r^2 = \exp(-\rho)\, \rho^2.$$

Set

$$\begin{aligned}
\frac{d}{d\rho}\rho^2 \exp(-\rho) &= 0 \\
\left(2\rho - \rho^2\right) \exp(-\rho) &= 0\,,
\end{aligned}$$

which gives

$$\rho = 2$$

and

$$r = 1.$$

The Bohr radius is the most probable radius of the ground state of the hydrogen atom. It is smaller than the mean radius, the expectation value, because the distribution extends to infinity - it is asymmetric about its peak at $r = 1$.

(b) For the 2s state,

$$\begin{aligned}
\langle \psi_{2\mathrm{s}} | r\psi_{2\mathrm{s}} \rangle &= \langle R_{2,0} | rR_{2,0} \rangle \\
&= \int_0^\infty R_{2,0}^2(r)\, r^3 dr \\
&= \int_0^\infty \left(8^{-1/2} (2-\rho) \exp\left(-\frac{\rho}{2}\right)\right)^2 \rho^3 d\rho, \qquad \rho = r \\
&= 8^{-1} \int_0^\infty \exp(-\rho)\, \left(4\rho^3 - 4\rho^4 + \rho^5\right) d\rho \\
&= 8^{-1} (4 \times 3! - 4 \times 4! + 5!) \\
&= 8^{-1} (24 - 96 + 120) \\
&= \frac{48}{8} = 6.
\end{aligned}$$

For the 2p state,

$$\begin{aligned}
\left\langle \psi_{2\mathrm{p}} \left| r\psi_{2\mathrm{p}} \right.\right\rangle &= \left\langle R_{2,1} \left| rR_{2,1} \right.\right\rangle \\
&= \int_0^\infty R_{2,1}^2(r)\, r^3 dr \\
&= \int_0^\infty \left(24^{-1/2}\rho \exp\left(-\frac{\rho}{2}\right)\right)^2 \rho^3 d\rho, \qquad \rho = r \\
&= 24^{-1} \int_0^\infty \exp(-\rho)\, \rho^5\, d\rho \\
&= 24^{-1} \times 5! \\
&= 5.
\end{aligned}$$

(c) The first lobe of the 2s state is in the interval, zero to the node in $R_{2\mathrm{s}}(r)$. The node occurs where $r = 2$. The probability that r is less than 2 is the probability that the electron is in the first lobe. This probability is given by

$$\begin{aligned}
&\int_0^2 R_{2,0}^2(r)\, r^2 dr \\
&= \int_0^2 \left(8^{-1/2}(2-\rho)\exp\left(-\frac{\rho}{2}\right)\right)^2 \rho^2 d\rho \\
&= 8^{-1}\int_0^2 \left(4\rho^2 - 4\rho^3 + \rho^4\right)\exp(-\rho)\, d\rho \\
&= 0.0527. \qquad \text{using wolfram alpha}
\end{aligned}$$

The electron is about 95% likely to be in the outer lobe.

Example 7.2. Potential and kinetic energy expectations for hydrogen-like ion energy eigenstates.

(a) What is the expectation value of the Coulomb potential energy for the 1s state of He^+? What is the relationship between this expectation value and the total energy of the 1s state?

(b) What is the expectation value of the kinetic energy for the 1s state of He^+? You can use the function $\varphi_{1,0}(r)$, for which the radial inner product has no r^2 factor in the integrand, and the radial kinetic energy is simply $-\frac{1}{2}d^2/dr^2$. The 1s state has only radial kinetic energy. You can also use the result of part a, and the known total energy. What is the relationship between the kinetic energy expectation and the total energy?

Solution 7.2.

(a) The expectation of the Coulomb potential energy for He^+ is

$$
\begin{aligned}
\left\langle \psi_{1\mathrm{s}} \left| \left(-\frac{2}{r}\right) \psi_{1\mathrm{s}} \right.\right\rangle &= \int_0^\infty \left(-\frac{2}{r}\right) R_{1,0}^2\left(r\right)\, r^2 dr \\
&= -8^{-1} \int_0^\infty \left(2^{3/2} 2 \exp\left(-\frac{\rho}{2}\right)\right)^2 \rho\, d\rho, \qquad \rho = 4r \\
&= -4 \int_0^\infty \exp\left(-\rho\right)\, \rho\, d\rho \\
&= -4 \\
&= 2 \times \left(-\frac{2^2}{2}\right) = 2E_1.
\end{aligned}
$$

The expectation value of potential energy is twice the total energy, for the 1s state. This is true for all states. It is a consequence of the virial theorem.

(b) The expectation of the radial kinetic energy, $\hat{K}_r$, for He^+ is

$$
\begin{aligned}
\left\langle \psi_{1\mathrm{s}} \left| \hat{K}_r \psi_{1\mathrm{s}} \right.\right\rangle &= \left\langle \varphi_{1,0} \left| \left(-\frac{1}{2}\frac{d^2}{dr^2}\right) \varphi_{1,0} \right.\right\rangle \underset{=1}{\left\langle Y_{1,0} \left| Y_{1,0} \right.\right\rangle} \\
&= \int_0^\infty 2^{3/2} 2r \exp\left(-2r\right) \left(-\frac{1}{2}\frac{d^2}{dr^2}\right) 2^{3/2} 2r \exp\left(-2r\right)\, dr \\
&= \frac{1}{2}\int_0^\infty \rho \exp\left(-\frac{\rho}{2}\right) \left(-8\frac{d^2}{d\rho^2}\right) \rho \exp\left(-\frac{\rho}{2}\right) d\rho \qquad \rho = 4r \\
&= -4 \int_0^\infty \rho \exp\left(-\frac{\rho}{2}\right) \frac{d}{d\rho} \left(1 - \frac{\rho}{2}\right) \exp\left(-\frac{\rho}{2}\right)\, d\rho \\
&= -4 \int_0^\infty \rho \exp\left(-\frac{\rho}{2}\right) \left(-\frac{1}{2} - \frac{1}{2}\left(1 - \frac{\rho}{2}\right)\right) \exp\left(-\frac{\rho}{2}\right)\, d\rho \\
&= -4 \int_0^\infty \left(-\rho + \frac{\rho^2}{4}\right) \exp\left(-\rho\right)\, d\rho \\
&= -4 \left(-1 + \frac{2}{4}\right) = 2 = -E_1.
\end{aligned}
$$

The expectation of kinetic energy is just minus the total energy. This also follows from the virial theorm. The expectation of kinetic energy could also be determined from

$$
\begin{aligned}
\left\langle \psi_{1\mathrm{s}} \left| \hat{K} \psi_{1\mathrm{s}} \right.\right\rangle &= \left\langle \psi_{1\mathrm{s}} \left| \left(\hat{H} - \hat{V}\right) \psi_{1\mathrm{s}} \right.\right\rangle \\
&= \left\langle \psi_{1\mathrm{s}} \left| \hat{H} \psi_{1\mathrm{s}} \right.\right\rangle - \left\langle \psi_{1\mathrm{s}} \left| \hat{V} \psi_{1\mathrm{s}} \right.\right\rangle \\
&= E_1 - (-4) \qquad \text{from part a} \\
&= -2 + 4 = 2.
\end{aligned}
$$

7.2 Transitions of the hydrogen atom.

Hydrogen atom transitions are subject to the same transition rule as rotations of a diatomic molecule - namely, $\Delta\ell = \pm 1$ and $\Delta m_\ell = 0$ or ± 1. This is because the angular dependence of the dipole operator has the same form in both cases - it is $\cos\theta$ ($\sin\theta\cos\phi$ for $\Delta m_\ell = \pm 1$ transitions). The radial transitions of the hydrogen atom are not subject to the selection rule of the diatomic vibrational transitions. The latter transitions are harmonic oscillator transitions. The transition integrals of radial transitions of the hydrogen atom,

$$\int_0^\infty R_{n',\ell\pm 1}(r)\, r R_{n,\ell}(r)\, r^2 dr,$$

are non-zero for all n' and n.

Because the energy level spacings in hydrogen decrease dramatically with increasing n, the transitions of hydrogen are labeled according to the lower level involved. For example, transitions to or from $n = 1$ levels are called Lyman transitions - the observed lines constitute the Lyman series. Transitions to or from $n = 2$ (except involving $n = 1$) are Balmer transitions, and so on (Paschen, Brackett, ...). The Lyman series lines are all in the UV portion of the EM spectrum. The Balmer lines are visible transitions. The other series are in the infrared, and beyond. After the Balmer series, the other series begin to overlap - increasingly so with larger smallest n.

7.3 Electron spin

Otto Stern and Walther Gerlach demonstrated the splitting of a beam of silver atoms into two beams by an inhomogeneous magnetic field, in 1922. In 1927, T.E. Phipps and J.B. Taylor split a beam of 1s hydrogen atoms into two beams with the same inhomogeneous magnetic field. These experiments reveal the existence of spin - since these atoms have no net orbital angular momentum. Excited states of hydrogen reveal the distinct m_ℓ states through splitting into additional beams. In the absence of a magnetic field, all the distinct m_ℓ states of an n, ℓ subshell have the same energy, E_n. When a magnetic field is applied, states with different m_ℓ have different energy - the *Zeeman effect*. It arises because of the interaction of the magnetic field and the magnetic dipole of the electron "orbiting" the nucleus. An orbiting electron has a *magnetic dipole* associated with its orbital angular momentum,

$$\hat{\boldsymbol{\mu}}_{m,\mathrm{orb}} = -\frac{e}{2m_e}\hat{\mathbf{L}}.$$

The interaction energy - the orbital *Zeeman energy* - is given by

$$\hat{H}^{(1)}_{\mathrm{orb,\ Z}} = -\hat{\boldsymbol{\mu}}_{m,\mathrm{orb}}\cdot\mathbf{B},$$

where **B** is the *magnetic field* - in Tesla (SI units) Choose the magnetic field to be along the z axis. In this case,

$$\hat{H}^{(1)}_{\text{orb, Z}} = \frac{e}{2m_e} B \hat{L}_z.$$

The eigenstates of this operator are the eigenstates of $\hat{L}_z$ - i.e., the distinct m_ℓ states of hydrogen. The Zeeman energy shifts are the eigenvalues of $\hat{H}^{(1)}_{\text{orb, Z}}$, given by

$$\begin{aligned} E^{(1)}_{\text{orb, Z, } m_\ell} &= \frac{e\hbar}{2m_e} B\, m_\ell \\ &= \mu_{\text{B}} B\, m_\ell, \end{aligned}$$

where

$$\mu_{\text{B}} = \frac{e\hbar}{2m_e} = 5.7883818066(38) \times 10^{-5} \text{ eV T}^{-1}$$

is called the *Bohr magneton* - it is the *atomic unit* of *magnetic dipole moment.* Subshells split into a set of distinct energy levels (with plus or minus energy shift in accord with the sign of m_ℓ) that are detected by the associated splitting of spectral lines when hydrogen is in a magnetic field. Spectral lines appear subject to the selection rules, $\Delta m_\ell = 0$ or ± 1. The associated frequencies are proportional to the applied magnetic field. This is the origin of the term *magnetic quantum number*, and the m_ℓ notation. However, in addition to the orbital Zeeman splitting experiments show splitting due to *electron spin* and *nuclear spin.*

An electron beam in an inhomogeneous magnetic field (plus a transverse electric field to counter the Lorentz force that causes the electrons to spiral) splits into two beams, This shows that electrons have two spin states, which corresponds to a *spin angular momentum quantum number*, $s = 1/2$. s is analogous to the orbital quantum number, ℓ . One key difference (other than the half-integral nature of s) is that there is only this one s state. The Zeeman interaction energy, accounting for both orbital and spin contributions, is

$$\begin{aligned} \hat{H}_{\text{Z}} &= -\left(\hat{\mu}_{\text{orb, } z} + \hat{\mu}_{\text{spin, } z}\right) B \\ &= \left(\frac{\mu_{\text{B}}}{\hbar}\hat{L}_z + \frac{g_e \mu_{\text{B}}}{\hbar}\hat{S}_z\right) B, \end{aligned}$$

where $g_e = 2.002319304361(53)$ is called the electron g-factor. Its value is known to extraordinary precision with experiment and theory in lock step. Classical relativity theory predicts a value of 2. Quantum electrodynamics predicts the correction as shown - in agreement with experiment. The eigenvalues of the Zeeman Hamiltonian are

$$E_{\text{Z}, m_\ell, m_s} = \mu_{\text{B}} B \left(m_\ell + g_e m_s\right),$$

where m_s ($= 1/2$ or $-1/2$) labels the two eigenstates of $\hat{S}_z$ (called α and β, respectively) in the same way m_ℓ labels the eigenstates of $\hat{L}_z$. Nuclear spin contributes an additional level of splitting sensitive to the M_I quantum number associated with the z component of nuclear spin angular momentum, $\hat{I}_z$. However, this splitting is orders of magnitude smaller. Nevertheless, the nuclear spin Zeeman splitting is the basis for *nuclear magnetic resonance* (NMR) spectroscopy and *magnetic resonance imaging* (MRI). The electron orbital and spin Zeeman splittings are the basis for *electron spin resonance* (ESR) spectroscopy (a.k.a. electron paramagnetic resonance - EPR - spectroscopy).

Example 7.3. Determine all Zeeman energies for an $\ell = 1$ electron.

Solution 7.3. For $\ell = 1$, $m_\ell = -1, 0, 1$. Also, $m_s = \pm\frac{1}{2}$. The energy levels are tabulated as follows:

$$\begin{array}{cccc} m_\ell & m_s & E_{\mathrm{Z},m_\ell,m_s} & E_{\mathrm{Z},m_\ell,m_s} \quad g_e = 2 \\ 1 & \frac{1}{2} & \mu_\mathrm{B} B\left(1+\frac{g_e}{2}\right) & 2\mu_\mathrm{B} B \\ 0 & \frac{1}{2} & \mu_\mathrm{B} B\frac{g_e}{2} & \mu_\mathrm{B} B \\ -1 & \frac{1}{2} & \mu_\mathrm{B} B\left(-1+\frac{g_e}{2}\right) & 0 \\ 1 & -\frac{1}{2} & \mu_\mathrm{B} B\left(1-\frac{g_e}{2}\right) & 0 \\ 0 & -\frac{1}{2} & -\mu_\mathrm{B} B\frac{g_e}{2} & -\mu_\mathrm{B} B \\ -1 & -\frac{1}{2} & \mu_\mathrm{B} B\left(-1-\frac{g_e}{2}\right) & -2\mu_\mathrm{B} B \end{array}.$$

Since $g_e/2$ is so close to 1, the energy levels are very close to the orbital Zeeman (i.e., with spin left out) energy levels of an $\ell = 2$ electron - except that the zero energy state is doubly degenerate.

7.4 The helium atom and the Pauli principle

The helium atom has two electrons. Both electrons interact with the nucleus in same fashion as the one electron in He$^+$. If the electrons did not interact, the energy eigenstates of the two electron helium atom would be the product of two one electron wavefunctions - one for each electron. The one electron states would be the energy eigenstates of He$^+$. There is one twist to this story, however. Electrons are identical particles - one electron is exactly like another. More than that, they are indistinguishable. The significance of this property is best understood by considering a product wavefunction for two electrons.

For example,

$$(1\mathrm{s}_\uparrow)(1)(1\mathrm{s}_\downarrow)(2)$$

is the ground state of a fictitious helium atom with no electron-electron repulsion and distinguishable electrons. Here, $(1\mathrm{s}_\uparrow)(1)$ is the wavefunction for electron 1 in the 1s orbital, with spin up ($m_s = +1/2$). The argument, 1, represents all the coordinates (spatial and spin) of electron 1. Electron 2 is in the 1s orbital with spin down. $(1\mathrm{s}_\uparrow)$ and $(1\mathrm{s}_\downarrow)$ are called *spin-orbitals* - they specify the complete state (orbital and spin) of one electron. Since the electrons are indistinguishable, they cannot be labeled. However, states of particles that cannot be labeled must be written in terms of states with labeled electrons - like the one above. The state of two indistinguishable particles must be such that if we measure an observable property of particle 1 it must yield the same result as the measurement of the same property of particle 2. This means that if we exchange the two particle labels, in the state of the system, the wavefunction can change - at most - by a phase factor, $e^{i\sigma}$. Wavefunctions differing only by a phase factor - i.e.,

$$\psi = e^{i\sigma}\varphi$$

- produce the same measurement outcomes.

Let $\hat{P}_{1,2}$ be the particle 1 and 2 *exchange operator*. Its effect on the above two particle product state is given by

$$\hat{P}_{1,2}(1\mathrm{s}_\uparrow)\,(1)\,(1\mathrm{s}_\downarrow)\,(2) = (1\mathrm{s}_\uparrow)\,(2)\,(1\mathrm{s}_\downarrow)\,(1)\,.$$

The product state is not an eigenfunction of the exchange operator. In the resultant state electron 1 is spin down and electron 2 is spin up. Admissable wavefunctions of a two indistinguishable particle system are eigenfunctions of $\hat{P}_{1,2}$, and the eigenvalues must be phase factors - i.e., they must have modulus 1. Since applying $\hat{P}_{1,2}$ twice returns the wavefunction to its initial form, we must have $e^{2i\sigma} = 1$ for the admissable phase factors. Thus, $e^{i\sigma} = 1$ or -1, and there are two kinds of indistinguishable particles. Particles whose many particle wavefunctions are eigenfunctions of the exchange operator (for any pair of particles) associated with eigenvalue 1 are called *bosons*, while those with eigenvalue -1 are called *fermions.* Electrons, and all the other elementary particles of matter, are fermions. They are *antisymmetric* with respect to exchange of two particles - the exchange eigenvalue is -1 for electrons. This is called the *Pauli principle.*

Antisymmetric wavefunctions can be constructed from product states like the one above. We define the antisymmetrizer operator, $\hat{A}$, for an n electron system as the sum over all permutation operators, $\hat{P}$, weighted by -1 if they have odd order, $p\left(\hat{P}\right)$. $p\left(\hat{P}\right)$ is the number of exchanges required to represent the permutation as a product of exchanges.

$$\hat{A} = \frac{1}{\sqrt{n!}} \sum_{\hat{P}} (-1)^{p\left(\hat{P}\right)}\,\hat{P}.$$

The $1/\sqrt{n!}$ is needed to give a normalized result when $\hat{A}$ is applied to a product of orthogonal one electron states occupied by different electrons. For a two electron system, there are only two permutations: the identity, represented here by the multiplication operator, 1, and the 1,2 exchange operator, $\hat{P}_{1,2}$. In this case,

$$\hat{A} = \frac{1}{\sqrt{2}}\left(1 - \hat{P}_{1,2}\right).$$

Applying the antisymmetrizer to the above product state gives a properly normalized, antisymmetric state of a helium atom;

$$\begin{aligned}\hat{A}(1\mathrm{s}_{\uparrow})\,(1)\,(1\mathrm{s}_{\downarrow})\,(2) &= \frac{1}{\sqrt{2}}\left(1-\hat{P}_{1,2}\right)(1\mathrm{s}_{\uparrow})\,(1)\,(1\mathrm{s}_{\downarrow})\,(2)\\ &= \frac{1}{\sqrt{2}}\left((1\mathrm{s}_{\uparrow})\,(1)\,(1\mathrm{s}_{\downarrow})\,(2)-(1\mathrm{s}_{\uparrow})\,(2)\,(1\mathrm{s}_{\downarrow})\,(1)\right).\end{aligned}$$

This state is an approximation to the ground state of helium - one that neglects the electron-electron repulsion. Note that either the two orbitals or the two spin states had to be different. Otherwise, the antisymmetrizer would have given zero when applied to the product state. Because electrons are fermions, no two electrons can occupy the same one electron state (spin-orbital). This gives rise to the *Aufbau principle*, and the shell structure of the ground state electron configurations of all the elements - i.e., the *periodic table*.

The above spin-orbitals are products of orbital and spin states;

$$(1\mathrm{s}_{\uparrow})\,(1)=(1\mathrm{s})\,(1)\,\alpha\,(1)$$

and

$$(1\mathrm{s}_{\downarrow})\,(1)=(1\mathrm{s})\,(1)\,\beta\,(1)\,.$$

Consequently, the above approximate ground state of helium takes the form,

$$\begin{aligned}\psi^{(0)}_{\mathrm{He\ g.s.}} &= \frac{1}{\sqrt{2}}\left((1\mathrm{s})\,(1)\,(1\mathrm{s})\,(2)\,\alpha\,(1)\,\beta\,(2)-(1\mathrm{s})\,(2)\,(1\mathrm{s})\,(1)\,\alpha\,(2)\,\beta\,(1)\right)\\ &= (1\mathrm{s})\,(1)\,(1\mathrm{s})\,(2)\,\frac{1}{\sqrt{2}}\left(\alpha\,(1)\,\beta\,(2)-\alpha\,(2)\,\beta\,(1)\right)\\ &= (1\mathrm{s})\,(1)\,(1\mathrm{s})\,(2)\,\chi_{0,0}\,(1,2)\,,\end{aligned}$$

where $\chi_{0,0}\,(1,2)$ is the two spin *singlet state* wherein the electrons have opposite spin and the total spin angular momentum is zero. The total spin quantum numbers are $(S, M_S) = (0,0)$. This state factors into an orbital part - both electrons in the 1s orbital - and a spin part - the two spin singlet state. The orbital part is symmetric with respect to exchange of the electrons, while the spin part is antisymmetric. The Pauli principle insists that the total two electron state be antisymmetric with repsect to exchange. When the two electron state splits into orbital and spin factors, either the two spin state or the two orbital state is antisymmetric, while the other is symmetric - to give an overall antisymmetric state. A many electron state factors into orbital and spin parts whenever *spin-orbit coupling* can be neglected. Spin-orbit coupling is small for light atoms. It is not small for inner shell electrons of heavy atoms. These electrons travel at relativistic speeds for which energy transfers between spin and orbital states.

Since helium (or any atom) has spherical symmetry, the total angular momentum - the sum of total spin and orbital angular momenta of the two electrons - must be conserved. This is because there is no net torque on the atom. Conservation of angular momentum, in quantum mechanics, means that total angular momentum operator, $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$, commutes with the Hamiltonian, and the energy eigenstates can be labeled by the quantum numbers that label the eigenstates of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$; namely, J and M_J. These quantum numbers are called *good quantum numbers*. Note that these are not diatomic rotational quantum numbers. $\hat{\mathbf{J}}$ is used for both total electronic angular momentum and molecular rotational angular momentum.

Since spin-orbit coupling is very small for helium (and all the light elements), total orbital and spin angular momenta of the atom are separately conserved. In this case, L, M_L, S and M_S are also good quantum numbers. This case is called *Russell-Saunders coupling*. Here, the total orbital and spin quantum numbers are good, whereas the individual electron orbital and spin quantum numbers are not. Electron-electron interactions are strong. They couple the orbital states of the two electrons, and (separately) their spin states. The above approximate helium ground state has $(L, M_L) = (0, 0)$.

With the above approximate ground state for helium, we can approximate the ground state energy by forming the expectation value of the helium Hamiltonian, including the electron-electron repulsion. Since the above state is an eigenstate of the Hamiltonian of the two electron Hamiltonian without electron-electron repulsion, this estimate corresponds to a first-order-perturbation-theory corrected (with respect to electron-electron repulsion) of the helium ground state energy. The helium Hamiltonian can be written (in atomic units) as

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}.$$

This Hamiltonian does not affect spin states. The ground state, $\psi^{(0)}_{\mathrm{He\ g.s.}}$, consequently factors into orbital and spin parts. The expectation value of the Hamiltonian for a helium atom in this state correspondingly factors into space and spin inner products,

$$\begin{aligned}
&E^{(0\ \&\ 1)}_{\text{He g.s.}} \\
= &\underbrace{\left\langle \psi^{(0)}_{\text{He g.s.}} \middle| \hat{H}\psi^{(0)}_{\text{He g.s.}} \right\rangle}_{\text{total inner product for 2 e}^{-}\text{s, space and spin}} \\
= &\underbrace{\left\langle (1\text{s})(1)(1\text{s})(2) \middle| \hat{H}(1\text{s})(1)(1\text{s})(2) \right\rangle}_{\substack{\text{this is a spatial inner product for 2 e}^{-}\text{s} \\ \text{a six dimensional integral}}} \underbrace{\left\langle \chi_{0,0}(1,2) \middle| \chi_{0,0}(1,2) \right\rangle}_{\substack{\text{this is a spin inner product for 2 e}^{-}\text{s} \\ =1,\ \chi_{0,0}(1,2)\text{ is normalized}}} \\
= &\left\langle (1\text{s})(1)(1\text{s})(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}} \right) (1\text{s})(1)(1\text{s})(2) \right\rangle \\
= &\left\langle (1\text{s})(1) \middle| \hat{H}_1 (1\text{s})(1) \right\rangle \left\langle (1\text{s})(2) \middle| (1\text{s})(2) \right\rangle + \left\langle (1\text{s})(1) \middle| (1\text{s})(1) \right\rangle \left\langle (1\text{s})(2) \middle| \hat{H}_2 (1\text{s})(2) \right\rangle \\
&+ \left\langle (1\text{s})(1)(1\text{s})(2) \middle| \frac{1}{r_{12}} (1\text{s})(1)(1\text{s})(2) \right\rangle \\
= &\underbrace{-\frac{2^2}{2\times 1^2}}_{1\text{s e}^{-}\text{ in He}^{+}} - \underbrace{\frac{2^2}{2\times 1^2}}_{1\text{s e}^{-}\text{ in He}^{+}} + J_{1\text{s},1\text{s}} \\
= &-4 + J_{1\text{s},1\text{s}},
\end{aligned}$$

where $J_{1\text{s},1\text{s}} > 0$ is called the 1s-1s Coulomb integral - it gives the repulsion energy of the two 1s electrons. Here, we see that the ground state energy of He is somewhat higher than -4 Hartree (the energy without electron-electron repulsion). The Coulomb integral,

$$J_{1\text{s},1\text{s}} = \int \frac{(1\text{s})^2(\mathbf{x}_1)(1\text{s})^2(\mathbf{x}_2)}{r_{12}} d\mathbf{x}_1 d\mathbf{x}_2,$$

is the six dimensional integral over the coordinates of electron 1 and electron 2. The integrand consists of the product of the 1s orbital probability density, $(1\text{s})^2$, for both electron 1 and electron 2, divided by the distance, $r_{12} = \|\mathbf{x}_1 - \mathbf{x}_2\|$, separating the electrons. The integrals can be evaluated analytically. The result is

$$J_{1\text{s},1\text{s}} = \frac{5}{4} \text{ Hartree (remember, we are in atomic units).}$$

The first order perturbation theory estimate of the ground state energy is

$$E^{(0\ \&\ 1)}_{\text{He g.s.}} = -4 + 1.25 = -2.75 \text{ Hartree} = -74.83 \text{ eV}$$

The true ground state energy of helium is

$$E_{\text{He g.s.}} = -2.9036 \text{ Hartree} = -79.01 \text{ eV}.$$

7.4.1 The variational principle

Our first order perturbation theory estimate of the ground state energy is higher than the true value. It is easy to prove that the expectation value of the Hamiltonian is a minimum when the state is the true ground state. Thus, any other state produces higher value - this is called the *variational principle.* We can estimate the ground state energy of a system by minimizing the expectation value of the Hamiltonian with respect to parameters that define the state. For example, the Coulomb integral defined above is quite large, making the ground state energy estimate too high. This is because the electrons partially shield each other from the nucleus - there is an effective nuclear charge, Z_{eff}, which is less than $Z = 2$ for helium. Consequently, if we replace the Z that appears in exponential portion of the radial wavefunction, $\exp(-Zr)$, and in the normalization constant, by variable Z_{eff}, then the above expectation value can be minimized with respect to Z_{eff} to give a better estimate of the ground state energy. We also get a better estimate of the ground state itself, and a "best" effective nuclear charge value for $1s^2$ electrons.

7.4.2 The ionization energy of helium

The ionization energy of helium is the change in energy change of the process,

$$\text{He} \rightarrow \text{He}^+ + \text{e}^-,$$

where the free electron on the right has zero kinetic energy. The ionization of helium is thus

$$\begin{aligned} I_{\text{He}} &= E_{\text{He}^+ \text{ g.s.}} - E_{\text{He g.s.}} \\ &= -2 - (-2.9036) \\ &= 0.9036 \text{ Hartree} = 24.59 \text{ eV}. \end{aligned}$$

Using our first order perturbation theory estimate for the ground state of He, the ionization energy is approximated as

$$\begin{aligned} & E_{\text{g.s. He}^+} - E^{(0\ \&\ 1)}_{\text{g.s. He}} \\ = {} & -2 - (-2.75) \\ = {} & 0.75 \text{ Hartree} = 20.41 \text{ eV}. \end{aligned}$$

7.5 Excited states of helium

The excited states of helium are similarly approximated as antisymmetrized product states. For example, consider the $1s^1 2s^1$ electron configuration. Here, there are two different orbitals, 1s and 2s, as well as two different spin states. Now, the two-orbital state can be either symmetric or antisymmetric (with respect to exchange). In the symmetric case, the spin state will be antisymmetric as in the ground state. However, if the two-orbital state is antisymmetric, the spin state must be symmetric. This corresponds to the three spin states,

$$\chi_{1,1}(1,2) = \alpha(1)\,\alpha(2)\,,$$

$$\chi_{1,0}(1,2) = \frac{1}{\sqrt{2}}\left(\alpha(1)\,\beta(2) + \alpha(2)\,\beta(1)\right)$$

and

$$\chi_{1,-1}(1,2) = \beta(1)\,\beta(2)\,.$$

These two-spin states are eigenstates of the total spin operators, $\hat{S}^2$ and $\hat{S}_z$, where

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2,$$

associated with the quantum numbers, S and M_S, listed in the two-spin state subscripts. All of the two-orbital states mentioned so far,

$$\psi^{(0)}_{\text{He g.s.}} = (1\text{s})\,(1)\,(1\text{s})\,(2)\,, \qquad \text{goes with singlet spin state - } S = 0$$

$$\begin{aligned} \psi^{(0)}_{\text{He 1st sing}} &= \frac{1}{\sqrt{2}}\left((1\text{s})\,(1)\,(2\text{s})\,(2) + (1\text{s})\,(2)\,(2\text{s})\,(2)\right) && \text{goes with singlet spin state - } S = 0 \\ &= \hat{B}(1\text{s})\,(1)\,(2\text{s})\,(2) && \begin{array}{c}\hat{B} \text{ is the symmetrizer operator} \\ = 2^{-1/2}\left(1 + \hat{P}_{1,2}\right)\end{array} \end{aligned}$$

and

$$\begin{aligned} \psi^{(0)}_{\text{He 1st trip}} &= \frac{1}{\sqrt{2}}\left((1\text{s})\,(1)\,(2\text{s})\,(2) - (1\text{s})\,(2)\,(2\text{s})\,(2)\right) && \text{goes with triplet spin state - } S = 1 \\ &= \hat{A}(1\text{s})\,(1)\,(2\text{s})\,(2) \end{aligned}$$

have zero orbital angular momentum, and total orbital angular momentum quantum numbers, $L = 0$ and $M_L = 0$. Of these, the antisymmetric two-orbital state has the lowest energy. This can be seen by evaluating the expectation value of the Hamiltonian for helium atoms in these two states. First consider the symmetric two-orbital state which pairs with the antisymmetric spin state

- the singlet state.

$$\begin{aligned}
& E^{(0\ \&\ 1)}_{\text{He 1st sing}} \\
= & \left\langle \psi^{(0)}_{\text{He 1st sing.}} \middle| \hat{H}\psi^{(0)}_{\text{He 1st sing}} \right\rangle \\
= & \left\langle \hat{B}(1\text{s})\,(1)\,(2\text{s})\,(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right) \hat{B}(1\text{s})\,(1)\,(2\text{s})\,(2) \right\rangle \qquad \text{two-spin state is normalized} \\
= & \left\langle (1\text{s})\,(1)\,(2\text{s})\,(2) \middle| \hat{B}\left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right) \hat{B}(1\text{s})\,(1)\,(2\text{s})\,(2) \right\rangle \qquad \hat{B} \text{ is Hermitian} \\
= & \left\langle (1\text{s})\,(1)\,(2\text{s})\,(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right) \hat{B}^2(1\text{s})\,(1)\,(2\text{s})\,(2) \right\rangle \qquad \hat{B} \text{ commutes with } \hat{H} \\
= & \left\langle (1\text{s})\,(1)\,(2\text{s})\,(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right) \sqrt{2}\hat{B}(1\text{s})\,(1)\,(2\text{s})\,(2) \right\rangle \qquad \hat{B}^2 = \sqrt{2}\hat{B} \\
= & \left\langle (1\text{s})\,(1)\,(2\text{s})\,(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right) ((1\text{s})\,(1)\,(2\text{s})\,(2) + (1\text{s})\,(2)\,(2\text{s})\,(2)) \right\rangle \\
= & \underbrace{-\frac{2^2}{2\times 1^2}}_{1\text{s e}^- \text{ in He}^+} \underbrace{-\frac{2^2}{2\times 2^2}}_{2\text{s e}^- \text{ in He}^+} + J_{1\text{s},2\text{s}} + K_{1\text{s},2\text{s}} \\
= & -\frac{5}{2} + J_{1\text{s},2\text{s}} + K_{1\text{s},2\text{s}},
\end{aligned}$$

where $J_{1\text{s},2\text{s}}$ and $K_{1\text{s},2\text{s}}$ are the 1s-2s *Coulomb* and *exchange* integrals:

$$J_{1\text{s},2\text{s}} = \int \frac{(1\text{s})^2\,(\mathbf{x}_1)\,(2\text{s})^2\,(\mathbf{x}_2)}{r_{12}}\, d\mathbf{x}_1 d\mathbf{x}_2$$

and

$$K_{1\text{s},2\text{s}} = \int \frac{(1\text{s})\,(\mathbf{x}_1)\,(2\text{s})\,(\mathbf{x}_1)\,(1\text{s})\,(\mathbf{x}_2)\,(2\text{s})\,(\mathbf{x}_2)}{r_{12}}\, d\mathbf{x}_1 d\mathbf{x}_2.$$

Whereas the Coulomb integral can be viewed as the electrostatic repulsion of two classical spherical charge distributions, the exchange integral has no classical analogue. Both of these terms are positive. The singlet state is shifted up by both Coulomb and exchange integrals.

When we construct the expectation value of the Hamiltonian with the first triplet state, the result is

$$\left\langle \psi^{(0)}_{\text{He 1st trip.}} \middle| \hat{H}\psi^{(0)}_{\text{He 1st trip}} \right\rangle = -\frac{5}{2} + J_{1\text{s},2\text{s}} - K_{1\text{s},2\text{s}}.$$

Here, the exchange energy enters with a minus sign. This means the triplet state has lower energy than the singlet state. This is an example of the first of *Hund's rules*: The states with highest spin multiplicity ($2S+1$, largest for largest S) have the lowest energy.

Evaluating the Coulomb and exchange integrals for the $1\text{s}^1 2\text{s}^1$ configuration gives

$$J_{1\text{s},2\text{s}} = 0.420$$

and

$$K_{\mathrm{1s,2s}} = 0.044.$$

The triplet and singlet state energies are estimated to be

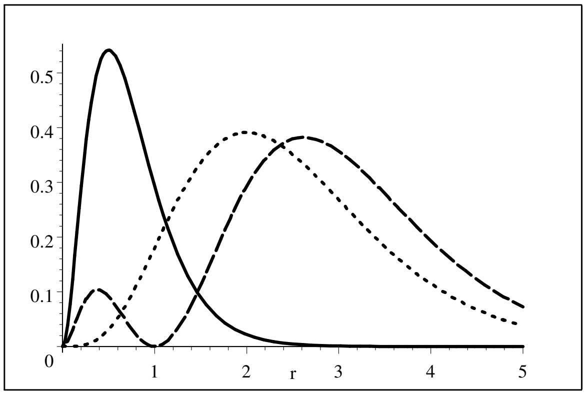
$$\begin{aligned} E^{(0\ \&\ 1)}_{\text{He 1st trip.}} &= -2.5 + 0.420 - 0.044 \\ &= -2.124 \text{ Hartree} = -57.80 \text{ eV} \end{aligned}$$

$$\begin{aligned} E^{(0\ \&\ 1)}_{\text{He 1st sing.}} &= -2.5 + 0.420 + 0.044 \\ &= -2.036 \text{ Hartree} = -55.40 \text{ eV}. \end{aligned}$$

The true values are −2.176 and −2.147 Hartree, respectively.

The next excited states of helium are the triplet and singlet states associated with the $1s^1 2p^1$ electron configuration. These states have total orbital angular momentum quantum number, $L = 1$. There are three distinct, degenerate triplet and singlet states, associated with the three 2p orbitals, $2p_1$, $2p_0$ and $2p_{-1}$. The total degeneracy of these states is $(2S + 1)(2L + 1) = 9$ for the $S = 1$ (triplet) states, and 3 for the $S = 0$ (singlet) states. These states are higher in energy than the $1s^1 2s^1$ states because the Coulomb integral is larger for these states. This energy difference is often attributed to the inner lobe of the 2s penetrating the 1s shell, and experiencing a larger effective nuclear charge. Here, the energy difference is attributed to $J_{\mathrm{1s,2s}} < J_{\mathrm{1s,2p}}$. This inequality, can be understood in terms of the two radial lobes of the 2s orbital as follows:

Figure 7.2 shows the radial electron density for the 1s, 2s and 2p orbitals: $r^2 |R_{1,0}|^2$ (solid), $r^2 |R_{2,0}|^2$ (dashed) and $r^2 |R_{2,1}|^2$ (dotted), respectively. We see that because of the inner lobe of 2s, the outer lobe is further from the nucleus and does not overlap the 1s electron density very much. The overlap of the 2p electron density with 1s is greater. It is in the overlapping regions where the electron-electron repulsion, $1/r_{12}$, is greatest. Note that the traditional first year explanation - that the inner lobe of 2s gives 2s a greater effective nuclear charge - becomes evident if we do a variational calculation. Specifically, if we minimized the expectation value of the Hamiltonian with respect to Z (use separate Z's in the exponential portion of the orbital radial functions), we would find that both 2s and 2p are pushed further from the nucleus (they have smaller effective nuclear charges than 1s), with 2p pushed further in order to avoid the overlap with 1s. This larger shift, to larger r, of 2p gives its a higher energy than 2s.



The radial probability density, $|\varphi(r)|^2 = r^2 |R(r)|^2$, for the 1s orbital (solid), the 2s orbital (dashed) and the 2p orbital (dotted). The units of r are Bohr = 52.918 pm.

Example 7.4. Consider the hydride ion, H^-.

(a) Write down an approximate ground state for H^-, with the independent electron approximation.

(b) Estimate the ground state energy of H^- using first order perturbation theory. Use $J_{1s,1s} = \frac{5}{4}$, for He, to get $J_{1s,1s}$ for H^-. [Hint: write the Coulomb integral in terms of universal hydrogen-like ion orbitals and make the Z dependence explicit.

(c) Estimate the electron affinity of hydrogen.

Solution 7.4.

(a) The independent-electron ground state is identical to that of He, except that the orbitals are for $Z = 1$ rather than $Z = 2$.

$$\begin{aligned}\psi^{(0)}_{H^-\ \mathrm{g.s.}} &= (1s)(1)(1s)(2)\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \alpha(2)\beta(1)) \\ &= (1s)(1)(1s)(2)\chi_{0,0}(1,2).\end{aligned}$$

These are H atom 1s orbitals with $Z = 1$.

(b) This calculation is identical to the calculation above of the ground state energy of helium, except for the value of Z. Specifically,

$$\begin{aligned}
& E^{(0\ \&\ 1)}_{\mathrm{H}^-\ \mathrm{g.s.}} \\
= \;& \underset{\text{total inner product for 2 e}^-\text{s, space and spin}}{\left\langle \psi^{(0)}_{\mathrm{H}^-\ \mathrm{g.s.}} \middle| \hat{H}\psi^{(0)}_{\mathrm{H}^-\ \mathrm{g.s.}} \right\rangle} \\
= \;& \underset{\substack{\text{this is a spatial inner product for 2 e}^-\text{s} \\ \text{a six dimensional integral}}}{\left\langle (1\mathrm{s})(1)(1\mathrm{s})(2) \middle| \hat{H}(1\mathrm{s})(1)(1\mathrm{s})(2) \right\rangle} \quad \underset{\substack{\text{this is a spin inner product for 2 e}^-\text{s} \\ =1,\ \chi_{0,0}(1,2)\text{ is normalized}}}{\left\langle \chi_{0,0}(1,2) \middle| \chi_{0,0}(1,2) \right\rangle} \\
= \;& \left\langle (1\mathrm{s})(1)(1\mathrm{s})(2) \middle| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{12}}\right)(1\mathrm{s})(1)(1\mathrm{s})(2) \right\rangle \\
= \;& \left\langle (1\mathrm{s})(1) \middle| \hat{H}_1(1\mathrm{s})(1) \right\rangle \left\langle (1\mathrm{s})(2) \middle| (1\mathrm{s})(2) \right\rangle + \left\langle (1\mathrm{s})(1) \middle| (1\mathrm{s})(1) \right\rangle \left\langle (1\mathrm{s})(2) \middle| \hat{H}_2(1\mathrm{s})(2) \right\rangle \\
& + \left\langle (1\mathrm{s})(1)(1\mathrm{s})(2) \middle| \frac{1}{r_{12}}(1\mathrm{s})(1)(1\mathrm{s})(2) \right\rangle \\
= \;& -\underset{1\mathrm{s\ e}^-\text{ in H}}{\frac{1^2}{2\times 1^2}} - \underset{1\mathrm{s\ e}^-\text{ in H}}{\frac{1^2}{2\times 1^2}} + J_{1\mathrm{s},1\mathrm{s}} \\
= \;& -1 + J_{1\mathrm{s},1\mathrm{s}}.
\end{aligned}$$

The Coulomb integral, $J_{1\mathrm{s},1\mathrm{s}}$, depends implicitly on Z. We can make the dependence explicit by noting that the hydrogen-like orbitals are universal when expressed in terms of radial coordinate, $\rho = 2Zr/n$. This corresponds to a scaling of both coordinates, $\mathbf{y}_1 = 2Z\mathbf{x}_1$ and $\mathbf{y}_2 = 2Z\mathbf{x}_2$, for 1s orbitals. Therefore,

$$\begin{aligned}
J_{1\mathrm{s},1\mathrm{s};\ Z} &= \int \frac{(1\mathrm{s})^2(\mathbf{y}_1)(1\mathrm{s})^2(\mathbf{y}_2)}{(2Z)^{-1}\rho_{12}}\, d\mathbf{y}_1 d\mathbf{y}_2 \\
&= 2Z \int \frac{(1\mathrm{s})^2(\mathbf{y}_1)(1\mathrm{s})^2(\mathbf{y}_2)}{\rho_{12}}\, d\mathbf{y}_1 d\mathbf{y}_2 \\
&= Z J_{1\mathrm{s},1\mathrm{s};\ 1},
\end{aligned}$$

where these are universal 1s orbitals. Note that factors introduced, upon change of variables, in the volume elements are absorbed into the normalization constant of the universal 1s orbitals. Since $J_{1\mathrm{s},1\mathrm{s}} = 5/4$ for He,

$$\begin{aligned}
J_{1\mathrm{s},1\mathrm{s};\ 2} &= \frac{5}{4} \\
&= 2J_{1\mathrm{s},1\mathrm{s};\ 1}.
\end{aligned}$$

Therefore,

$$J_{1\mathrm{s},1\mathrm{s};\ 1} = \frac{5}{8}$$

is the Coulomb integral for H^-, and

$$E^{(0\ \&\ 1)}_{\mathrm{H}^-\ \mathrm{g.s.}} = -1 + 0.625 = -0.375\ \text{Hartree} = -10.2\ \text{eV}.$$

The true value is -0.52728 Hartree $= -14.348$ eV. This estimate is not too bad.

(c) The electron affinity is the energy change in the process,

$$\mathrm{H} + \mathrm{e}^- \rightarrow \mathrm{H}^-,$$

where the free electron on the left has zero kinetic energy. The electron affinity of hydrogen is thus

$$\begin{aligned} A_{\mathrm{H}} &= E_{\mathrm{H}^-\ \mathrm{g.s.}} - E_{\mathrm{H\ g.s.}} \\ &= -0.52728 - \left(-\frac{1}{2}\right) \\ &= -0.02728 \text{ Hartree} \\ &= -0.7423 \text{ eV} \\ &= -71.6 \text{ kJ mol}^{-1}. \end{aligned}$$

There electron is weakly bound. There is only one bound state of H^-. All excited states are above the threshold for ejection of the electron. These states appears as resonances which decay with a characteristic rate - allowing them to exist for a time.

7.6 Many electron atoms - term symbols

For many electron atoms, the distinct states are found first by identifying the various possible electron configurations, then by identifying the *term symbols* associated with each electron configuration. A term symbol corresponds to a set of states with specific L, S and J. For a given L and S, each distinct J value corresponds to $2J+1$ degenerate states which are different in energy from the other J states due to spin-orbit coupling. Since spin-orbit coupling is small for light elements, the distinct J states are very close in energy. The term symbol consists of a letter specifying the L value: S ($L = 0$), P ($L = 1$), D ($L = 2$), F ($L = 3$), G ($L = 4$), H ($L = 5$), I ($L = 6$), K ($L = 7$), L ($L = 8$), M ($L = 10$), ..., a left superscript giving the spin multiplicity, $2S+1$, and a right subscript giving the J value. For example, ${}^2D_{5/2}$ corresponds to $L = 2$, $S = 1/2$ and $J = L + S = 5/2$. ${}^2D_{3/2}$ is the other term symbol associated with $L = 2$ and $S = 1/2$. The term symbol letters (the first five are well known by chemistry students) can be remembered using the mnemonic, "Sober physicists don't find giraffes hiding in kitchens like mine".

The spin and orbital angular momenta are vectors that can add to make longer or shorter vectors as long as $|L - S| \leq J \leq L + S$. All J values subject to this constraint occur; i.e., $J = |L - S|, |L - S| + 1, \ldots L + S$. The term symbols of the states of helium discussed above are ${}^1\mathrm{S}_0$ for the ground state, ${}^3\mathrm{S}_1$ and ${}^1\mathrm{S}_0$ for the triplet and singlet states associated with the $1\mathrm{s}^1 2\mathrm{s}^1$ electron configuration, and ${}^3\mathrm{P}$ ($= {}^3\mathrm{P}_2$, ${}^3\mathrm{P}_1$ and ${}^3\mathrm{P}_0$) and ${}^1\mathrm{P}_1$ for the triplet and singlet states associated with the $1\mathrm{s}^1 2\mathrm{p}^1$ electron configuration. The term symbols are easy to deduce for these electron configurations. However, for configurations like 2p^2 or 3d^2 many term symbols arise. This is because there are three p orbitals and 5 d orbitals which can be occupied by two electrons in many different ways.

To get the term symbols (just get the L and S values - distinct J term symbols can be found after all the L and S values are identified) associated with an electron configuration, one must consider all the ways the electrons can occupy the given orbitals without violating the Pauli principle - i.e., all the *box diagrams*. For each box diagram, add up the individual electron m_ℓ values to get M_L, and the m_s values to get M_S. Since every $M_L < 0$ state is accompanied by a $-M_L$ state, it is sufficient to consider only states with $M_L \geq 0$. Similarly, we can restrict ourselves to $M_S \geq 0$. Consider the p^2 electron configuration box diagrams:

	p_1	p_0	p_{-1}	M_L	M_S
^{1}D	↑↓			2	0
^{1}D		↑↓		0	0
^{3}P	↑	↑		1	1
^{3}P	↑		↑	0	1
^{1}D	↑	↓		1	0
^{3}P	↑		↓	0	0
^{1}S	↓		↑	0	0
^{3}P	↓	↑		1	0

The next step is to identify the largest M_L value (and the largest M_S value that comes with that M_L value). Here, the largest M_L value is +2. The largest M_S value with $M_L = 2$ is 0. This means that there is a ^{1}D term (specifically, ^{1}D$_2$). Cross out the rows that come with this symbol - three rows with $M_L = 2, 1, 0$ (all with $M_S = 0$). These rows are labeled on the left with ^{1}D. Repeat the process with the remaining rows - i.e., find the next largest M_L and M_S, identify the term symbol and cross out the rows that come with this symbol. Here, we find $M_L = 1$ and $M_S = 1$ give the next term symbol found - namely, ^{3}P. This symbol corresponds to $M_L = 1$ and 0 rows for each of $M_S = 1$ and 0 - four rows are crossed out. The last remaining row has $M_L = 0$ and $M_S = 0$. It is the ^{1}S term. Thus, p^2 = ^{1}D + ^{3}P + ^{1}S = ^{1}D$_2$ + ^{3}P$_0$ + ^{3}P$_1$ + ^{3}P$_2$ + ^{1}S$_0$.

The p^4 electron configuration produces the same terms as p^2. To see this, simply replace every row in a completed p^2 table with a filled p subshell, p^6, then remove an electron from each orbital with an electron in the p^2 table (spin up or down as in the p^2 table). This produces a table identical to the p^2 table except that the signs of M_L and M_S are reversed. This sign reversal has no impact on the identification of terms since each term has both plus and minus (or zero) M_L and M_S values. Similarly, the p^5 configuration produces the same terms as p^1 - namely, ^{2}P (= ^{2}P$_{1/2}$ and ^{2}P$_{3/2}$).

If the two electrons are in different subshells, it is easy to get the terms in this case. All

$$(2L_1 + 1)\,(2L_2 + 1)\,(2S_1 + 1)\,(2S_2 + 1)$$

states of the two electron system are admissable. For example, the $2p^1 3p^1$ electron configuration has term symbols, ^{1}D, ^{3}D,^{1}P, ^{3}P,^{1}S, ^{3}S. These are the possible term symbols with $L = |1-1| = 0$ to $1+1=2$, and $S = \left|\frac{1}{2}-\frac{1}{2}\right| = 0$ to $\frac{1}{2}+\frac{1}{2}=1$. The number of states represented by each of these term symbols is 5×1, 5×3, 3×1, 3×3, 1×1 and 1×3, respectively. The total number of states is

$$\begin{aligned} 5+15+3+9+1+3 &= 36 \\ &= 3\times 3\times 2\times 2 \end{aligned}$$

in this case. The second line verifies that all states are taken into account.

Once the term symbols are determined, *Hund's rules* determine the ordering of the term symbols in energy.

Rule 1. Terms are ordered according to spin multiplicity, with the highest multiplicity terms having the lowest energy.

Rule 2. Terms with the same multiplicity are ordered according to orbital angular momentum, with the highest L states having the lowest energy.

Rule 3. The J levels within each term are ordered with (a) the lowest J giving the lowest energy, if the outermost subshell is half-filled or less, and (b) the highest J giving the lowest energy, if the outermost subshell is more than half-filled.

The first rule is a generalization of the effect seen above for the excited states of helium. The triplet was the lowest energy excited state because the exchange integral appeared with a minus sign. This resulted because the orbital state was antisymmetric. An antisymmetric orbital state is small whenever the electrons are near. The electrons are generally further apart in the anitsymmetric orbital state, and consequently have reduced Coulomb repulsion. Higher multiplicity terms have more unpaired electrons. The orbital state must be antisymmetric with respect to the exchange of any two unpaired electrons. Higher multiplicity terms have orbital states that are more highly antisymmetric, with greater associated reduction in repulsion energy. This is the origin of Hund's rule 1. Hund's rule 2 results because the electrons are more spread out when they have a larger orbital angular momentum. This reduces the repulsion energy. Different electron repulsion is the only difference in energy between terms with the same electron configuration. Hund's rule 3 is more subtle.

For many electron atoms, the electron configuration can be very complex. However, the ground state and lowest excited states involve electron configurations with subshells closed as much as possible. Each closed subshell is equivalent to an empty subshell (with respect to term symbols), and has only the term, ^{1}S. This corresponds to zero orbital and spin angular momentum. Thus, in compound electron configurations, closed shells can be ignored when determining term symbols. For example, consider the ground state electron configuration of nitrogen, $1s^2 2s^2 2p^3$. Only the 2p subshell is not closed. The allowed term symbols are those of the p^3 electron configuration, determined by the following table. Rows are crossed out by attributing them to a term symbol in the order specified by the rules - the term symbol is indicated on the left.

	p_1	p_0	p_{-1}	M_L	M_S
^{2}D	$\uparrow\downarrow$	$\uparrow$		2	$\frac{1}{2}$
^{2}D	$\uparrow\downarrow$		$\uparrow$	1	$\frac{1}{2}$
^{2}P	$\uparrow$	$\uparrow\downarrow$		1	$\frac{1}{2}$
^{4}S	$\uparrow$	$\uparrow$	$\uparrow$	0	$\frac{3}{2}$
^{2}D	$\uparrow$	$\uparrow$	$\downarrow$	0	$\frac{1}{2}$
^{2}P	$\uparrow$	$\downarrow$	$\uparrow$	0	$\frac{1}{2}$
^{4}S	$\downarrow$	$\uparrow$	$\uparrow$	0	$\frac{1}{2}$

The table shows that $p^3 = {}^2D + {}^2P + {}^4S$. Hund's rules identify the ^{4}S term as the ground state. Specifically, the ground state consists of the four degenerate $S = 3/2$ states, with $L = 0$. The ground state of nitrogen is a quartet state. The first excited state is ^{2}D, while ^{2}P gives the next excited state. The first excited quartet state is associated with an excited electron configuration. Ordering the excited states, and computing their energies, permits the assignment of atomic emission spectra - subject to the electric dipole selection rules. In the case of Russell-Saunders coupling, the selection rules are $\Delta\ell = \pm1, \Delta L = 0, \pm1, \Delta J = 0, \pm1$ (except not $J = 0$ to $J = 0$) and $\Delta S = 0$. Here, the electron configuration must change with a transition of one electron with $\Delta\ell = \pm1$. After that, because the photon either adds or removes a quantum of angular momentum, both the orbital and total angular momentum quantum numbers cannot change by more than 1.

Atomic emission spectroscopy is routinely used to do elemental analysis of metals. The metal is vaporized by a high voltage arc. Atoms in the arc emit characteristic frequencies of light, in characteristic proportions. Light from each element can be quantified to give the composition of the sample - to ensure desired component elements and impurities have abundances within specifications.

Example 7.5. Find the terms associated with the following electron configurations. Order them with increasing energy, according to Hund's rules. Determine the total number of states, and break it down according to the number of states in each term.

(a) pd

(b) d^8

Solution 7.5.

(a) The two electrons are in different subshells. All

$$(2L_1 + 1)(2L_2 + 1) = 3 \times 5 = 15$$

orbital states of the two electron system are admissable, for all

$$(2S_1 + 1)(2S_2 + 1) = 2 \times 2 = 4$$

spin states. Thus, L varies from $|L_1 - L_2| = 1$ to $L_1 + L_2 = 3$. For each L value, S varies from $|S_1 - S_2| = 0$ to $S_1 + S_2 = 1$. Consequently, the terms are $^3\mathrm{F}_2$,$^3\mathrm{F}_3$, $^3\mathrm{F}_4$, $^1\mathrm{F}_3$, $^3\mathrm{D}_1$, $^3\mathrm{D}_2$, $^3\mathrm{D}_3$, $^1\mathrm{D}_2$, $^3\mathrm{P}_0$, $^3\mathrm{P}_1$, $^3\mathrm{P}_2$ and $^1\mathrm{P}$ - listed in order of increasing energy according to Hund's rule's. The total number of states is $15 \times 4 = 60$. We will check that this number emerges when we tally states for each term. For the $^3\mathrm{F}$, $^1\mathrm{F}$, $^3\mathrm{D}$ and $^1\mathrm{D}$ terms, there are $3 \times 7 = 21$, $1 \times 7 = 7$, $3 \times 5 = 15$, $1 \times 5 = 5$, $3 \times 3 = 9$ and $1 \times 3 = 3$ states, respectively. These numbers add up to 60. The 21 states of $^3\mathrm{F}$ are further broken into $5 + 7 + 9$ states, with these numbers corresponding to the three subterms, $J = 2, 3$ and 4. The other terms have similar breakdowns into subterms.

(b) The d^8 configuration is the same as d^2, except for the energy ordering of subterms - Hund's rule 3. The two electrons are in the same subshell. The d^2 electron configuration box diagrams:

	d_2	d_1	d_0	d_{-1}	d_{-2}	M_L	M_S
$^1\mathrm{G}$	$\uparrow\downarrow$					4	0
$^1\mathrm{G}$	$\uparrow$	$\downarrow$				3	0
$^3\mathrm{F}$	$\downarrow$	$\uparrow$				3	0
$^1\mathrm{G}$	$\uparrow$		$\downarrow$			2	0
$^3\mathrm{F}$	$\downarrow$		$\uparrow$			2	0
$^1\mathrm{G}$	$\uparrow$			$\downarrow$		1	0
$^3\mathrm{F}$	$\downarrow$			$\uparrow$		1	0
$^1\mathrm{G}$	$\uparrow$				$\downarrow$	0	0
$^3\mathrm{F}$	$\downarrow$				$\uparrow$	0	0
$^1\mathrm{D}$		$\uparrow\downarrow$				2	0
$^1\mathrm{D}$		$\uparrow$	$\downarrow$			1	0
$^3\mathrm{P}$		$\downarrow$	$\uparrow$			1	0
$^1\mathrm{D}$		$\uparrow$		$\downarrow$		0	0
$^3\mathrm{P}$		$\downarrow$		$\uparrow$		0	0
$^1\mathrm{S}$			$\uparrow\downarrow$			0	0
$^3\mathrm{F}$	$\uparrow$	$\uparrow$				3	1
$^3\mathrm{F}$	$\uparrow$		$\uparrow$			2	1
$^3\mathrm{F}$	$\uparrow$			$\uparrow$		1	1
$^3\mathrm{F}$	$\uparrow$				$\uparrow$	0	1
$^3\mathrm{P}$		$\uparrow$	$\uparrow$			1	1
$^3\mathrm{P}$		$\uparrow$		$\uparrow$		0	1

Following the rules, we identify the terms in the following order: $^1\mathrm{G}$, $^3\mathrm{F}$, $^1\mathrm{D}$, $^3\mathrm{P}$ and $^1\mathrm{S}$. Hund's rules orders these terms (and subterms), according to increasing energy, as follows: $^3\mathrm{F}_4 < {}^3\mathrm{F}_3 < {}^3\mathrm{F}_2 < {}^3\mathrm{P}_2 < {}^3\mathrm{P}_1 < {}^3\mathrm{P}_0 < {}^1\mathrm{G}_4 < {}^1\mathrm{D}_2 < {}^1\mathrm{S}_0$. The number of states for these symbols is 9, 7, 5, 5, 5, 3, 1, 9, 5, 1. These add up to 45. By considering only $M_l > 0$ and $M_S > 0$, we cannot simply tally the number of rows in the table. The total number of states is more easily ascertained from the terms - not subterms. The terms, ^{3}F, ^{3}P, ^{1}G, ^{1}D and ^{1}S have 3×7, 3×3, 1×9, 1×5 and 1×1 states, respectively. These add up to 45.

Chapter 8

Electronic structure of molecules

8.1 The hydrogen molecule

Atoms form covalent bonds with shared valence electrons. The simplest molecule is H_2^+, with only one electron, held together by the attraction of the electron to the two nuclei. We begin with the next simplest molecule, H_2.

Consider two hydrogen atoms, in their ground states, approaching each other. The electron in each atom experiences Coulombic attraction to the other nucleus. There are also repulsions between the nuclei, and between the electrons. At large distance, the net effect is always attractive - each atom is an electric dipole, and dipole-dipole interactions lower energy. At distance below the bond length, the net force is repulsive. The bond length is the equilibrium distance where the energy is a minimum.

To model the H_2 molecule using quantum mechanics, we introduce *molecular orbitals* - orbitals spread over two or more atoms. We use atomic orbitals as basis functions in which to expand the molecular orbitals. This approach is called LCAO - *linear combination of atomic orbitals.* Thus,

$$\psi_+ (1) = N^{-1/2} \left((1\mathrm{s})_\mathrm{A} (1) + (1\mathrm{s})_\mathrm{B} (1) \right),$$

is a molecular orbital constructed as the superposition of states $(1\mathrm{s})_\mathrm{A}$ and $(1\mathrm{s})_\mathrm{B}$ - the (1s) states of atoms A and B, respectively. The normalization constant here is not 2 because $(1\mathrm{s})_\mathrm{A}$ and $(1\mathrm{s})_\mathrm{B}$ are not orthogonal when the atoms are not far apart. Instead,

$$\begin{aligned}
\langle \psi_+ | \psi_+ \rangle &= 1 \\
&= N^{-1} \langle (1\mathrm{s})_\mathrm{A} + (1\mathrm{s})_\mathrm{B} | (1\mathrm{s})_\mathrm{A} + (1\mathrm{s})_\mathrm{B} \rangle \\
&= N^{-1} \Big(\underset{=1}{\langle (1\mathrm{s})_\mathrm{A} | (1\mathrm{s})_\mathrm{A} \rangle} + 2 \underset{=u}{\langle (1\mathrm{s})_\mathrm{A} | (1\mathrm{s})_\mathrm{B} \rangle} + \underset{=1}{\langle (1\mathrm{s})_\mathrm{B} | (1\mathrm{s})_\mathrm{B} \rangle} \Big) \\
&= N^{-1} (2 + 2u),
\end{aligned}$$

where

$$u = \langle (1\mathrm{s})_{\mathrm{A}}\ | (1\mathrm{s})_{\mathrm{B}} \rangle$$

is called the overlap integral, and

$$N = 2\,(1+u)\,.$$

To get the ground state of H_2, we use the independent electron approximation, and write the two electron wavefunction as the product of the symmetric orbital state, $\psi\,(1)\,\psi\,(2)$, and the antisymmetric singlet spin state, $\chi_{0,0}\,(1,2)$. Here, ψ is some appropriate lowest energy molecular orbital.

$$\psi_{\mathrm{H}_2\ \mathrm{g.s.}}\,(1,2) = \psi\,(1)\,\psi\,(2)\,\chi_{0,0}\,(1,2)\,.$$

The molecular orbital, ψ, is written as

$$\psi\,(1) = c_1 (1\mathrm{s})_{\mathrm{A}}\,(1) + c_2 (1\mathrm{s})_{\mathrm{B}}\,(1)\,.$$

The coefficients are determined by substituting this expression into the expectation value of the Hamiltonian, and minimizing with respect to variations in the coefficients. This gives the best possible wavefunction of this form. The Hamiltonian for H_2 is

$$\begin{aligned} \hat{H} &= \hat{H}_1 + \hat{H}_2 + \frac{1}{r_{\mathrm{A,\ B}}} + \frac{1}{r_{12}} \qquad (8.1) \\ &= \hat{H}_{\mathrm{A,\ 1}} - \frac{1}{r_{\mathrm{B,\ 1}}} + \hat{H}_{\mathrm{B,\ 2}} - \frac{1}{r_{\mathrm{A,\ 2}}} + \frac{1}{r_{\mathrm{A,\ B}}} + \frac{1}{r_{12}}. \end{aligned}$$

In the first line, $\hat{H}$ is expressed as the sum of one electron Hamiltonians, $\hat{H}_1$ and $\hat{H}_2$ - one for each electron. In the second line, $\hat{H}_1$ is written as the H atom Hamiltonian for electron 1 in atom A, $\hat{H}_{\mathrm{A,\ 1}}$, plus the attraction to nucleus B. Similarly, electron 2 is associated with atom B, with the other - equivalent - attraction written separately. $r_{\mathrm{B,\ 1}}$ is the distance from nucleus B to electron 1, and similarly for $r_{\mathrm{A,\ 2}}$. $\hat{H}_1$ and $\hat{H}_2$ can also be written with electron 1 associated with atom B and electron 2 associated with atom A. We will use which ever way is most convenient.

The expectation value of the Hamiltonian is given by

$$\begin{aligned}
& \left\langle \psi_{\mathrm{H_2\ g.s.}} \left| \hat{H} \psi_{\mathrm{H_2\ g.s.}} \right\rangle \right. \\
= {} & \left\langle \psi(1)\,\psi(2)\,\chi_{0,0}(1,2) \left| \hat{H}\psi(1)\,\psi(2)\,\chi_{0,0}(1,2) \right\rangle \right. \\
= {} & \left\langle \psi(1)\,\psi(2) \left| \hat{H}\psi(1)\,\psi(2) \right\rangle \right. \\
= {} & \left\langle \psi(1)\,\psi(2) \left| \left(\hat{H}_1 + \hat{H}_2 + \frac{1}{r_{\mathrm{A,\ B}}} + \frac{1}{r_{12}} \right) \psi(1)\,\psi(2) \right\rangle \right. \\
= {} & \left\langle \psi(1)\,\psi(2) \left| \hat{H}_1\psi(1)\,\psi(2) \right\rangle \right. + \left\langle \psi(1)\,\psi(2) \left| \hat{H}_2\psi(1)\,\psi(2) \right\rangle \right. \\
& + \underbrace{\left\langle \psi(1)\,\psi(2) \left| \frac{1}{r_{\mathrm{A,\ B}}}\psi(1)\,\psi(2) \right\rangle \right.}_{=\frac{1}{r_{\mathrm{A,\ B}}}\langle\psi(1)\psi(2)\,|\psi(1)\psi(2)\rangle=\frac{1}{r_{\mathrm{A,\ B}}}} + \left\langle \psi(1)\,\psi(2) \left| \frac{1}{r_{12}}\psi(1)\,\psi(2) \right\rangle \right. \\
= {} & \left\langle \psi(1) \left| \hat{H}_1\psi(1) \right\rangle \right. \underbrace{\langle \psi(2)\,|\psi(2)\rangle}_{=1} \\
& + \underbrace{\langle \psi(1)\,|\psi(1)\rangle}_{=1} \left\langle \psi(2) \left| \hat{H}_2\psi(2) \right\rangle \right. + \frac{1}{r_{\mathrm{A,\ B}}} + J_{\psi,\psi} \\
= {} & \left\langle \psi(1) \left| \hat{H}_1\psi(1) \right\rangle \right. + \underbrace{\left\langle \psi(2) \left| \hat{H}_2\psi(2) \right\rangle \right.}_{\substack{= \left\langle \psi(1) \left| \hat{H}_1\psi(1) \right\rangle \right. \\ \text{only the electron label is different}}} + \frac{1}{r_{\mathrm{A,\ B}}} + J_{\psi,\psi} \\
= {} & 2 \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} \left\langle (1\mathrm{s})_\mathrm{A} \left| \hat{H}_1 (1\mathrm{s})_\mathrm{A} \right\rangle \right. & \left\langle (1\mathrm{s})_\mathrm{A} \left| \hat{H}_1 (1\mathrm{s})_\mathrm{B} \right\rangle \right. \\ \left\langle (1\mathrm{s})_\mathrm{B} \left| \hat{H}_1 (1\mathrm{s})_\mathrm{A} \right\rangle \right. & \left\langle (1\mathrm{s})_\mathrm{B} \left| \hat{H}_1 (1\mathrm{s})_\mathrm{B} \right\rangle \right. \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
& + \frac{1}{r_{\mathrm{A,\ B}}} + J_{\psi,\psi}. \qquad (8.2)
\end{aligned}$$

For now, we neglect $J_{\psi,\psi}$. It can be put back into the energy - in approximate fashion - using first perturbation.

The expectation value is written above as a quadratic form. When the molecular orbital is written as a linear combination of the two atomic orbitals, it is given a two dimensional vector representation - the two coefficients are expressed here as a column vector. The one electron Hamiltonians become the same 2 × 2 matrix in this representation.

Consider the matrix representation of $\hat{H}_1$ (or $\hat{H}_2$)

$$
\begin{aligned}
& \begin{pmatrix}
\left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \hat{H}_1 (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \hat{H}_1 (1\mathrm{s})_{\mathrm{B}} \right\rangle \\
\left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \hat{H}_1 (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \hat{H}_1 (1\mathrm{s})_{\mathrm{B}} \right\rangle
\end{pmatrix} \\
= & \begin{pmatrix}
\left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \left(\hat{H}_{\mathrm{A},\,1} - \frac{1}{r_{\mathrm{B},\,1}}\right) (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \left(\hat{H}_{\mathrm{B},\,1} - \frac{1}{r_{\mathrm{A},\,1}}\right) (1\mathrm{s})_{\mathrm{B}} \right\rangle \\
\left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \left(\hat{H}_{\mathrm{A},\,1} - \frac{1}{r_{\mathrm{B},\,1}}\right) (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \left(\hat{H}_{\mathrm{B},\,1} - \frac{1}{r_{\mathrm{A},\,1}}\right) (1\mathrm{s})_{\mathrm{B}} \right\rangle
\end{pmatrix} \\
= & \begin{pmatrix}
\left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \hat{H}_{\mathrm{A},\,1} (1\mathrm{s})_{\mathrm{A}} \right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \frac{-1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \hat{H}_{\mathrm{B},\,1} (1\mathrm{s})_{\mathrm{B}} \right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \frac{-1}{r_{\mathrm{A},\,1}} (1\mathrm{s})_{\mathrm{B}} \right\rangle \\
\left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \hat{H}_{\mathrm{A},\,1} (1\mathrm{s})_{\mathrm{A}} \right\rangle + \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \frac{-1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| \hat{H}_{\mathrm{B},\,1} (1\mathrm{s})_{\mathrm{B}} \right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| \frac{-1}{r_{\mathrm{A},\,1}} (1\mathrm{s})_{\mathrm{B}} \right\rangle
\end{pmatrix} \\
= & \begin{pmatrix}
E_1^{(0)} + \alpha_{\mathrm{A,B}} & E_1^{(0)} u + \beta_{\mathrm{B,A}} \\
E_1^{(0)} u + \beta_{\mathrm{A,B}} & E_1^{(0)} + \alpha_{\mathrm{B,A}}
\end{pmatrix} \\
= & \begin{pmatrix}
-\frac{1}{2} + \alpha_{\mathrm{A,B}} & -\frac{1}{2} u + \beta_{\mathrm{A,B}} \\
-\frac{1}{2} u + \beta_{\mathrm{A,B}} & -\frac{1}{2} + \alpha_{\mathrm{A,B}}
\end{pmatrix} \qquad \begin{array}{c} \alpha_{\mathrm{B,A}} = \alpha_{\mathrm{A,B}} \text{ and } \beta_{\mathrm{B,A}} = \beta_{\mathrm{A,B}} \\ \text{both atoms are H atoms} \end{array} \\
= & \begin{pmatrix}
\alpha' & \beta' \\
\beta' & \alpha'
\end{pmatrix}.
\end{aligned}
$$

The expectation value of the Hamiltonian - with electron repulsion neglected - takes the form,

$$
\begin{aligned}
& \left\langle \psi_{\mathrm{H_2\ g.s.}} \middle| \hat{H}^{(0)} \psi_{\mathrm{H_2\ g.s.}} \right\rangle \\
= \; & 2 \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ \beta' & \alpha' \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \frac{1}{r_{\mathrm{A,\,B}}} \\
= \; & 2 \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ \beta' & \alpha' \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \; + \frac{1}{r_{\mathrm{A,\,B}}} \\
= \; & 2\mathbf{c}^T \mathsf{H}' \mathbf{c} + \frac{1}{r_{\mathrm{A,\,B}}} \qquad (8.3)
\end{aligned}
$$

Since $1/r_{\mathrm{A,\,B}}$ is a constant, it can be left out of the minimization, then put back into the total energy at the end. Our task is to find the vector which minimizes $2\mathbf{c}^T\mathsf{H}'\mathbf{c}$, subject to the constraint that the molecular orbital be normalized. The constraint has the form,

$$
\begin{aligned}
1 \; = \; & \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{A}} \middle| (1\mathrm{s})_{\mathrm{B}} \right\rangle \\ \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| (1\mathrm{s})_{\mathrm{A}} \right\rangle & \left\langle (1\mathrm{s})_{\mathrm{B}} \middle| (1\mathrm{s})_{\mathrm{B}} \right\rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
= \; & \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} 1 & u \\ u & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
= \; & \mathbf{c}^T \mathsf{U} \mathbf{c}
\end{aligned}
$$

Matrix U is called the overlap matrix.

The constrained minimization problem is solved using the method of *Lagrange multipliers*. Since there is one constraint, we introduce one multiplier,

λ. The method works by minimizing

$$\begin{aligned} F &= 2\mathbf{c}^T \mathsf{H}'\mathbf{c} + \lambda\left(1 - \mathbf{c}^T \mathsf{U}\mathbf{c}\right) \\ &= 2\sum_{j=1}^{2}\sum_{k=1}^{2} c_j \mathsf{H}'_{j,k} c_k + \lambda\left(1 - \sum_{j=1}^{2}\sum_{k=1}^{2} c_j \mathsf{U}_{j,k} c_k\right) \end{aligned}$$

with respect to variation of $\mathbf{c}$ and λ. Setting $\partial F/\partial\lambda = 0$ returns the constraint equation. Setting $\partial F/\partial c_n = 0$ gives

$$\begin{aligned} 0 &= 2\sum_{j=1}^{2}\sum_{k=1}^{2}\left(\delta_{j,n}\mathsf{H}'_{j,k}c_k + c_j\mathsf{H}'_{j,k}\delta_{k,n}\right) - \lambda\sum_{j=1}^{2}\sum_{k=1}^{2}\left(\delta_{j,n}\mathsf{U}_{j,k}c_k + c_j\mathsf{U}_{j,k}\delta_{k,n}\right) \\ &= 2\left(\sum_{k=1}^{2}\mathsf{H}'_{n,k}c_k + \sum_{j=1}^{2}c_j\mathsf{H}'_{j,n}\right) - \lambda\left(\sum_{k=1}^{2}\mathsf{U}_{n,k}c_k + \sum_{j=1}^{2}c_j\mathsf{U}_{j,n}\right) \\ &= 4\sum_{k=1}^{2}\mathsf{H}'_{n,k}c_k - 2\lambda\sum_{k=1}^{2}\mathsf{U}_{n,k}c_k \qquad \mathsf{H}' \text{ and } \mathsf{U} \text{ are symmetric} \end{aligned}$$

where

$$\delta_{k,n} = \begin{cases} 1, & k = n \\ 0, & k \neq n \end{cases} .$$

Writing the above equation in vector notation - $n = 1, 2$ - gives

$$0 = 2\mathsf{H}'\mathbf{c} - \lambda\mathsf{U}\mathbf{c},$$

or

$$\mathsf{H}'\mathbf{c} = \frac{\lambda}{2}\mathsf{U}\mathbf{c}.$$

This is called a *generalized eigenvalue equation.* $\lambda/2$ is the eigenvalue. It can be converted into an eigenvalue equation by letting

$$\mathbf{b} = \mathsf{U}^{1/2}\mathbf{c},$$

and finding

$$\mathsf{H}'\mathsf{U}^{-1/2}\mathbf{b} = \frac{\lambda}{2}\mathsf{U}^{1/2}\mathbf{b},$$

or

$$\begin{aligned} \mathsf{U}^{-1/2}\mathsf{H}'\mathsf{U}^{-1/2}\mathbf{b} &= \frac{\lambda}{2}\mathbf{b} \\ \mathsf{H}\mathbf{b} &= \epsilon\mathbf{b}, \end{aligned}$$

where

$$\begin{aligned} \mathsf{H} &= \mathsf{U}^{-1/2}\mathsf{H}'\mathsf{U}^{-1/2} \\ &= \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}. \end{aligned}$$

$\mathsf{H}\mathbf{b} = \epsilon\mathbf{b}$ is a regular *eigenvalue equation.* Here, $\epsilon = \lambda/2$ is an eigenvalue of H, the one electron Hamiltonian with $\mathsf{U}^{-1/2}$ on either side. $\mathsf{U}^{-1/2}$ exists because U is positive definite (all its eigenvalues are greater than zero). $\lambda = 2\epsilon$ is the energy of two electrons in the associated molecular orbital - neglecting electron repulsion. Substituting eigenvector $\mathbf{b}_1$, associated with lowest eigenvalue ϵ_1, into the $\hat{H}^{(0)}$ expectation - including the constant $1/r_{\mathrm{A,\ B}}$ term gives

$$\begin{aligned}
E^{(0)}_{\mathrm{H_2\ g.s.}} &\cong \left\langle \psi_{\mathrm{H_2\ g.s.}} \middle| \hat{H}^{(0)} \psi_{\mathrm{H_2\ g.s.}} \right\rangle \\
&= 2\mathbf{b}_1^T \mathsf{H} \mathbf{b}_1 + \frac{1}{r_{\mathrm{A,\ B}}} \\
&= 2\epsilon_1 + \frac{1}{r_{\mathrm{A,\ B}}},
\end{aligned}$$

the minimum expectation value possible with the atomic basis set of two states. The ground state of H_2 has two electrons in the lowest energy molecular orbital.

The molecular orbital energy levels are the eigenvalues of H. The associated determinantal equation is

$$\begin{aligned}
0 &= \begin{vmatrix} \alpha - \epsilon & \beta \\ \beta & \alpha - \epsilon \end{vmatrix} \\
&= (\alpha - \epsilon)^2 - \beta^2
\end{aligned}$$

or

$$\alpha - \epsilon = \pm\beta$$

which gives

$$\epsilon_1 = \alpha + \beta$$

and

$$\epsilon_2 = \alpha - \beta.$$

ϵ_1 is the lowest energy because α and β are both negative.

The eigenvector associated with eigenvalue, ϵ, is given from the first row of the eigenvalue equation,

$$\begin{pmatrix} \alpha - \epsilon & \beta \\ \beta & \alpha - \epsilon \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

i.e.,

$$(\alpha - \epsilon)\, b_1 + \beta b_2 = 0,$$

or

$$b_1 = \frac{\beta}{\epsilon - \alpha} b_2.$$

For the lowest energy eigenvalue,

$$b_{1,1} = b_{2,1},$$

and the normalized eigenvector is

$$\begin{pmatrix} b_{1,1} \\ b_{2,1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

To get the associated molecular orbital, we must apply $\mathsf{U}^{-1/2}$ to get the atomic orbital coefficients. Specifically,

$$\mathbf{c}_1 = \mathsf{U}^{-1/2}\mathbf{b}_1.$$

However, in this case, the eigenvectors of H are also eigenvectors U. This is because U and H have the same structure. U corresponds to replacing α by 1 and β by u, in H. The eigenvalues of U associated with $\mathbf{b}_1$ and $\mathbf{b}_2$ are $1+u$ and $1-u$. Consequently,

$$\mathbf{c}_1 = (1+u)^{-1/2}\,\mathbf{b}_1$$

and

$$\mathbf{c}_2 = (1-u)^{-1/2}\,\mathbf{b}_2.$$

Applying $\mathsf{U}^{-1/2}$, in this case, simply corrects the normalization, accounting for the non-zero overlap of the basis functions. Otherwise, the $\mathbf{c}$ vectors are the same as the $\mathbf{b}$ vectors.

In general, when there are more atomic orbitals, and - in particular - different atomic orbitals, the $\mathbf{c}$ vectors will be different from the $\mathbf{b}$ vectors. However, aside from correcting the normalization, $\mathsf{U}^{-1/2}$ only mixes coefficents of overlapping (i.e., neighboring) orbitals. In particular, each $\mathbf{b}$ unit vector is the superposition of the associated $\mathbf{c}$ unit vector with small amounts of the $\mathbf{c}$ unit vectors associated with the overlapping orbitals. In practice, we work with the $\mathbf{b}$ vectors, and interpret them as though they were the coefficients of the atomic orbitals.

The lowest energy molecular orbital found above,

$$\psi_1(1) = (\sigma 1\mathrm{s})(1) = (2(1+u))^{-1/2}\left((1\mathrm{s})_{\mathrm{A}}(1) + (1\mathrm{s})_{\mathrm{B}}(1)\right), \tag{8.4}$$

is the σ1s molecular orbital. It is a *bonding molecular orbital* which has constructive interference, between the two atomic orbitals, in the space between the atoms. It is a σ orbital because there is zero orbital angular momentum about the axis of the molecule. Diatomic molecules are not spherically symmetric. However, they are cylindrically symmetric. Consequently, the quantum number associated with the angular momentum component along the molecular axis is a good quantum number. This quantum number is called Λ. States of the diatomic molecular are labeled as Σ, Π, Δ, $\ldots$, for $\Lambda = 0$, 1, 2, and so on. The quantum number specific to individual molecular orbitals is given the associated lower case Greek letter: σ, π, δ, $\ldots$. σ bonds are formed from atomic orbitals with zero angular momentum about the molecular axis - they are both cylindrically symmetric.

The other molecular orbital,

$$\psi_2(1) = (\sigma^*1\mathrm{s})(1) = (2(1-u))^{-1/2}\left((1\mathrm{s})_\mathrm{A}(1) - (1\mathrm{s})_\mathrm{B}(1)\right), \tag{8.5}$$

is the σ^*1s *antibonding molecular orbital* - here, * does not represent complex conjugation. The two atomic orbitals destructively interfere between the atoms, and the energy associated with this orbital, $\epsilon_2 = \alpha - \beta$, is higher than the energy of either atomic orbital. Note that α is the expectation value of H associated with unit $\mathbf{b}$ vectors, $\mathbf{b}^T = (1,0)$ or $(0,1)$. To the extent that these vectors correspond to the atomic orbitals, we interpret α as the 1s orbital energy. This is a reference energy. The energy of the $\sigma1$s orbital, $\epsilon_1 = \alpha + \beta$, is lower than α. An electron in a bonding orbital is stabilized relative to the atomic orbitals. In contrast, an electron in an antibonding bonding orbital is destabilized relative to the atomic orbitals. The destabilization of the antibonding orbital equals the stabilization of the bonding orbital.

Since $\sigma1$s has capacity for two electrons, and H_2 has two electrons, the ground state electron configuration of H_2 is $(\sigma1\mathrm{s})^2$.

The orbital energy parameters, α and β, were introduced in a series of steps above. In terms of the matrix elements of the attraction of a hydrogen atom electron to a second nucleus, $\alpha_{A,B}$ and $\beta_{A,B}$, we have

$$\begin{aligned}\alpha &= \frac{-\frac{1}{2} + \alpha_{A,B} - \left(-\frac{1}{2}u + \beta_{A,B}\right)u}{1-u^2} \\ \beta &= \frac{-\left(-\frac{1}{2} + \alpha_{A,B}\right)u - \frac{1}{2}u + \beta_{A,B}}{1-u^2}.\end{aligned}$$

With electron repulsion neglected, the ground state energy of the hydrogen molecule is approximated by

$$\begin{aligned}E^{(0)}_{\mathrm{H_2\ g.s.}} &= 2\epsilon_1 + \frac{1}{r_{\mathrm{A,\ B}}} \\ &= 2(\alpha+\beta) + \frac{1}{r_{\mathrm{A,\ B}}} \\ &= \frac{-\left(1-u^2\right) + 2(1-u)\left(\alpha_{A,B}+\beta_{A,B}\right)}{1-u^2} + \frac{1}{r_{\mathrm{A,\ B}}} \\ &= -1 + \frac{2\left(\alpha_{A,B}+\beta_{A,B}\right)}{1+u} + \frac{1}{r_{\mathrm{A,\ B}}}, \end{aligned} \tag{8.6}$$

The matrix elements, $\alpha_{A,B}$ and $\beta_{A,B}$, are three dimensional integrals that can be evaluated using elliptic coordinates. The result depends upon the distance between the nuclei. We obtain a better estimate of the ground state energy if we instead use hydrogen-like ion orbitals with a Z dependence. The integrals are evaluated in the Appendix, Sec. 8.7. There, the ground state energy is minimized with respect to Z to obtain the best ground state energy estimate with this type of orbital - in accord with the variational principle. The result,

$E_{\mathrm{H_2\ g.s.}}(r_{\mathrm{A,\ B}})$, is the potential energy for the vibrational nuclear motion. Minimizing this energy with respect to $r_{\mathrm{A,\ B}}$ gives the equilibrium bond distance.

In Sec. 8.7, the ground state energy estimate is shown to take the form,

$$E^{(0)}_{\mathrm{H_2\ g.s.}}(r) = \min_Z \left[-Z^2 + 2\frac{Z^2 - Z - r^{-1} + (1+Zr)\left(Z(Z-2)e^{-Zr} + r^{-1}e^{-2Zr}\right)}{1 + (1 + Zr + Z^2r^2/3)e^{-Zr}} + \frac{1}{r} \right].$$

Minimizing this energy, with respect to Z, requires numerical solution of $\partial E^{(0)}_{\mathrm{H_2\ g.s.}}/\partial Z = 0$. The optimal Z value varies with r. For large r, terms with exponential factors can be neglected;

$$\begin{aligned} E^{(0)}_{\mathrm{H_2\ g.s.}}(r) &= \min_Z \left[Z^2 - 2Z - \frac{1}{r} \right] \\ &= -1 - \frac{1}{r} \qquad\qquad Z = 1 \text{ is optimal} \\ &= 2E_{\mathrm{H\ g.s.}} - \frac{1}{r}. \end{aligned}$$

Adding back the electron repulsion - essentially $1/r$ - gives the energy of two hydrogen atoms. For small r, the exponential factors approach one, and terms with positive powers of r can be neglected. The ground state energy becomes

$$\begin{aligned} & E^{(0)}_{\mathrm{H_2\ g.s.}}(r) \\ =\ & \min_Z \left[-Z^2 + Z^2 - Z - r^{-1} + (1+Zr)\left(Z(Z-2)(1-Zr) + (1+Zr)r^{-1}(1-2Zr)\right) + \frac{1}{r} \right] \\ =\ & \min_Z \left[Z^2 - 4Z + \frac{1}{r} \right] \\ =\ & -4 + \frac{1}{r} \qquad\qquad Z = 2 \text{ is optimal} \\ =\ & E^{(0)}_{\mathrm{He\ g.s.}} \qquad\qquad \frac{1}{r} \text{ becomes part of nuclear energy.} \end{aligned}$$

Here, we get the energy of a helium atom with electron repulsion neglected.

Putting the elecMinimizing $E^{(0)}_{\mathrm{H_2\ g.s.}}(r)$ with respect to r gives the equilibrium bond distance of H_2^+ which has no electron repulsionThe ground state energy decreasesestimates - let $r_{\mathrm{B,\ A}} = r_{\mathrm{eq}} = 71.14$ pm (this must be converted to Bohr). Include electron repulsion at the level of first order perturbation theory. Estimate the *equilibrium dissociation energy*, D_{eq}, of H_2 by comparing your result with the energy of two isolated hydrogen atoms, in their ground states. The *bond dissociation energy*, D_0, is given by

$$D_0 = D_{\mathrm{eq}} - \frac{\hbar\omega}{2},$$

where $\hbar\omega/2$ is the zero point vibrational energy. The ground state of H_2 has zero point vibrational energy which makes the dissociation energy a little smaller than D_{eq}. The vibrational wavenumber of $^1H_2$ is 4400 cm^{-1}. Estimate the bond dissociation energy for H_2.

Solution 8.3. The ground state energy of H_2, without electron repulsion, is justThe energy of two isolated hydrogen atoms, in their ground states, is

$$2 \times \left(-\frac{1}{2}\right) = -1.$$

The *equilibrium dissociation* energy of H_2 the energy change for the process,

$$H_2(g) \rightarrow 2\ H(g);$$

i.e.,

$$\begin{aligned} D_{\mathrm{eq}}\,(\text{H-H}) &\cong -1 - \left(-1 - \frac{5}{3}\frac{1}{r_{\mathrm{eq}}}\right) \\ &= \frac{5}{3}\frac{1}{r_{\mathrm{eq}}}. \end{aligned}$$

Thus, in spite of the nuclear repulsion, the attraction of two electrons to the two nuclei makes bond formation favorable. The bond dissociation energy is positive - energy is required to break the bond. Here, r_{eq} is units of Bohr;

$$\begin{aligned} r_{\mathrm{eq}} &= 71.14\ \text{pm} \times 0.01889726\ \text{Bohr pm}^{-1} \\ &= 1.344\ \text{Bohr}. \end{aligned}$$

We get

$$D_{\mathrm{eq}}\,(\text{H-H}) \cong \frac{5}{3}\frac{1}{1.344} = 1.24\ \text{Hartree}.$$

To get the bond dissociation energy, we need the zero point energy,

$$\begin{aligned} \frac{\hbar\omega}{2} &= \frac{h\nu}{2} = \frac{hc\bar{\nu}}{2} \\ &= \frac{6.626 \times 10^{-34}\ \text{J s} \times 2.998 \times 10^{10}\ \text{cm s}^{-1} \times 4400\ \text{cm}^{-1}}{2} \\ &= 4.370 \times 10^{-20}\ \text{J} \\ &= 4.370 \times 10^{-20}\ \text{J} \times \left(4.359744 \times 10^{-18}\ \text{J Hartree}^{-1}\right)^{-1} \\ &= 0.01002\ \text{Hartree}. \end{aligned}$$

Therefore,

$$\begin{aligned} D_0\,(\text{H-H}) &= D_{\mathrm{eq}}\,(\text{H-H}) - \frac{\hbar\omega}{2} \\ &= 1.23\ \text{Hartree} \\ &= 5.36 \times 10^{-18}\ \text{J} \\ &= 3228\ \text{kJ mol}^{-1} \end{aligned}$$

Example 8.1. Consider H_2 in its ground state. If atom is centered at the origin, while the other is centered at $(x, y, z) = (0, 0, r_{\mathrm{eq}})$, the probability that an electron is between the two atoms is the expectation value of

$$\theta_{0<z<r_{\mathrm{eq}}}(x, y, z) = \left\{ \begin{array}{ll} 1, & 0 < z < r_{\mathrm{eq}} \\ 0, & \text{otherwise} \end{array} \right. ,$$

for the orbital the electron occupies. This probability can be evaluated for H_2 orbitals, with the following reasonable approximations: Neglect $(1\mathrm{s})_{\mathrm{A}}(\mathbf{x})$ for $z > r_{\mathrm{eq}}$, and $(1\mathrm{s})_{\mathrm{B}}(\mathbf{x})$ for $z < 0$. This means that all overlap between the orbitals comes from $0 < z < r_{\mathrm{eq}}$; i.e., from between the two atoms, as defined above.

(a) What is the probability an electron in $\sigma 1\mathrm{s}$ is between the two atoms? What is the enhancement of this value compared to the same probability for an electron in either $(1\mathrm{s})_{\mathrm{A}}$ or $(1\mathrm{s})_{\mathrm{B}}$?

(b) What is the probability an electron in $\sigma^* 1\mathrm{s}$ is between the two atoms? What is the depletion of this value compared to the same probability for an electron in either $(1\mathrm{s})_{\mathrm{A}}$ or $(1\mathrm{s})_{\mathrm{B}}$?

Solution 8.1.

(a) The expectation of $\theta_{0<z<r_{\mathrm{eq}}}(x, y, z)$ for an electron in $\sigma 1\mathrm{s}$ is given by

$$\begin{aligned} P &= \int_{0<z<r_{\mathrm{eq}}} \left|(\sigma 1\mathrm{s})(\mathbf{x})\right|^2 d\mathbf{x} \\ &= \frac{1}{2(1+u)} \int_{0<z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{A}}(\mathbf{x}) + (1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\right)^2 d\mathbf{x} \\ &= \frac{1}{2(1+u)} \left(\int_{0<z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{A}}(\mathbf{x})\right)^2 d\mathbf{x} + \int_{0<z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\right)^2 d\mathbf{x} \right. \\ &\quad \left. +2 \int_{0<z<r_{\mathrm{eq}}} (1\mathrm{s})_{\mathrm{A}}(\mathbf{x})(1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\, d\mathbf{x} \right) \\ &= \frac{1}{2(1+u)} \left(\underbrace{\int_{0<z} \left((1\mathrm{s})_{\mathrm{A}}(\mathbf{x})\right)^2 d\mathbf{x}}_{=1/2} + \underbrace{\int_{z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\right)^2 d\mathbf{x}}_{=1/2} \right. \\ &\quad \left. +2 \underbrace{\int (1\mathrm{s})_{\mathrm{A}}(\mathbf{x})(1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\, d\mathbf{x}}_{=u} \right) \\ &= \frac{1+2u}{2(1+u)} = \frac{1}{2} + \frac{u}{1+u}. \end{aligned}$$

The enhancement of this probability for the bonding orbital is $u/(1+u)$. The probability for an electron in either $(1\mathrm{s})_{\mathrm{A}}$ or $(1\mathrm{s})_{\mathrm{B}}$ is $1/2$.

(b) The expectation of $\theta_{0<z<r_{\mathrm{eq}}}(x, y, z)$ for an electron in $\sigma^{*}1\mathrm{s}$ is given by

$$\begin{aligned}
P &= \int_{0<z<r_{\mathrm{eq}}} \left|(\sigma^{*}1\mathrm{s})(\mathbf{x})\right|^{2} d\mathbf{x} \\
&= \frac{1}{2(1-u)} \int_{0<z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{A}}(\mathbf{x}) - (1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\right)^{2} d\mathbf{x} \\
&= \frac{1}{2(1-u)} \left(\underbrace{\int_{0<z} \left((1\mathrm{s})_{\mathrm{A}}(\mathbf{x})\right)^{2} d\mathbf{x}}_{=1/2} + \underbrace{\int_{z<r_{\mathrm{eq}}} \left((1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\right)^{2} d\mathbf{x}}_{=1/2} \right. \\
&\quad \left. -2 \underbrace{\int (1\mathrm{s})_{\mathrm{A}}(\mathbf{x})(1\mathrm{s})_{\mathrm{B}}(\mathbf{x})\, d\mathbf{x}}_{=u} \right) \\
&= \frac{1-2u}{2(1-u)} = \frac{1}{2} - \frac{u}{1-u}.
\end{aligned}$$

The depletion of this probability for the bonding orbital is $u/(1-u)$. Note that, since $u > 0$, electron density in the region between the atoms is depleted more in the antibonding orbital, than it is enhanced in the bonding orbital.

Example 8.2. Estimate the lowest electronic transition frequency for H_2 in terms α and β.

Solution 8.2. The first excited electronic state of H_2 corresponds to the electron configuration, $(\sigma 1\mathrm{s})^{1}(\sigma^{*}1\mathrm{s})^{1}$. The difference in energy between excited state and ground state is

$$\begin{aligned}
& E_{\mathrm{H}_2\ \mathrm{e.s.}} - E_{\mathrm{H}_2\ \mathrm{g.s.}} \\
&= \epsilon_1 + \epsilon_2 + \frac{1}{r_{\mathrm{A,\ B}}} - \left(2\epsilon_1 + \frac{1}{r_{\mathrm{A,\ B}}}\right) \\
&= \epsilon_2 - \epsilon_1 \\
&= \alpha - \beta - (\alpha + \beta) \\
&= -2\beta;
\end{aligned}$$

i.e., it is the spacing between bonding and antibonding orbital energies, -2β.

8.2 Diatomic molecules

More general diatomic molecules can be understood in terms of what we learned above about the hydrogen molecule. The first step is to consider the case where two inequivalent atomic orbitals are used. The simplest diatomic of this sort is HeH^{+}. This system is identical to H_2, except that that the two 1s orbitals are inequivalent. The analysis proceeds as above with the one electron Hamiltonian matrix,

$$\mathsf{H} = \begin{pmatrix} \alpha_1 & \beta \\ \beta & \alpha_2 \end{pmatrix},$$

where α_1 and α_2 are effectively the energies of an electron in the He 1s and H 1s orbitals, respectively.

The determinantal equation for the eigenvalues of H is

$$\begin{aligned} 0 &= \begin{vmatrix} \alpha_1 - \epsilon & \beta \\ \beta & \alpha_2 - \epsilon \end{vmatrix} \\ &= (\alpha_1 - \epsilon)(\alpha_2 - \epsilon) - \beta^2 \\ &= \epsilon^2 - (\alpha_1 + \alpha_2)\,\epsilon + \alpha_1\alpha_2 - \beta^2. \end{aligned}$$

The *molecular orbitial energies* - the eigenvalues of H - are given by

$$\begin{aligned} \epsilon &= \frac{\alpha_1 + \alpha_2 \pm \left((\alpha_1 + \alpha_2)^2 - 4\left(\alpha_1\alpha_2 - \beta^2\right)\right)^{1/2}}{2} \\ &= \frac{\alpha_1 + \alpha_2 \pm \left(\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2 - 4\alpha_1\alpha_2 + 4\beta^2\right)^{1/2}}{2} \\ &= \frac{\alpha_1 + \alpha_2 \pm \left((\alpha_1 - \alpha_2)^2 + 4\beta^2\right)^{1/2}}{2} \\ &= \frac{\alpha_1 + \alpha_2}{2} \pm \left(\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2 + \beta^2\right)^{1/2} \end{aligned}$$

We consider two extremes. If the energy separation, $|\alpha_1 - \alpha_2|$, is much smaller than the coupling, $|\beta|$, then

$$\begin{aligned} \left(\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2 + \beta^2\right)^{1/2} &= \beta\left(1 + \frac{1}{\beta^2}\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2\right)^{1/2} \\ &\cong \beta\left(1 + \frac{1}{2\beta^2}\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2\right) \qquad \alpha_1 - \alpha_2 \ll \beta \\ &= \beta + \frac{1}{2\beta}\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2. \end{aligned}$$

In this case,

$$\begin{aligned} \epsilon &= \frac{\alpha_1 + \alpha_2}{2} \pm \left(\beta + \frac{1}{2\beta}\left(\frac{\alpha_1 - \alpha_2}{2}\right)^2\right) \\ &= \frac{\alpha_1 + \alpha_2}{2} \pm \beta \pm \frac{1}{2\beta}\left(\frac{\alpha_2 - \alpha_1}{2}\right)^2. \end{aligned}$$

The lowest energy eigenvalue is

$$\begin{aligned}
\epsilon_1 &= \frac{\alpha_1+\alpha_2}{2}+\beta+\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)^2 \\
&= \alpha_1+\beta+\frac{\alpha_2-\alpha_1}{2}+\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)^2 \\
&= \alpha_1+\beta+\left(\frac{\alpha_2-\alpha_1}{2}\right)\left(1+\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)\right), \qquad (8.7)
\end{aligned}$$

where $\alpha_1 < \alpha_2$. Similarly, the highest energy energy eigenvalue is

$$\begin{aligned}
\epsilon_2 &= \frac{\alpha_1+\alpha_2}{2}-\beta-\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)^2 \\
&= \alpha_2-\beta-\frac{\alpha_2-\alpha_1}{2}-\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)^2 \\
&= \alpha_2-\beta-\left(\frac{\alpha_2-\alpha_1}{2}\right)\left(1+\frac{1}{2\beta}\left(\frac{\alpha_2-\alpha_1}{2}\right)\right). \qquad (8.8)
\end{aligned}$$

These results are similar to the case of equivalent orbitals. $\epsilon_1 = \alpha_1 + \beta$, with a small positive correction less than $(\alpha_2 - \alpha_1)/2$. Remember that $\beta < 0$. $\epsilon_2 = \alpha_2 - \beta$, with a small negative correction with magnitude less than $(\alpha_2 - \alpha_1)/2$.

The above results are interpreted as follows. The lower atomic orbital energy - here, the first - is lowered in energy by an amount less than $|\beta|$ to give the energy of the bonding molecular orbital. The higher atomic orbital energy is raised by an amount less than $|\beta|$ to give the energy of the antibonding molecular orbital. The energy mismatch reduces the splitting energy of the orbitals. $2|\beta|$ is the maximum splitting, which is seen in case of equivalent orbitals.

Now consider the case of $|\alpha_1 - \alpha_2| \gg |\beta|$. The eigenvalues are given by

$$\begin{aligned}
\epsilon &= \frac{\alpha_1+\alpha_2}{2}\pm\left(\left(\frac{\alpha_2-\alpha_1}{2}\right)^2+\beta^2\right)^{1/2} \\
&= \frac{\alpha_1+\alpha_2}{2}\pm\left(\frac{\alpha_2-\alpha_1}{2}\right)\left(1+\left(\frac{2\beta}{\alpha_2-\alpha_1}\right)^2\right)^{1/2} \\
&\cong \frac{\alpha_1+\alpha_2}{2}\pm\left(\frac{\alpha_2-\alpha_1}{2}\right)\left(1+\frac{1}{2}\left(\frac{2\beta}{\alpha_2-\alpha_1}\right)^2\right) \\
&= \frac{\alpha_1+\alpha_2}{2}\pm\left(\frac{\alpha_2-\alpha_1}{2}\right)\pm\frac{1}{2}\frac{2\beta^2}{\alpha_2-\alpha_1} \\
&= \left\{\begin{array}{l}\alpha_2+\frac{\beta^2}{\alpha_2-\alpha_1} \\ \alpha_1-\frac{\beta^2}{\alpha_2-\alpha_1}\end{array}\right. .
\end{aligned}$$

In this case, the eigenvalues are only slightly shifted by the bonding. The lower orbital energy is shifted lower, and the higher orbital energy is shifted higher.

The shifts are smaller than $|\beta|$ by the factor, $|\beta| / (\alpha_2 - \alpha_1) \ll 1$. Specifically,

$$\epsilon_1 = \alpha_1 - \frac{\beta^2}{\alpha_2 - \alpha_1} \tag{8.9}$$

and

$$\epsilon_2 = \alpha_2 + \frac{\beta^2}{\alpha_2 - \alpha_1}. \tag{8.10}$$

The overall conclusion is that bonding is strongest when the energies of the atomic orbitals are well matched. Atomic orbitals well separated in energy show little stabilization due to bonding.

In the case of the ground state of HeH^+, the atomic orbitals are not well matched in energy. The nuclear charge on He is twice that of H. Also, since these are 1s orbitals, the lowest energy orbitals, energy differences are greatest. The H 1s orbital energy matches better with second shell orbital energies in period two atoms. Because of the large energy mismatch in HeH^+, the bond is very weak. HeH^+ is beyond a superacid - it could only exist not exposed to anything else.

The above sets the frame work for bonding between atoms with many electrons. Atomic orbitals on the two atoms form molecular orbitals, when there is good coupling - $|\beta|$ is not too small, or zero due to symmetry - and the atomic orbital energies are matched. We consider elements from the second period. The inner shell 1s orbitals on the two atoms are very low in energy - with greater mismatches between elements. These orbitals have little overlap - and associated coupling - because each is close to its own nucleus, and the distance between nuclei must accommodate the valence shells of the two atoms. To the extent that these orbitals split, the bonding and antibonding molecular orbitals formed are very low in energy, and are both filled in the ground state electron configuration of the molecule. There is no net bonding because the stabilization of the bonding electrons is canceled by the destabilization of the anitbonding electrons. Thus, the inner shell orbitals do not form bonds. Bonding is the sharing of valence shell electrons, as only valence shell orbitals have good overlap. This is a cornerstone of chemistry.

The next step is to group atomic orbitals according to angular momentum about the molecular axis. For second period elements, this means the 2s and 2p orbitals. Orbitals 2s and $2p_z$ have zero angular momentum about the z axis, and form σ bonds. Orbitals $2p_x$ and $2p_y$ have $m_\ell = \pm 1$, and form π bonds. These groups of atomic orbitals can be considered separately. Coupling between σ and π-type atomic orbitals is zero due to symmetry. The σ-type orbitals do not vary with angle about the molecular axis, whereas π-type orbitals have one node. Since $\hat{H}$ commutes with $\hat{L}_z$, applying $\hat{H}$ to an orbital does not change its symmetry. The inner product of σ and π-type wavefunctions is the integral of a function with π symmetry - it integrates to zero. Consequently, H matrix elements between such orbitals are zero.

8.2.1 The σ system

Since the 2s and $2\mathrm{p}_z$ orbitals both have σ symmetry, there are four atomic orbitals that can form σ bonds. However, there is limited room between the atoms, and the these orbitals have density on the wrong side of each atom, as well as between them. The σ system is constructed in two stages. First, consider the two atoms approaching each other - but, not yet bonded. Each atom acts as a perturbation on the other atom. Because the perturbation is localized to one side of the atom, the atom no longer has spherical symmetry. The perturbation couples 2s and $2\mathrm{p}_z$. Since the two orbitals are on the same atom, they are already orthogonal; $\mathsf{U} = \mathsf{1}$ and $\mathsf{U}^{-1/2} = \mathsf{1}$. The eigenvectors of the 2×2 H matrix, associated with 2s and $2\mathrm{p}_z$, provide new atomic orbitals - *hybrid orbitals* - that account for the presence of the approaching other atom. When the other atom is far, the coupling is small and the the hybrid orbitals are much like the 2s and $2\mathrm{p}_z$. The lower energy hybrid orbital is a superposition of mostly 2s, with a little of $+2\mathrm{p}_z$.

$$(2\mathrm{sp}_c) = \frac{1}{\sqrt{1+c^2}}\left((2\mathrm{s}) + c(2\mathrm{p}_z)\right)$$

The constructive interference, on the side of the incoming atom, polarizes 2s so that it is shifted toward the incoming atom. This enhances the lowering of energy due to the incoming atom. The higher energy hybrid orbital is mostly $2\mathrm{p}_z$, with a little of $-2\mathrm{s}$. This orbital polarizes in the direction opposite to that of the incoming atom, and increases in energy.

When the incoming atom approaches more closely, a covalent bond forms - if the atoms do not already have closed shells. The two $2\mathrm{sp}_c$ orbitals, which face each other, overlap and have a negative (i.e., favorable) energy coupling. These orbitals split into bonding and antibonding orbitals. The remaining twist is to note that the energy is optimized when the overlap between the two $2\mathrm{sp}_c$ orbitals is optimized. We determine the value of c that produces the greatest polarization of the $2\mathrm{sp}_c$ orbital.

The probability that an electron in the $2\mathrm{sp}_c$ orbital has positive z coordinate is given by

$$\begin{aligned}
P_{z>0} &= \int_{\mathbb{R}^{3N}} (2\mathrm{sp}_c)(\mathbf{x})\,\theta_{z>0}(\mathbf{x})\,(2\mathrm{sp}_c)(\mathbf{x})\,d\mathbf{x} \\
&= \langle (2\mathrm{sp}_c)\,|\theta_{z>0}(2\mathrm{sp}_c)\rangle \\
&= \frac{1}{1+c^2}\langle (2\mathrm{s}) + c(2\mathrm{p}_z)\,|\theta_{z>0}\left((2\mathrm{s}) + c(2\mathrm{p}_z)\right)\rangle \\
&= \frac{1}{1+c^2}\left(\langle (2\mathrm{s})\,|\theta_{z>0}(2\mathrm{s})\rangle + c^2\langle (2\mathrm{p}_z)\,|\theta_{z>0}(2\mathrm{p}_z)\rangle + 2c\langle (2\mathrm{s})\,|\theta_{z>0}(2\mathrm{p}_z)\rangle\right) \\
&= \frac{1}{1+c^2}\left(\frac{1}{2} + c^2\frac{1}{2} + 2c\eta\right) \\
&= \frac{1}{2} + \frac{2c\eta}{1+c^2}. \qquad (8.11)
\end{aligned}$$

Due to symmetry of the 2s orbital under reflection through the the x y plane (i.e., $z \to -z$), $\langle (2\text{s}) \, | \theta_{z<0}(2\text{s}) \rangle = \langle (2\text{s}) \, | \theta_{z>0}(2\text{s}) \rangle$. Since the sum of these two probabilities is 1, they must both equal 1/2. This was used in the second last line above for the 2s and 2p_z orbitals. The probability, $P_{z>0}$, is a maximum when

$$\begin{aligned} \frac{d}{dc}\frac{2c\eta}{1+c^2} &= 0 \\ \frac{2\eta\left(1+c^2\right) - 4c^2\eta}{\left(1+c^2\right)^2} &= 0 \end{aligned}$$

or

$$\begin{aligned} 2\eta\left(1+c^2\right) - 4c^2\eta &= 0 \\ 1 + c^2 - 2c^2 &= 0 \\ c^2 &= 1 \\ c &= \pm 1. \end{aligned}$$

$c = +1$ gives the maximum polarization. $c = -1$ gives the maximum polarization in the opposite direction - $P_{z>0}$ is a minimum. These are the *2sp hybrid orbitals*, the hybridization adopted by bonding atoms. The 2sp orbitals facing each other, on the two bonding atoms, form bonding and antibonding molecular orbitals. The magnitude of the coupling is largest for sp atomic orbitals - larger than unhybridized s and p orbitals. The two 2sp orbitals that face away from the other atom have little coupling. They are *non-bonding orbitals* on their respective atoms.

In summary, the σ system consists of bonding and anitbonding orbitals, σ2sp and σ^*2sp, and the two non-bonding orbitals, both n2sp. For homonuclear diatomics, the orbital energies of the two bonding 2sp orbitals are the same, and the orbital splitting is greatest. Energy mismatch reduces stabilization of the bonding orbital, and destabilization of the antibonding orbital.

8.2.2 The π system

The π system consists of the two 2p_x and two 2p_y orbitals on the bonding atoms - four orbitals in total. However, these orbitals come in two pairs. The 2p_x orbitals are anitsymmetric with to reflection through the y z plane ($x \to -x$). The 2p_y orbitals are symmetric with to this reflection. Since the Hamiltonian commutes with the associated reflection operator, it does not couple symmetric and antisymmetric states. The 2p_x and 2p_y orbitals form two independent π systems, orthogonal to each other. Each π system has two atomic orbitals that split into a bonding and antibonding pair. π bonding is weaker than σ bonding because the atomic 2p_x and 2p_y orbitals do not overlap as much - along the z direction - and have weaker couplings.

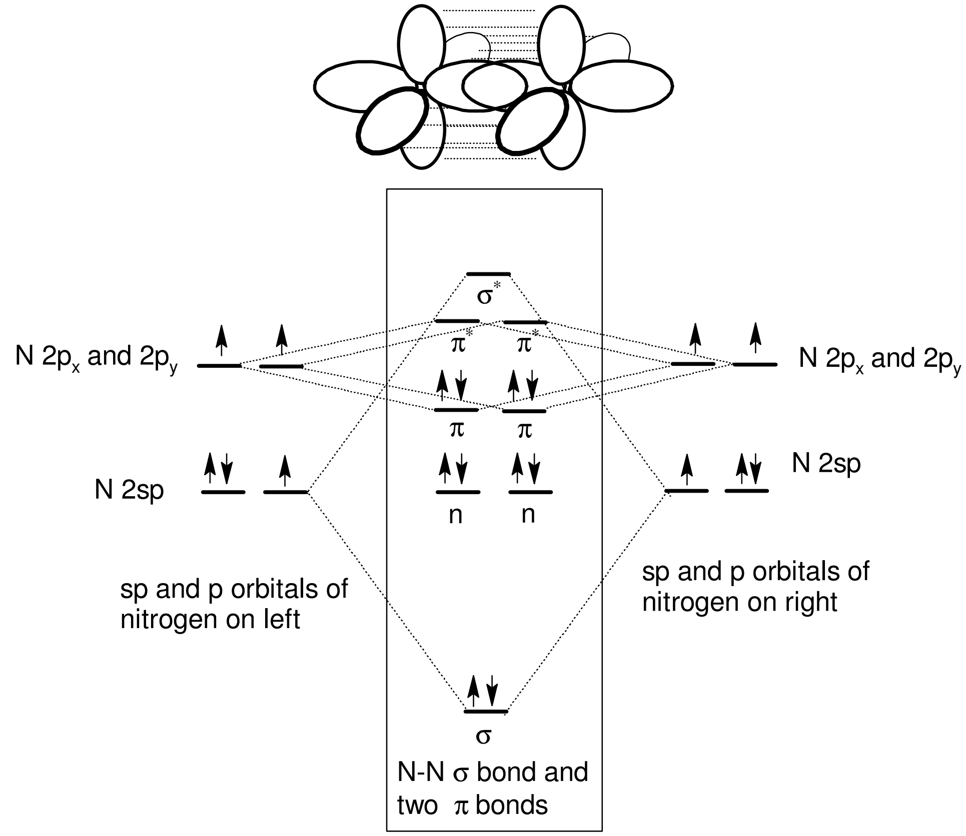


Figure 8.1: A schematic molecular orbital diagram for the ground state of N_2. There are non-bonding electron pairs on both atoms, and three bonding pairs - one σ and two π. Nitrogen has a triple bond.

8.2.3 Molecular orbital diagrams

Approximate diatomic molecule ground states can be constructed by filling the bonding, non-bonding and antibonding orbitals described above, with lower energy orbitals filled first. Figure 8.1 shows a schematic molecular orbital diagram for the ground state of N_2. The nitrogen molecule has ten valence electrons - five from each nitrogen atom. Four of these electrons remain in non-bonding - two electrons in each of the non-bonding 2sp orbitals. The other six electrons occupy the σ and two π orbitals, in pairs. The nitrogen molecule has a triple bond.

In the case of F_2 there are fourteen valence electrons. In this case, the two π^* orbitals are doubly occupied. The destabilization of these π^* electrons cancels the stabilization of the π electrons, and there are no net π bonds. Since

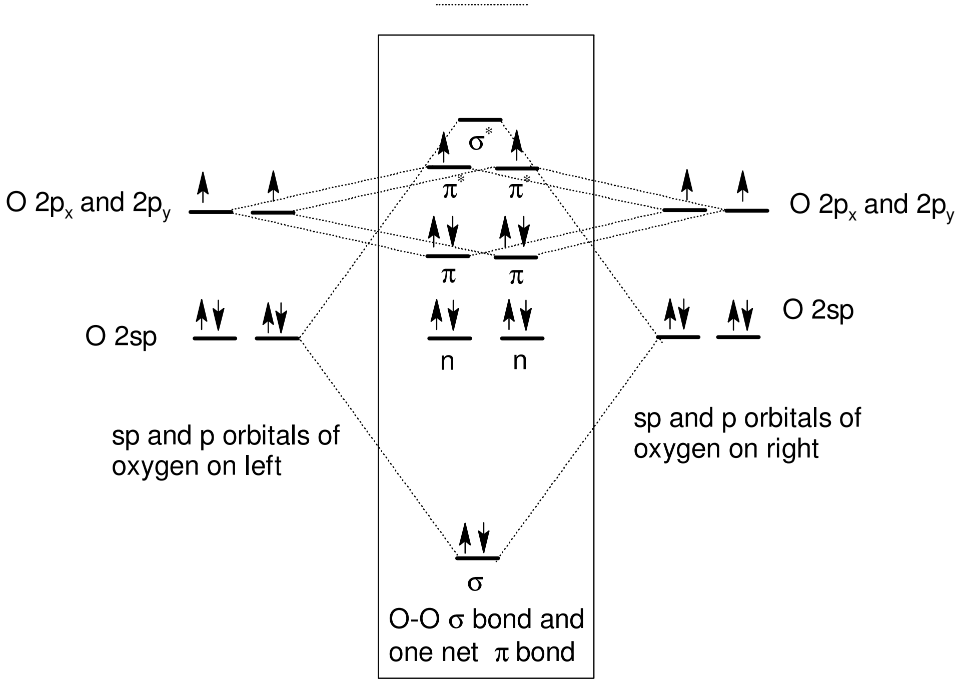


Figure 8.2: A schematic molecular orbital diagram for the ground state of oxygen. The ground state is the triplet state with the unpaired electron spins in the degenerate π^* orbitals.

the higher energy σ^* orbital is unoccupied, there is a net σ bond. The fluorine molecule has a single bond. Oxygen molecules have twelve valence electrons. In this case, the two π^* orbitals are singly occupied. In accord with Hund's rule 1 for atomic ground states, the ground state of the oxygen molecule is the triplet state with the two π^* electrons in the same spin state. Figure 8.2 shows the molecular orbital diagram for O_2.

The singlet state of O_2 is its first electronic excited state. Singlet O_2 is very reactive. *Photodynamic therapy* (PDT) uses the targeted production of singlet oxygen in cancer tissue to kill cancer cells. In the first step, a *photosensitizer* is administered to the patient. The photosensitizer is a nontoxic chromophore which can be targeted to cancer tissues. It absorbs light to produce an excited state that transfers energy to neighboring triplet oxygen (present in all living tissues) to produce singlet oxygen. The singlet oxygen reacts rapidly with neighboring molecules, damaging chromosomes and crippling the life processes of cells. The light is introduced externally for near the surface tissues (e.g., skin cancer, or acne - PDT also kills bacteria), or through narrow fiber optic cables.

In the case of heteronuclear diatomics, the σ-σ^* and π-π^* splittings are not as large - because of orbital energy mismatch. This is depicted in Figure 8.3. Nevertheless, the bond dissociation energy of CO (1076 kJ mol^{-1}) is higher than that of N_2 (945 kJ mol^{-1}). However, there is no inconsistency. If we tally the electron energy differences for

$$\mathrm{N_2(g) \rightarrow N(g) + N(g)}$$

and

$$\mathrm{CO(g) \rightarrow C(g) + O(g)},$$

we get

$$\Delta E\,(\mathrm{N_2}) = -4\beta_\pi\,(\mathrm{N_2}) - 2\beta_\sigma\,(\mathrm{N_2})$$

for N_2 and

$$\Delta E\,(\mathrm{CO}) = -4\beta_\pi\,(\mathrm{CO}) - 2\beta_\sigma\,(\mathrm{CO}) + 2\left(\alpha_{\mathrm{2p}}\,(\mathrm{C}) - \alpha_{\mathrm{2p}}\,(\mathrm{O})\right).$$

Here, we assume $|\beta_\pi\,(\mathrm{CO})|$ is smaller than $\alpha_{\mathrm{2p}}\,(\mathrm{C}) - \alpha_{\mathrm{2p}}\,(\mathrm{O})$, so that the π bond stabilization is still $2\beta_\pi\,(\mathrm{CO})$ per π orbital. The extra stabilization of the carbon $2\mathrm{p}_x$ and $2\mathrm{p}_y$ electrons can compensate for the smaller values of $|\beta_\pi\,(\mathrm{CO})|$ and $|\beta_\sigma\,(\mathrm{CO})|$, the respective couplings of the π and σ orbitals. In general, more accurate quantum chemistry calculations are required to reliably account for bond dissociation energies.

Excited state electron configurations can be generated from the above molecular orbital diagrams by considering specific excitations of one (or more) electrons. Except in cases of multiple terms associated with the ground state electron configuration (e.g., oxygen triplet and singlet), the first excited state is associated with the *HOMO* to *LUMO* transition of a single electron. HOMO is the highest occupied molecular orbital, and LUMO is the lowest unoccupied molecular orbital. In the case of nitrogen, the HOMO to LUMO transition is a π to π^* transition. Measuring the associated transition energy, via spectroscopy, provides a means of determining β_π - the transition energy is estimated to be $2\,|\beta_\pi|$.

8.3 Polyatomic molecules

Electron configurations of polyatomic molecules can be constructed by simple extensions of the diatomic case. First, consider two atoms approaching a single atom, A. As before, the atom polarizes. However, the polarization depends on the angle between the lines of approach of the two atoms. The optimal angle between these lines depends on the number of electrons in the valence shell of A. The atom must accommodate non-bonding electrons, and the atoms approach. This gives rise to the valence shell electron pair repulsion theory (VSEPR) rules for predicting approximate molecular geometries. Applying VSEPR to each atom determines a geometry for the whole molecule - subject to steric constraints, and other non-bonding interactions (e.g., hydrogen bonding).

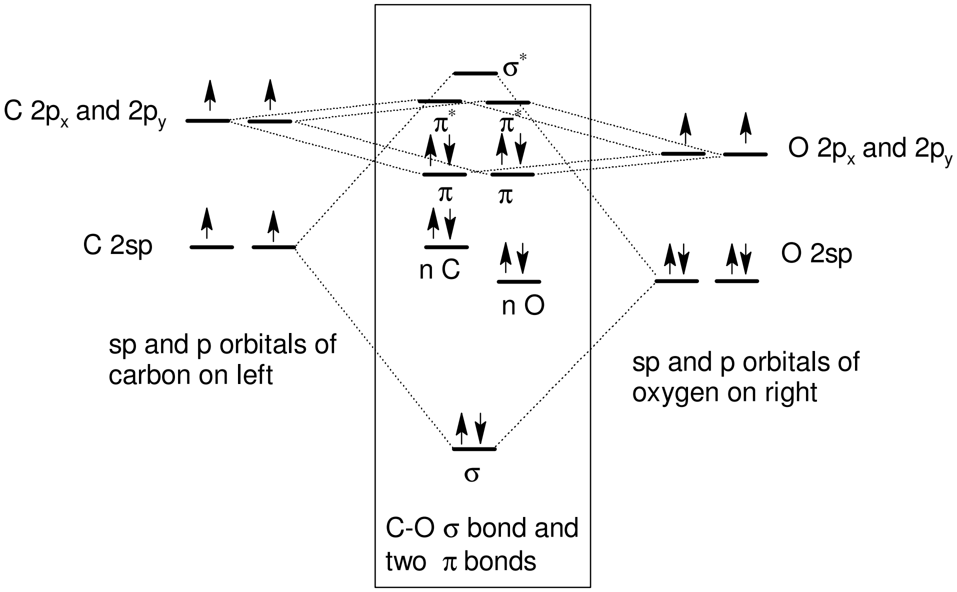


Figure 8.3: A schematic molecular orbital diagram for CO. The atomic orbital energies in oxygen are lower than those of carbon because oxygen has a larger effective nuclear charge.

VSEPR theory predicts linear, trigonal planar and tetrahedral geometries about second period atoms. In the case of the linear geometry, the optimal angle between the lines of approach of the two approaching atoms is 180°. Atom A forms two sp orbitals that form σ bonds with orbitals from the two approaching atoms. Additional π bonds are formed with one or both of the approaching atoms, depending on the occupation of the π and π^* orbitals.

When the geometry is trigonal planar, the optimal hybridization for σ bonding is sp^2. The three sp^2 orbitals are given by

$$(\mathrm{sp}^2) = \begin{cases} \frac{1}{\sqrt{3}}\left((\mathrm{s}) + \sqrt{2}(\mathrm{p}_z)\right) \\ \frac{1}{\sqrt{3}}\left((\mathrm{s}) + \sqrt{\frac{3}{2}}(\mathrm{p}_y) - \sqrt{\frac{1}{2}}(\mathrm{p}_z)\right) \\ \frac{1}{\sqrt{3}}\left((\mathrm{s}) - \sqrt{\frac{3}{2}}(\mathrm{p}_y) - \sqrt{\frac{1}{2}}(\mathrm{p}_z)\right) \end{cases} \tag{8.12}$$

They are orthogonal. Each one has 1/3 s character and 2/3 p character. The associated probability densities are concentrated in directions pointing to the vertices of an equilateral triangle in the y z plane - with one vertex along the $+z$ axis. Rotating the coordinate system, about the x axis, by 120° maps each orbital to the one below - with the last mapping to the first. These orbitals are optimal for forming two or three bonds σ bonds, with the case of two bonds arising when the third sp^2 orbital is occupied by non-bonding electrons. There is one remaining p orbital which can produce a π bond with one of the σ-bonded atoms - depending on the occupation of π and π^*.

When the geometry is tetrahedral the hybridization is sp^3. The sp^3 orbitals are given by

$$(\mathrm{sp}^3) = \begin{cases} \frac{1}{2}\left((\mathrm{s}) + \sqrt{3}(\mathrm{p}_z)\right) \\ \frac{1}{2}\left((\mathrm{s}) + \sqrt{\frac{8}{3}}(\mathrm{p}_y) - \sqrt{\frac{1}{3}}(\mathrm{p}_z)\right) \\ \frac{1}{2}\left((\mathrm{s}) + \sqrt{2}(\mathrm{p}_x) - \sqrt{\frac{2}{3}}(\mathrm{p}_y) - \sqrt{\frac{1}{3}}(\mathrm{p}_z)\right) \\ \frac{1}{2}\left((\mathrm{s}) - \sqrt{2}(\mathrm{p}_x) - \sqrt{\frac{2}{3}}(\mathrm{p}_y) - \sqrt{\frac{1}{3}}(\mathrm{p}_z)\right) \end{cases} . \tag{8.13}$$

Each one of these has 1/4 s character and 3/4 p character. These orbitals are optimal for forming two, three or four bonds - with two, one or zero non-bonding electron pairs, respectively.

The trigonal bipyramidal and octahedral geometries associated with $\mathrm{sp}^3\mathrm{d}$ and $\mathrm{sp}^3\mathrm{d}^2$ hydrid orbitals.

8.4 Hartree-Fock theory

So far, we have neglected the electron repulsion term in the Hamiltonian, except to say that we could include its effect via perturbation theory. Here, we refer back to Eq. 8.2 which included the electron repulsion term for the ground state of H_2, and consider the more general case of a molecular (or atomic) system with n electrons. *Hartee-Fock theory* provides the best orbital-based description of a molecular or atomic system. In the most general formulation, one begins with the antisymmetrizer for n electrons acting on a product of spin-orbitals - a spin-orbital is a product of orbital and spin functions. If spin-orbit coupling is negligible, the many electron wavefunction factors into the product of orbital and spin wavefunctions. In the case of closed shell systems, where all orbitals are doubly occupied, *restricted Hartree-Fock* arises. In this case, spin up and spin down electrons occupy the same orbitals. In the open shell case, there are unpaired electrons, and the spin up orbitals will be slightly different from the spin down orbitals - due to different antisymmetrizations for the different spin state orbitals.

In the case of closed shells, in the independent electron approximation, the orbital part of the many electron wavefunction takes the form,

$$\psi\left(1,2,\ldots,n\right)=\hat{A}_{\alpha}\hat{A}_{\beta}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right),$$

where $\hat{A}_{\alpha}$ is the antisymmetrizer for the spin up electrons - those with odd index - and $\hat{A}_{\beta}$ is the anitsymmetrizer for the spin down electrons - those with even index. The $n/2$ molecular orbitals are orthogonal and normalized one electron wavefunctions.

The Hamiltonian for the n electron system is

$$\begin{aligned}\hat{H} &= -\frac{1}{2}\sum_{j=1}^{n}\left(\nabla^{2}-\sum_{a=1}^{N}\frac{Z_{a}}{r_{j,a}}\right)+\sum_{j<k}\frac{1}{r_{j,k}}+\sum_{a<b}\frac{Z_{a}Z_{b}}{r_{a,b}}\\ &= \sum_{j=1}^{n}\hat{H}_{j}+\sum_{j<k}\frac{1}{r_{j,k}}+\sum_{a<b}\frac{Z_{a}Z_{b}}{r_{a,b}}\end{aligned}$$

where a and b label the N nuclei in the system, while j and k label the electrons. The expectation value of the Hamiltonian for the above wavefunction is

$$\begin{aligned}&\left\langle\psi\left|\hat{H}\psi\right.\right\rangle\\ =\;&\left\langle\hat{A}_{\alpha}\hat{A}_{\beta}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\left|\hat{H}\hat{A}_{\alpha}\hat{A}_{\beta}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\right.\right\rangle\\ =\;&\underbrace{\left\langle\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\left|\hat{A}_{\beta}\hat{A}_{\alpha}\hat{H}\hat{A}_{\alpha}\hat{A}_{\beta}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\right.\right\rangle}_{\text{antisymmetrizers are Hermitian}}\\ =\;&\underbrace{\left\langle\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\left|\hat{H}\hat{A}_{\alpha}^{2}\hat{A}_{\beta}^{2}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\right.\right\rangle}_{\text{antisymmetrizers commute with Hamiltonian}}\\ =\;&\left(\frac{n}{2}\right)!\left\langle\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\left|\hat{H}\hat{A}_{\alpha}\hat{A}_{\beta}\psi_{1}\left(1\right)\psi_{1}\left(2\right)\ldots\psi_{n/2}\left(n\right)\right.\right\rangle.\end{aligned}$$

In the last line, we used the formula

$$\hat{A}^{2}=\sqrt{n!}\hat{A}, \tag{8.14}$$

where

$$\hat{A}=\frac{1}{\sqrt{n!}}\sum_{\hat{P}}\left(-1\right)^{p\left(\hat{P}\right)}\hat{P}$$

is the antisymmetrizer for n electrons. The antisymmetrizers above are for $n/2$ electrons.

To prove Eq. 8.14, note that

$$\begin{aligned} \hat{A}^2 &= \frac{1}{n!}\sum_{\hat{P}}(-1)^{p\left(\hat{P}\right)}\hat{P}\sum_{\hat{P}'}(-1)^{p\left(\hat{P}'\right)}\hat{P}' \\ &= \frac{1}{n!}\sum_{\hat{P}}\sum_{\hat{P}'}(-1)^{p\left(\hat{P}\right)+p\left(\hat{P}'\right)}\hat{P}\hat{P}' \end{aligned}$$

The product of any two permutations is another permutation. The parity of a product of permutations is the sum of the parities. Therefore,

$$\begin{aligned} \hat{A}^2 &= \frac{1}{n!}\sum_{\hat{P}}\sum_{\hat{P}'}(-1)^{p\left(\hat{P}\hat{P}'\right)}\hat{P}\hat{P}' \\ &= \frac{1}{n!}\sum_{\hat{P}}\sum_{\hat{P}'}(-1)^{p\left(\hat{P}'\right)}\hat{P}' \\ &= \frac{n!}{n!}\sum_{\hat{P}'}(-1)^{p\left(\hat{P}'\right)}\hat{P}' \\ &= \sqrt{n!}\hat{A}. \end{aligned}$$

In the second line, we note that the sum over all $\hat{P}'$ can be written as the sum over all $\hat{P}\hat{P}'$, since each distinct permutation is counted only once in the sum over $\hat{P}\hat{P}'$. The product, $\hat{P}\hat{P}'$, can then be relabeled as $\hat{P}'$. The sum over $\hat{P}$ accounts for the $n!$ factor which cancels the prefactor.

The expectation of the Hamiltonian is expressed as the sum of three terms - the expectation value of the three terms in the Hamiltonian.

$$\begin{aligned} & \left\langle\psi\left|\hat{H}\psi\right.\right\rangle \\ = \; & \left(\frac{n}{2}\right)!\left\langle\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right| \qquad (8.15) \\ & \left|\left(\sum_{j=1}^{n}\hat{H}_j+\sum_{j<k}\frac{1}{r_{j,k}}+\sum_{a<b}\frac{Z_aZ_b}{r_{a,b}}\right)\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \end{aligned}$$

The third term is straightforward. The repulsions between nuclear appear here as a constant term. The associated matrix element is just a matrix element of the product of antisymmetrizers. Since any permutation of the n electron wavefunction produces an orthogonal n electron wavefunction, only the $\hat{P}_\alpha = \hat{P}_\beta = 1$ term produces a non-zero inner product. For example, consider a simple exchange operator of electrons 1 and 3. This yields the inner product,

$$\begin{aligned} & \left\langle\psi_1(1)\,\psi_1(2)\,\psi_2(3)\ldots\psi_{n/2}(n)\left|\psi_1(3)\,\psi_1(2)\,\psi_2(1)\ldots\psi_{n/2}(n)\right.\right\rangle \\ = \; & \underset{=0}{\langle\psi_1(1)\,|\psi_2(1)\rangle}\underset{=1}{\langle\psi_1(2)\,|\psi_1(2)\rangle}\underset{=0}{\langle\psi_2(3)\,|\psi_1(3)\rangle}\ldots\underset{=1}{\left\langle\psi_{n/2}(n)\left|\psi_{n/2}(2)\right.\right\rangle} \\ = \; & 0. \end{aligned}$$

Altogether,

$$
\begin{aligned}
&\left(\frac{n}{2}\right)! \left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n) \middle| \sum_{a<b}\frac{Z_aZ_b}{r_{a,b}}\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \\
=\;& \sum_{a<b}\frac{Z_aZ_b}{r_{a,b}}\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n) \middle| \sum_{\hat{P}_\alpha}(-1)^{p\left(\hat{P}_\alpha\right)}\right. \\
&\left.\hat{P}_\alpha\sum_{\hat{P}_\beta}(-1)^{p\left(\hat{P}_\beta\right)}\hat{P}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \\
=\;& \sum_{a<b}\frac{Z_aZ_b}{r_{a,b}}.
\end{aligned}
$$

The first term in Eq. 8.15 is

$$
\begin{aligned}
&\left(\frac{n}{2}\right)! \left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n) \middle| \sum_{j=1}^{n}\hat{H}_j\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \\
=\;& \sum_{j=1}^{n}\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right| \\
&\left|\hat{H}_j\sum_{\hat{P}_\alpha}(-1)^{p\left(\hat{P}_\alpha\right)}\hat{P}_\alpha\sum_{\hat{P}_\beta}(-1)^{p\left(\hat{P}_\beta\right)}\hat{P}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \\
=\;& \sum_{j=1}^{n}\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right| \\
&\left|\sum_{\hat{P}_\alpha}(-1)^{p\left(\hat{P}_\alpha\right)}\hat{P}_\alpha\sum_{\hat{P}_\beta}(-1)^{p\left(\hat{P}_\beta\right)}\hat{P}_\beta\psi_1(1)\,\psi_1(2)\ldots\hat{H}_j\psi_{\lceil\frac{j}{2}\rceil}(j)\ldots\psi_{n/2}(n)\right\rangle
\end{aligned}
$$

This is a sum of inner products between a product state, $\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)$, and a permuted product state with the j the factor replaced by $\hat{H}_j\psi_{\lceil\frac{j}{2}\rceil}(j)$. This is the only factor that is not orthogonal to the other factors. Thus, any permutation of the product state on the right of the inner product produces a zero inner product, and this one electron term becomes

$$\begin{aligned}
& \sum_{j=1}^{n} \left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n) \,\middle|\, \psi_1(1)\,\psi_1(2)\ldots\hat{H}_j\psi_{\lceil \frac{j}{2}\rceil}(j)\ldots\psi_{n/2}(n)\right\rangle \\
= \; & \sum_{j=1}^{n} \langle\psi_1(1)\underset{=1}{|}\psi_1(1)\rangle\langle\psi_1(2)\underset{=1}{|}\psi_1(2)\rangle\ldots \\
& \underset{=\left\langle \psi_{\lceil \frac{j}{2}\rceil}(1)\,\middle|\,\hat{H}_1\psi_{\lceil \frac{j}{2}\rceil}(1)\right\rangle}{\left\langle \psi_{\lceil \frac{j}{2}\rceil}(j)\,\middle|\,\hat{H}_j\psi_{\lceil \frac{j}{2}\rceil}(j)\right\rangle}\ldots\left\langle\psi_{n/2}(n)\underset{=1}{\middle|}\psi_{n/2}(n)\right\rangle \\
= \; & \sum_{j=1}^{n} \left\langle \psi_{\lceil \frac{j}{2}\rceil}(1)\,\middle|\,\hat{H}_1\psi_{\lceil \frac{j}{2}\rceil}(1)\right\rangle \\
= \; & 2\sum_{m=1}^{n/2} \left\langle\psi_m(1)\,\middle|\,\hat{H}_1\psi_m(1)\right\rangle .
\end{aligned}$$

Finally, we consider the second term in the expectation value of $\hat{H}$.

$$\begin{aligned}
& \left(\frac{n}{2}\right)!\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\,\middle|\,\sum_{j<k}\frac{1}{r_{j,k}}\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle \\
= \; & \sum_{j<k}\left\langle\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right| \\
& \left|\frac{1}{r_{j,k}}\sum_{\hat{P}_\alpha}(-1)^{p\left(\hat{P}_\alpha\right)}\hat{P}_\alpha\sum_{\hat{P}_\beta}(-1)^{p\left(\hat{P}_\beta\right)}\hat{P}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right\rangle .
\end{aligned}$$

In this case, the operator $1/r_{j,k}$ depends only on the coordinates of electrons j and k. Permutations which exchange any other indices produce a zero inner product. The only allowed permutations are the identity and the exchange of j and k. There are two cases. If electrons j and k have opposite spin, there is no permutation that exchanges j and k. If electrons j and k have the same spin, then there is an exchange term. Specifically,,

$$\begin{aligned}
&\left(\frac{n}{2}\right)!\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\left|\sum_{j<k}\frac{1}{r_{j,k}}\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right.\right\rangle \\
= &\sum_{\substack{j<k\\ \text{opposite spin}}}\left\langle \psi_{\lceil\frac{j}{2}\rceil}(j)\,\psi_{\lceil\frac{k}{2}\rceil}(k)\left|\frac{1}{r_{j,k}}\psi_{\lceil\frac{j}{2}\rceil}(j)\,\psi_{\lceil\frac{k}{2}\rceil}(k)\right.\right\rangle \\
&+\sum_{\substack{j<k\\ \text{same spin}}}\left\langle \psi_{\lceil\frac{j}{2}\rceil}(j)\,\psi_{\lceil\frac{k}{2}\rceil}(k)\left|\frac{1}{r_{j,k}}\left(1-\hat{P}_{j,k}\right)\psi_{\lceil\frac{j}{2}\rceil}(j)\,\psi_{\lceil\frac{k}{2}\rceil}(k)\right.\right\rangle \\
= &\sum_{\substack{j<k\\ \text{opposite spin}}}\left\langle \psi_{\lceil\frac{j}{2}\rceil}(1)\,\psi_{\lceil\frac{k}{2}\rceil}(2)\left|\frac{1}{r_{1,2}}\psi_{\lceil\frac{j}{2}\rceil}(1)\,\psi_{\lceil\frac{k}{2}\rceil}(2)\right.\right\rangle \qquad (8.16) \\
&+\sum_{\substack{j<k\\ \text{same spin}}}\left\langle \psi_{\lceil\frac{j}{2}\rceil}(1)\,\psi_{\lceil\frac{k}{2}\rceil}(2)\left|\frac{1}{r_{1,2}}\left(1-\hat{P}_{1,2}\right)\psi_{\lceil\frac{j}{2}\rceil}(1)\,\psi_{\lceil\frac{k}{2}\rceil}(2)\right.\right\rangle .
\end{aligned}$$

The first sum in Eq. 8.16 has terms, $(j,k) = (1,2),(1,4),(1,6),\ldots,(2,3),(2,5),\ldots,$ $(3,4),(3,6),\ldots$. These electron indices correspond to orbital indices, $(m,m') =$ $(1,1),(1,2),(1,3),\ldots,\ (1,2),(1,3),\ldots,\ (2,2),(2,3),\ldots$, respectively. The $m = m'$ pairs appear once, while the $m < m'$ pairs appear twice. This arises because there are two spin states. The second sum in Eq. 8.16 includes only $m < m'$ pairs, which also appear twice. The result is

$$\begin{aligned}
&\left(\frac{n}{2}\right)!\left\langle \psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\left|\sum_{j<k}\frac{1}{r_{j,k}}\hat{A}_\alpha\hat{A}_\beta\psi_1(1)\,\psi_1(2)\ldots\psi_{n/2}(n)\right.\right\rangle \\
= &\sum_{m}\left\langle \psi_m(1)\,\psi_m(2)\left|\frac{1}{r_{1,2}}\psi_m(1)\,\psi_m(2)\right.\right\rangle + 2\sum_{m<m'}\left\langle \psi_m(1)\,\psi_{m'}(2)\left|\frac{1}{r_{1,2}}\psi_m(1)\,\psi_{m'}(2)\right.\right\rangle \\
&+2\sum_{m<m'}\left\langle \psi_m(1)\,\psi_{m'}(2)\left|\frac{1}{r_{1,2}}\left(1-\hat{P}_{1,2}\right)\psi_m(1)\,\psi_{m'}(2)\right.\right\rangle \\
= &\sum_{m}\left\langle \psi_m(1)\,\psi_m(2)\left|\frac{1}{r_{1,2}}\left(2-\underset{=\psi_m(1)\psi_m(2)}{\hat{P}_{1,2}\right)\psi_m(1)\,\psi_m(2)}\right.\right\rangle \\
&+2\sum_{m<m'}\underset{\text{symmetric with respect to swapping } m \text{ and } m'}{\left\langle \psi_m(1)\,\psi_{m'}(2)\left|\frac{1}{r_{1,2}}\left(2-\hat{P}_{1,2}\right)\psi_m(1)\,\psi_{m'}(2)\right.\right\rangle} \\
= &\sum_{m=1}^{n/2}\sum_{m'=1}^{n/2}\left\langle \psi_m(1)\,\psi_{m'}(2)\left|\frac{1}{r_{1,2}}\left(2-\hat{P}_{1,2}\right)\psi_m(1)\,\psi_{m'}(2)\right.\right\rangle \\
= &\sum_{m=1}^{n/2}\sum_{m'=1}^{n/2}\left(2J_{m,m'}-K_{m,m'}\right).
\end{aligned}$$

where $J_{m,m'}$ is the *Coulomb integral* and $K_{m,m'}$ is the *exchange integral*. The Coulomb integral,

$$\begin{aligned} J_{m,m'} &= \left\langle \psi_m(1)\,\psi_{m'}(2) \left| \frac{1}{r_{1,2}} \psi_m(1)\,\psi_{m'}(2) \right.\right\rangle \\ &= \int\!\!\int \frac{1}{r_{1,2}} \left|\psi_m(1)\right|^2 \left|\psi_{m'}(2)\right|^2 d\mathbf{x}_1 d\mathbf{x}_2, \end{aligned}$$

is the repulsion energy experienced by one electron in orbital m due to another in orbital, m'. It is the repulsion energy averaged over the two electron probability distribution. The exchange integral,

$$\begin{aligned} K_{m,m'} &= \left\langle \psi_m(1)\,\psi_{m'}(2) \left| \frac{1}{r_{1,2}} \psi_m(2)\,\psi_{m'}(1) \right.\right\rangle \\ &= \int\!\!\int \frac{1}{r_{1,2}} \psi_m^*(1)\,\psi_{m'}(1)\,\psi_m(2)\,\psi_{m'}^*(2)\; d\mathbf{x}_1 d\mathbf{x}_2, \end{aligned}$$

does not have a classical interpretation - the distribution, $\psi_m^*(1)\,\psi_{m'}(1)$, is not always positive, and cannot be interpreted as a probability distribution. Equation 8.15 now takes the form,

$$\begin{aligned} &\left\langle \psi \left| \hat{H}\psi \right.\right\rangle \\ &= 2\sum_{m=1}^{n/2} \left\langle \psi_m(1) \left| \hat{H}_1 \psi_m(1) \right.\right\rangle + \sum_{m=1}^{n/2}\sum_{m'=1}^{n/2} \left(2J_{m,m'} - K_{m,m'}\right) + \sum_{a<b} \frac{Z_a Z_b}{r_{a,b}} \\ &= \sum_{m=1}^{n/2} \left\langle \psi_m(1) \left| \left(2\hat{H}_1 + \sum_{m'=1}^{n/2}\left(2\hat{J}_{m'} - \hat{K}_{m'}\right)\right) \psi_m(1) \right.\right\rangle + \sum_{a<b} \frac{Z_a Z_b}{r_{a,b}} \\ &= 2\sum_{m=1}^{n/2} \left\langle \psi_m(1) \left| \hat{H}_{\mathrm{HF}} \psi_m(1) \right.\right\rangle - \sum_{m=1}^{n/2}\sum_{m'=1}^{n/2} \left(2J_{m,m'} - K_{m,m'}\right) \qquad (8.17) \\ &\quad + \sum_{a<b} \frac{Z_a Z_b}{r_{a,b}} \end{aligned}$$

where

$$\hat{H}_{\mathrm{HF}} = \hat{H}_1 + \sum_{m'=1}^{n/2} \left(2\hat{J}_{m'} - \hat{K}_{m'}\right)$$

is the *Hartree-Fock operator*,

$$\hat{J}_{m'}\psi_m(1) = \left(\int \frac{1}{r_{1,2}} \left|\psi_{m'}(2)\right|^2 d\mathbf{x}_2\right) \psi_m(1)$$

is the *Coulomb operator*, and

$$\hat{K}_{m'}\psi_m(1) = \left(\int \frac{1}{r_{1,2}} \psi_m(2)\,\psi_{m'}^*(2)\; d\mathbf{x}_2\right) \psi_{m'}(1)$$

is the *exchange operator*.

Minimizing $\langle\psi\left|\hat{H}\psi\right\rangle$ with respect to variations in the molecular orbitals, $\psi_m(1)$, subject to the constraint that the orbitals be orthogonal leads to the eigenvalue equation for $\hat{H}_{\mathrm{HF}}$,

$$\hat{H}_{\mathrm{HF}}\psi_m(1) = \epsilon_m\psi_m(1)\,.$$

The best ground state, in the independent electron approximation, is the wavefunction constructed with double occupation of the $n/2$ lowest energy molecular orbitals. The energy of this state is given by twice (the orbitals are doubly occupied) the sum of the orbital energies, ϵ_m, minus the electron repulsion term.

$$2\sum_{m=1}^{n/2}\epsilon_m - \sum_{m=1}^{n/2}\sum_{m'=1}^{n/2}\left(2J_{m,m'} - K_{m,m'}\right) + \sum_{a<b}\frac{Z_aZ_b}{r_{a,b}}$$

The sum of the electron energies (the first term) counts the electron repulsion term twice. So, the electron repulsion term must be subtracted out to correct the overcounting.

The Hartree-Fock operator is a one electron operator which is represented by a matrix when a basis is chosen for the $\psi_m(1)$. Finding the eigenvalues and eigenvectors of a one electron matrix is a managable computation. The only twist here is that the Coulomb and exchange operators depend upon the occupied molecular orbitals. To do the computation, one must start with a reasonable guess for the molecular orbitals - perhaps the orbitals obtained with electron repulsion neglected. The guessed molecular orbitals are used to construct the Coulomb and exchange operators. The eigenvectors of the associated Hartree-Fock operator then provide the next guess for the molecular orbitals. The process is repeated until the molecular orbitals stop changing. In practice, a more elaborate iteration is required to quickly obtain *self-consistent* molecular orbitals. In Hartree-Fock theory, electron repulsion is treated as a self-consistent mean field.

The next step is to introduce an atomic orbital basis set. In practice, Z is replaced by Z_{eff}, specific to each orbital, with Z_{eff} determined by minimizing energy with respect to variation of Z_{eff}. More generally, superpositions of atomic orbitals with multiple Z_{eff} values are used. For example, when three different exponentially decaying terms are used, the basis is called a triple zeta basis. In any case, the molecular orbitals are written as linear combinations of some set of primitive basis functions which, in general, is not orthogonal.

$$\psi(1) = \sum_{j=1}^{d} c_j\varphi_j(1)\,,$$

where d is the number of basis functions.

To obtain the minimum total energy, we minimize the expectation value of

$\hat{H}_{\mathrm{HF}}$ with respect to variations in the coefficients.

$$\begin{aligned}
&\left\langle \psi(1) \left| \hat{H}_{\mathrm{HF}} \psi(1) \right.\right\rangle \\
=\ &\left\langle \sum_{j_1=1}^{d} c_{j_1}\varphi_{j_1}(1) \left| \hat{H}_1 \sum_{j_2=1}^{d} c_{j_2}\varphi_{j_2}(1) \right.\right\rangle \\
&+\sum_{m'=1}^{n/2} \left\langle \sum_{j_1} c_{j_1}\varphi_{j_1}(1)\,\psi_{m'}(2) \left| \frac{1}{r_{1,2}}\left(2-\hat{P}_{1,2}\right) \sum_{j_2} c_{j_2}\varphi_{j_2}(1)\,\psi_{m'}(2) \right.\right\rangle \\
=\ &\sum_{j_1,j_2} c_{j_1}^{*} c_{j_2} \left\langle \varphi_{j_1}(1) \left| \hat{H}_1 \varphi_{j_2}(1) \right.\right\rangle + \sum_{j_1,j_2} c_{j_1}^{*} c_{j_2} \sum_{m'=1}^{n/2} \left(2J_{m',j_1,j_2} - K_{m',j_1,j_2}\right) \\
=\ &\mathbf{c}^T \left(\mathsf{h} + \sum_{m'=1}^{n/2} \left(2\mathsf{J}_{m'} - \mathsf{K}_{m'}\right)\right) \mathbf{c}. \qquad \mathbf{c} \text{ is real} \qquad (8.18)
\end{aligned}$$

Minimizing the above quadratic form subject to the constraint that the molecular orbital be normalized gives rise to the generalized eigenvalue problem,

$$\left(\mathsf{h} + \sum_{m=1}^{n/2} \left(2\mathsf{J}_{m} - \mathsf{K}_{m}\right)\right) \mathbf{c} = \epsilon \mathsf{U}\mathbf{c},$$

where U is the overlap matrix. These equations (there is an equation for each eigenvector) are called the *Roothaan equations*. The eigenvector, $\mathbf{c}_1$, associated with the lowest orbital energy, ϵ_1, produces the minimum expectation of $\hat{H}_{\mathrm{HF}}$. The next lowest energy eigenvector, $\mathbf{c}_2$, produces the minimum expectation of $\hat{H}_{\mathrm{HF}}$ among vectors, $\mathbf{c}$, such that

$$\mathbf{c}_1^T \mathsf{U}\mathbf{c} = 0;$$

i.e., $\mathbf{c}$ represents molecular orbitals orthogonal to $\psi_1(1)$. This continues until there are enough molecular orbitals to accommodate all of the electrons.

In summary, one finds enough molecular orbitals to accommodate all the electrons. These orbitals are then used to construct new Coulomb and exchange operators. The new generalized eigenvalue equation is then solved for new molecular orbitals, and so. This procedure is iterated to convergence to self-consistent molecular orbitals. In practice, convergence of the iteration is accelerated by using the results of some number of successive steps to construct a quasi-Newton iteration (an approximate Newton-Rhaphson iteration in a low dimensional subspace).

Hartree-Fock excited states can also be constructed. The procedure is the same, except that the orbitals are occupied with an excited electron configuration. For example, an electron from one of the doubly occupied orbitals of the ground state is placed in a ground state unoccupied orbital. Iterating the Roothaan equations, starting with J_m and K_m computed with these different occupations, produces the molecular orbitals for the associated atomic or molecular electronic excited state. Excited states with one electron excited are generally the lowest energy excited states. These are the excited states accessible in electronic spectroscopy. Excited states with two or more electron excitations produce higher energy excited states. To observe two electron excitations with spectroscopy requires high intensity light.

The ionization, electron affinity and absorption frequencies are estimated using Hartree-Fock orbital energies, if we neglect the changes in J_m and K_m that result from removal of an electron, addition of an electron, or the excitation of an electron. For *ionization energy*, the result is

$$I_{\mathrm{HF},\,0} = -\epsilon_{\mathrm{HOMO}}.$$

This is called *Koopman's theorem.* Additional higher energy ionization thresholds - observable in *photoelectron spectra* - correspond to $-\epsilon_m$, for lower energy orbital. Photoelectron spectroscopy resolves ejected electrons according to kinetic energy. Multiple peaks are observed, corresponding to all orbitals with ionization energy below the photon energy of the light source. Ionization from lower energy orbitals gives electrons with lower kinetic energy.

Electron affinity is similarly approximated.

$$A_{\mathrm{HF},\,0} = -\epsilon_{\mathrm{LUMO}}.$$

Electronic transition frequencies are approximated by

$$\nu = \frac{\epsilon_{m'} - \epsilon_m}{h},$$

where ψ_m is an occupied orbital and $\psi_{m'}$ is an unoccupied orbital. Of course, transitions are subject to selection rules. Selection rules are generally associated with symmetry, and are derived using *group theory.*

8.5 Configuration interaction

The Hartree-Fock approximation provides a valuable benchmark for electronic computations. But, it neglects *electron correlation.* Electrons in molecules are not independent - i.e., molecular electronic wavefunctions are not antisymmetrized products one one electron functions. The most straightforward way to account for electron correlation is to generate a number of Hartree-Fock states. In addition to the Hartree-Fock ground state, Hartree-Fock excited states are constructed with excited electron configurations with single, double or even triple electron excitations. For example, one can consider all single and double electron excitations up to a maximum orbital energy, ϵ_{max}. Having constructed a suitable set of Hartree-Fock states, the matrix representation of the n electron Hamiltonian is computed for this basis. The eigenvalues of this matrix are the energy levels of the molecule. The associated eigenvectors provide coefficients for the expansion of the eigenstates of the molecule in terms of the Hartree-Fock states. The states of the system are expressed here as superpositions of Hartree-Fock states, each with a specific electron configuration. States which have more than one significant Hartree-Fock state component are said to exhibit *configuration interaction.*

Converged configuration interaction computations quickly become computationally intractable. The computation time scales as $n!$, and quickly gets out of hand. Consequently, many other methods have been developed to account for electron exchange with more tractable scaling with number of electrons. *Density functional theory* provides the lowest cost methods that accounts for some electron correlation. It can give good results for ground states when well calibrated. Density functional theory is generally semi-empirical, with parameters chosen to fit known data. More expensive methods include *coupled cluster theory* and *density matrix renormalization group theory.* These methods scale exponentially with number of electrons rather than factorially - the scaling of configuration interaction.

8.6 Hückel theory

Historically, it was important to simplify the computation of Hartree-Fock orbitals, via various approximations, to make the computations tractable. Sometimes these simplified approaches also provide a clearer view of electronic structure. This simplest approach is Hückel theory. Orginally, it was developed to describe only the π system of a molecule. Extended Hückel theory provides a full calculation including all the atomic orbitals. Here, we conisder only π systems.

While π bonding is weaker than σ bonding, it has the unique ability to form bonding networks over many atoms. σ bonds form with atomic orbitals pointing directly at each other. This makes σ bonds localized between two atoms. A p orbital on an atom overlaps equally well with the p orbitals on neighboring atoms in all directions, within the plane perpendicular to the p orbital. This allows the formation of bonding networks, and gives rise to the special stability of aromatic compounds.

The Hückel method starts with an empirical one electron Hamiltonian. It can be viewed as an empirical Hartree-Fock Hamiltonian, except that in Hückel theory the sum of the electron energies is not corrected for overcounting of electron repulsion. Electron repulsion with the σ system is not overcounted - it can be viewed as included in the the empirical parameters of the Hückel Hamiltonian. Electron repulsion between π electrons is essentially treated at the level of first order perturbation - it is included only in the diagonal elements of the Hückel Hamiltonian. No correction for overcounting is required because the electron repulsion energy is considered to be split between the electrons.

Polyenes, aromatic compounds, nucleic acids and other compounds have π systems arising from a planar σ system constructed from pairs of atomic $2sp^2$ orbitals (or $2sp^2$ and hydrogen 1s). The $2p_x$ orbitals orthogonal to the molecular plane provide the atomic basis for the π system. The Hückel Hamiltonian is a matrix whose diagonal elements are the energies of electrons occupying the $2p_x$ orbitals. These diagonal elements are denoted by α. Subscripts are used to distinguish different types of $2p_x$ orbital - e.g., on different types of atom. The off–diagonal elements - the couplings between the orbitals - are zero unless the atoms are directly bonded. Couplings decrease dramatically with distance, so we can neglect couplings of atoms two or more bonds apart. The couplings between bonded atoms are denoted by β, with subscripts distinguishing different types of bonds. Since the basis of $2p_x$ orbitals is non-orthogonal, there is a non-trivial overlap matrix U - i.e., it is not just the identity matrix. We suppose that the Hamiltonian constructed with α and β values, as just described, has already been transformed - i.e., $\mathsf{H} \to \mathsf{U}^{-1/2}\mathsf{H}\mathsf{U}^{-1/2}$ - to act on the $\mathbf{b}$ vectors which give the molecular orbital expansion coefficients, in terms of orthogonalized basis functions. Each orthogonalized basis function is a superposition of a $2p_x$ orbital and small amounts of neighboring $2p_x$ orbitals. We ignore this distinction and talk about the eigenvectors of the Hückel Hamiltonian, as though they were expansion coefficients in terms of atomic orbitals. The α and β values are ultimately empirical values which are chosen to give binding energies and absorption frequencies in accord with experiment.

The simplest Hückel treatments consider unsaturated hydrocarbons, and use a single α value for all the carbon $2p_x$ orbitals, and a single β value for every carbon-carbon π bond. The problem is simply to find the eigenvalues and eigenvectors of the Hückel Hamiltonian. In the π system for the ground state, electrons occupy the π molecular orbitals - in pairs - from the lowest energy orbital to HOMO. HOMO is determined by the total number of π electrons. The advantage of all the simplfying assumptions of Hückel theory becomes apparent when one considers examples.

8.6.1 Butadiene

1,3-Butadiene is the simplest molecule that demonstrates stabilization due to conjugation of double bonds. The Hückel Hamiltonian is the 4×4 matrix,

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

showing the connectivity of atoms - atom 1 is connected to 2, 2 to 3, and 3 to 4. The eigenvectors of this matrix do not depend on the values of α and β. Only the associated eigenvalues depend on α and β. This is because H can be written as

$$\mathsf{H} = \alpha 1 + \beta \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The $\alpha 1$ simply shifts the eigenvalues of the other term by α. The molecular orbitals of butadiene are the eigenvectors of

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The associated eigenvalues are scaled by β, and added to α, to give the molecular orbitals energies.

The Hückel eigenvalues and unnormalized eigenvectors for butadiene are tabulated as follows:

m	1	2	3	4
ϵ_m	$\alpha + \frac{\sqrt{5}+1}{2}\beta$	$\alpha + \frac{\sqrt{5}-1}{2}\beta$	$\alpha - \frac{\sqrt{5}-1}{2}\beta$	$\alpha - \frac{\sqrt{5}+1}{2}\beta$
$\mathbf{b}_m$	1	-1	1	-1
	$\frac{\sqrt{5}+1}{2}$	$\frac{-\sqrt{5}+1}{2}$	$\frac{-\sqrt{5}+1}{2}$	$\frac{\sqrt{5}+1}{2}$
	$\frac{\sqrt{5}+1}{2}$	$\frac{\sqrt{5}-1}{2}$	$\frac{-\sqrt{5}+1}{2}$	$\frac{-\sqrt{5}-1}{2}$
	1	1	1	1

The ground π state of butadiene is given by putting two electrons in each of $\mathbf{b}_1$ and $\mathbf{b}_2$. The associated π energy is

$$\begin{aligned} E_{\pi \text{ g.s.}} &= 2\left(\alpha + \frac{\sqrt{5}+1}{2}\beta\right) + 2\left(\alpha + \frac{\sqrt{5}-1}{2}\beta\right) \\ &= 4\alpha + 2\sqrt{5}\beta \\ &= 4\alpha + 4.472\beta. \end{aligned}$$

The π bonding energy is 4.472β. It is the stabilization of the π system relative to the four electrons isolated in $2\mathrm{p}_x$ orbitals. To better appreciate this energy, we constrast it with the energy of two isolated C-C π bonds - i.e., as in ethene. The π bonding energy in this case is 4β. The additional 0.472β of binding energy in butadiene is the conjugation energy of the two double bonds.

The molecular orbitals of butadiene are shown in Figure 8.4.

8.6.2 Allyl radical, cation and anion

The allyl radical π system is even simpler than that of 1,3-butadiene. Since there are only three π electrons, it is a radical system. The Lewis structure of the radical shows an isolated $2\mathrm{p}_x$ electron next to a double bond. Chemists say that resonance structures stabilize the $2\mathrm{p}_x$ electron, stabilizing the radical. We demonstrate this effect using Hückel theory.

The Hückel Hamiltonian for the allyl radical is

$$\mathsf{H} = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}.$$

The eigenvalues and normalized eigenvectors for allyl are tabulated as follows:

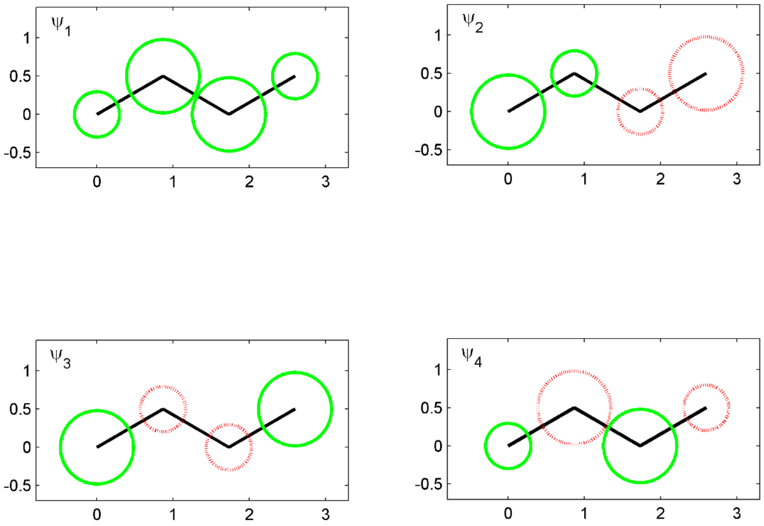


Figure 8.4: Hückel theory molecular orbitals for 1,3-butadiene. Orbital coefficients are repesented by circles with radius in proportion to $|b_{m,j}|$. A green circle corresponds to a coefficient with the same sign as the first coefficient on the left. Otherwise, the circle is red.

m	1	2	3
ϵ_m	$\alpha+\sqrt{2}\beta$	α	$\alpha-\sqrt{2}\beta$
$\mathbf{b}_m$	$\frac{1}{2}$ $\frac{\sqrt{2}}{2}$ $\frac{1}{2}$	$\frac{1}{\sqrt{2}}$ 0 $-\frac{1}{\sqrt{2}}$	$\frac{1}{2}$ $-\frac{\sqrt{2}}{2}$ $\frac{1}{2}$

The ground state π energy of the allyl radical, in Hückel theory, is

$$\begin{aligned} E_{\pi \text{ g.s.}} &= 2\left(\alpha+\sqrt{2}\beta\right)+\alpha \\ &= 3\alpha+2\sqrt{2}\beta \\ &= 3\alpha+2.828\beta. \end{aligned}$$

An isolated π bond and isolated $2\mathrm{p}_x$ electron have energy, $3\alpha+2\beta$. The additional 0.828β is the stabilization of the radical due to the neighboring π bond. In the calculation, the unpaired electron still has energy α. It is the energy of the two π bond electrons that is stabilized. This stabilizes the allyl radical, allowing it to form more readily than an alkyl radical.

If we subtract one electron, we get the allyl cation. The electron comes from ψ_2 of the π system - the highest energy occupied orbital. The cation π energy is $2\alpha+2\sqrt{2}\beta$, which is lower than the π energy of an isolated double bond, $2\alpha+2\beta$, by 0.828β. The allyl cation is stabilized relative to an alkyl cation by the stabilization of the π bonding electrons. Above, we see that the two bonding electrons are spread over all three $2\mathrm{p}_x$ orbitals - the increased space for these electrons lowers their energy.

If we add one electron, we get the allyl anion. The π bonding energy is again stabilized by 0.828β, with the effect of stabilizing the allyl anion relative to an alkyl carbanion.

8.6.3 Benzene and aromaticity

The Hückel Hamiltonian for benzene is the 6×6 matrix,

$$\mathsf{H}=\begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha \end{pmatrix}.$$

The unnormalized molecular orbitals and associated energies are tabulated as follows:

m	1	2	3	4	5	6
ϵ_m	$\alpha + 2\beta$	$\alpha + \beta$	$\alpha + \beta$	$\alpha - \beta$	$\alpha - \beta$	$\alpha - 2\beta$
$\mathbf{b}_m$	1	0	2	0	-2	-1
	1	1	1	1	1	1
	1	1	-1	-1	1	-1
	1	0	-2	0	-2	1
	1	-1	-1	1	1	-1
	1	-1	1	-1	1	1

The six-fold rotational symmetry gives rise to simple molecular orbitals, and degeneracy of the middle two orbital energies: $\epsilon_2 = \epsilon_3 = \alpha + \beta$ and $\epsilon_4 = \epsilon_5 = \alpha - \beta$. In fact, we could have used symmetry - *group theory* - to split the six dimensional vector space into the four distinct eigenspaces - two of them two dimensional.

The π energy of benzene is given by

$$2\,(\alpha + 2\beta) + 4\,(\alpha + \beta) = 6\alpha + 8\beta.$$

This is lower than the π energy of three isolated double bonds, $6\alpha + 6\beta$, by 2β. Benzene exhibits significant stability, in comparison with alkenes. It is the first member of the class of *aromatic* compounds - cyclic and polycyclic polyene compounds that have significant stabilization, and undergo different types of reactions than alkenes.

The molecular orbitals of butadiene are shown in Figure 8.5.

Consider the case of cyclobutadiene, with molecular orbitals and energies given by

m	1	2	3	4
ϵ_m	$\alpha + 2\beta$	α	α	$\alpha - 2\beta$
$\mathbf{b}_m$	1	1	0	-1
	1	0	-1	1
	1	-1	0	-1
	1	0	1	1

The ground state π energy is

$$2\,(\alpha + 2\beta) + 2\alpha = 4\alpha + 4\beta.$$

This is the same π energy as two isolated bonds. Cyclobutadiene has no stabilization due to conjugation. Cyclobutadiene is not a planar molecule - it is puckered. No stabilization results from combining the two π bonds into a single π system. It puckers to relieve strain. Cyclobutadiene is *antiaromatic.* At the same time, cyclobutadiene is highly strained. It is forced to use non-optimal hybridization to form the σ bonds with 90° bond angles. Consequently, its σ bonds are weaker than σ bonds with optimal bond angles.

We have seen that benzene has special stability, whereas cyclobutadiene has reduced stability. The difference results from the orbital energies, and their occupation. Cyclobutadiene has one pair of degenerate non-bonding orbitals that are both singly occupied in its ground state. The associated energy is α. Only the two lowest energy electrons in cyclobutadiene experience bonding

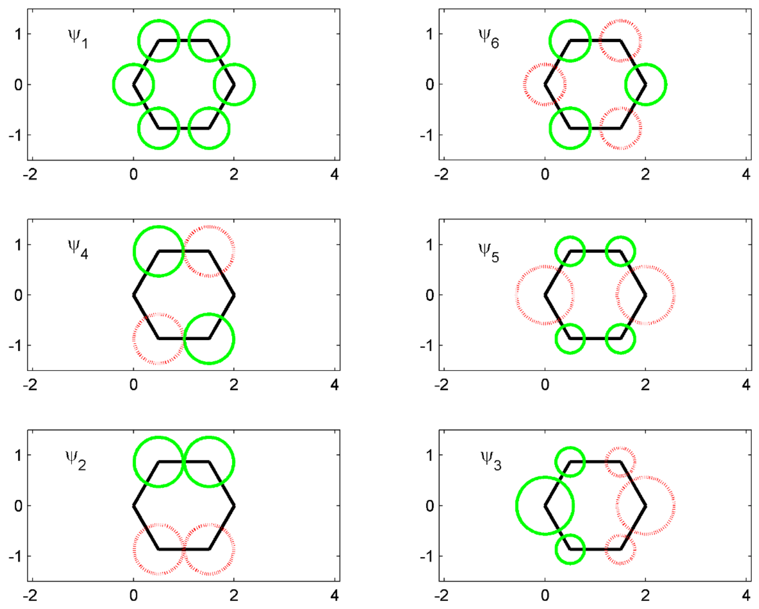


Figure 8.5: The molecular orbitals of benzene. The bottom two panels give the two degenerate HOMO states, while the middle two panels give the two degenerate LUMO states.

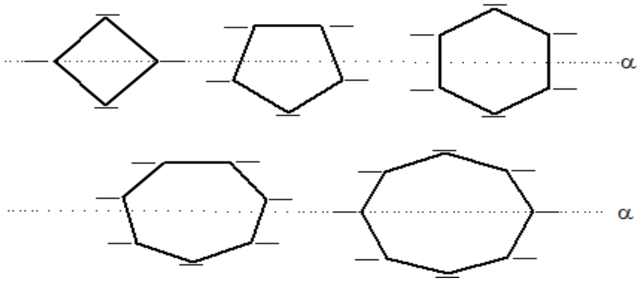


Figure 8.6: The energy levels - vertical axis - of cyclic π systems with four to eight atoms.

stabilization. Benzene has two pairs of degenerate orbitals. The two energies are split above and below α. However, only the lower energy orbitals are occupied (both doubly occupied) in neutral benzene. All electrons in benzene experience bonding stabilization. The two lowest energy π electrons experience extra stabilization - it is the source of the special stability of benzene.

Examing the eigenvalues of many cyclic π systems leads to the following observations: The lowest orbital energy of a cyclic π system is always non-degenerate. The associated molecular orbital has equal coefficients for all the atomic orbitals. Successive orbital energies are doubly degenerate. If the number of atoms in the ring is even, then the top orbital energy is non-degenerate. The energy levels of four to eight atom rings are shown in Fig. 8.6. The energies are the y coordinates of the vertices of a regular polygon arranged symmetrically about a vertical axis with a vertex at the bottom.

The stability of these species depends upon the orbital occupation. Stable species are formed when there are only closed shells. This means that the lowest energy molecular orbital and some number of subsequent shells - each shell consisting of two orbitals - are occupied. This corresponds with $4n + 2$ electrons, $n = 0, 1, 2, \ldots$ ($n = 0$ corresponds to the cyclopropenyl cation.) This is the origin of Hückel's rule for aromaticity. Cyclic π systems with $4n+2$ electrons have special stability - they are *aromatic*. Odd numbered electron systems are radicals. The other possibility is $4n$ electrons. These systems are diradicals, and no experience aromatic stabilization - they are *antiaromatic*. The two electrons at the highest energy are in non-bonding orbitals (even numbered rings), or weakly bonding orbitals (odd numbered rings). Note that odd numbered rings form form aromatic anions or cations - an electron must be added or removed to achieve $4n + 2$ electrons.

8.6.4 Atomic charge distribution

Hückel theory provides more than just molecular orbitals and associated energies. Most notably, one can determine the charge distribution within the molecule. The atomic cores, and the σ system, determine a molecular skeleton consisting of atoms with net core charge, $q_{\text{core},\ k}$ equal to one or two - two for oxygen atoms with two σ bonds, and nitrogen atoms with three σ bonds. The π system adds electron density to this skeleton. The total number of π electrons determines the net charge on the species. In Hückel theory, this charge can be broken into atomic components - i.e., we compute the net charge on each atom. Here, n resumes its role as the number of electrons (i.e., as opposed to the Hückel number introduced above). The net charge on atom, k, is

$$q_k = q_{\text{core},\ k} - \sum_{m=1}^{n} n_m \left| b_{m,k} \right|^2 ,$$

where $b_{m,k}$ is the coefficient associated with atom k in the m th molecular orbital. n_m is the occupation of the m th molecular orbital. The charge on the atom is expressed as the core charge plus the negative π electron charge. The latter equals the sum, over occupied orbital contributions, of the product of minus the orbital occupation and the probability a ψ_m electron is found at atom k. This formula can be applied to ground state or excited state electron configurations by changing the occupations.

For the ground state of neutral polyenes, $q_k = 0$ for all atoms - the charge is evenly spread throughout the molecule. When there are electronegative heteroatoms (e.g., O or N), the lower associated α values gives these atoms more electron density. This arises because the low α orbitals are featured more prominantly in the lower energy orbitals. This produces polarized ground state charge distributions, with negative charge accumulated onto heteroatoms (unless they have core charge two which is generally not completely canceled). Polarized charge distributions also arise in polycyclic systems with only carbon atoms.

The charge distribution of the ground state allyl radical is given by

$$\begin{aligned} q_1 &= 1 - \left(2\left(\frac{1}{2}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 \right) = 0 \\ q_2 &= 1 - \left(2\left(\frac{\sqrt{2}}{2}\right)^2 + (0)^2 \right) = 0 \\ q_3 &= 1 - \left(2\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 \right) = 0. \end{aligned}$$

The electron density is spread evenly across the allyl radical. However, the π electron density of the middle atom comes entirely from bonding electrons. The higher energy non-bonding electron is shared equally by the first and third atoms - i.e., these atoms equally share the radical character of allyl. These atoms will be the sites of reactivity.

In the case of the allyl cation,

$$\begin{aligned} q_1 &= 1-2\left(\frac{1}{2}\right)^2 = \frac{1}{2} \\ q_2 &= 1-2\frac{\sqrt{2}}{2}^2 = 0 \\ q_3 &= 1-2\left(\frac{1}{2}\right)^2 = \frac{1}{2}. \end{aligned}$$

The +1 charge is shared equally by the outer two atoms. The allyl cation acts as an electrophile at these sites.

In the case of the allyl anion,

$$\begin{aligned} q_1 &= 1-\left(2\left(\frac{1}{2}\right)^2+2\left(\frac{1}{\sqrt{2}}\right)^2\right) = -\frac{1}{2} \\ q_2 &= 1-\left(2\left(\frac{\sqrt{2}}{2}\right)^2+2\,(0)^2\right) = 0 \\ q_3 &= 1-\left(2\left(\frac{1}{2}\right)^2+2\left(-\frac{1}{\sqrt{2}}\right)^2\right) = -\frac{1}{2}. \end{aligned}$$

The −1 charge is shared equally by the outer two atoms. The allyl anion acts as a nucleophile at these sites. The charge distributions for the allyl cation and anion are consistent with the principle of like charge separation. In a conductor, any net charge repels itself into the surface of the conductor, so as to maximize the distance between like charges.

In the case of ethenoxide, CH_2CHO^-, the Hückel Hamiltonian is the same as that of allyl, except that the α value for $2p_x$ on oxygen is lower - oxygen is more electronegative. - and the C-O bond has a different β value. Using $\alpha\,(C) = -6.5$, $\alpha\,(O) = -11.9$, $\beta\,(\text{C-C}) = -2.7$ and $\beta\,(\text{C-O}) = -2.16$ (all in eV), we get

m	1	2	3
ϵ_m	−12.806	−8.586	−3.508
	0.1634	−0.7351	0.6580
$\mathbf{b}_m$	0.3816	−0.5680	−0.7292
	0.9098	0.3702	0.1877

and

$$\begin{aligned} q_1 &= 1-\left(2\,(0.1634)^2+2\,(-0.7351)^2\right) = -0.134 \\ q_2 &= 1-\left(2\,(0.3816)^2+2\,(-0.5680)^2\right) = 0.064 \\ q_3 &= 1-\left(2\,(0.9098)^2+2\,(0.3702)^2\right) = -0.930. \end{aligned}$$

The core charge on oxygen, and the number of π electrons it contributes, can be deduced from the Lewis structure of CH_2CHO^-. In that structure, oxygen has one σ bond, and three lone pairs - one in a 2sp orbital, one in $2p_y$, and one in $2p_x$. Subtracting the π system $2p_x$ electrons from oxygen gives it a formal charge of $+1$. This is the core charge of oxygen. Since oxygen contributes two electrons to the π system, while each carbon atom contributes one, there are four π electrons in total.

As expected, the -1 charge of ethenoxide is mostly on the oxygen atom. However, some of the charge is shared with the first carbon atom. Moreover, the minus charge on oxygen polarizes the C-C π bond such that the carbon next to the oxygen has a net positive charge. This puts more negative charge at the first carbon atom. The negative charge on the first carbon is consistent with ethenoxide acting as a nucleophile at this carbon atom - ethenoxide has carbanion character.

8.6.5 Self-consistent Hückel theory

The above calculations reveal a limitation of Hückel theory. The same α value is used for a neutral atom, and atoms with net positive or negative charge. Differences in electron repulsion of the π electrons on the atom are neglected. The simplest improvement of Hückel theory is to treat the electron repulsion on each atom at the level of first order perturbation theory - i.e., it affects only the diagonal elements of the Hamiltonian matrix. Thus, we let α_k depend on the charge on atom k. This simplest dependence is linear. Specifically,

$$\alpha_k = \alpha_k^{(0)} - \lambda_k q_k, \qquad \lambda_k > 0.$$

$\alpha_k = \alpha_k^{(0)}$ for a neutral atom. If the charge is positive, α_k is reduced by λ_k, per unit charge. If the charge is negative, α_k is larger than the neutral reference value. This charge dependence of the α_k requires an iteration to yield self-consistent solutions - i.e., the charges used to get the α_k are the same as those produced by the α_k. This is *self-consistent Hückel theory.*

Applying self-consistent Hückel theory with $\lambda_k = 3.78$ for all three atoms gives molecular orbitals and energies tabulated as follows:

m	1	2	3
ϵ_m	-10.55	-7.47	-3.10
$\mathbf{b}_m$	0.362	-0.662	0.657
	0.611	-0.364	-0.703
	0.704	0.656	0.272

The associated self-consistent charge distribution is

$$\begin{aligned} q_1 &= -0.14 \\ q_2 &= -0.01 \\ q_3 &= -0.85. \end{aligned}$$

The negative charge is shared to a greater extent with the carbon atoms, than in the previous calculation. Here, we account for electron repulsion on an oxygen atom with charge near −1. This raises the α value of oxygen closer to that of the carbon atoms, and reduces accumulation of negative charge on oxygen. Again, we see that the minus charge is primarily shared by the first carbon atom, consistent with reactivity at carbon atom - as a nucleophile.

8.7 Appendix: ground state energy estimate for $\mathbf{H_2^+}$ and $\mathbf{H_2}$

In Sec. 8.1 - see Eq. 8.6 - the ground state energy of $\mathrm{H_2}$ was estimated to be the energy of two H atoms plus interaction energies expressed in terms of matrix elements of $1/r_{\mathrm{B}}$. Here, we evaluate the matrix elements using hydrogen-like ion orbitals with a Z dependence, so that the effective nuclear charge can be adjusted to give the best possible energy estimate - in accord with the variational principle.

With electron repulsion neglected, the ground state energy of $\mathrm{H_2}$ in the approximate ground state, provided by Eq. 8.4 , has the form,

$$E_{\mathrm{H_2\ g.s.}} = 2\left(\alpha_Z + \beta_Z\right) + \frac{1}{r_{\mathrm{A,\ B}}},$$

where

$$\begin{aligned} &\alpha_Z \\ = \;& \frac{\left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{A,\ 1}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle - \left(\left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{B,\ 1}} (1\mathrm{s})_{\mathrm{B}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{B}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle \right) u}{1 - u^2} \\ &\beta_Z \\ = \;& \frac{-\left(\left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{A,\ 1}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle \right) u + \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{B,\ 1}} (1\mathrm{s})_{\mathrm{B}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{B}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle}{1 - u^2}. \end{aligned}$$

Thus,

$$\begin{aligned} &E_{\mathrm{H_2\ g.s.}} \\ = \;& 2\frac{\left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{A,\ 1}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \hat{H}_{\mathrm{B,\ 1}} (1\mathrm{s})_{\mathrm{B}} \right.\right\rangle + \left\langle (1\mathrm{s})_{\mathrm{B}} \left| \frac{-1}{r_{\mathrm{B,\ 1}}} (1\mathrm{s})_{\mathrm{A}} \right.\right\rangle}{1 + u} + \frac{1}{r_{\mathrm{A,\ B}}} \end{aligned}$$

In Sec. 8.1, the atomic Hamiltonian expectations simplified because the atomic orbitals were eigenfunctions of the atomic Hamiltonians. Here, the 1s orbitals have variable atomic number, Z. They are eigenfunctions of hydrogen-like ion Hamiltonians. We write

$$\begin{aligned} \hat{H}_{\mathrm{A,\ 1}} &= -\frac{\nabla_1^2}{2} - \frac{1}{r_{\mathrm{A,\ 1}}} \\ &= -\frac{\nabla_1^2}{2} - \frac{Z}{r_{\mathrm{A,\ 1}}} + \frac{Z-1}{r_{\mathrm{A,\ 1}}}, \end{aligned}$$

and note that

$$\begin{aligned}
&\left\langle (1\mathrm{s})_\mathrm{A} \left| \hat{H}_{\mathrm{A},\,1} (1\mathrm{s})_\mathrm{A} \right.\right\rangle \\
= \;& \left\langle (1\mathrm{s})_\mathrm{A} \left| \left(-\frac{\nabla_1^2}{2} - \frac{Z}{r_{\mathrm{A},\,1}}\right)(1\mathrm{s})_\mathrm{A} + \frac{Z-1}{r_{\mathrm{A},\,1}}(1\mathrm{s})_\mathrm{A} \right.\right\rangle \\
= \;& \left\langle (1\mathrm{s})_\mathrm{A} \left| \left(-\frac{Z^2}{2}\right)(1\mathrm{s})_\mathrm{A} + \frac{Z-1}{r_{\mathrm{A},\,1}}(1\mathrm{s})_\mathrm{A} \right.\right\rangle \\
= \;& -\frac{Z^2}{2} \underbrace{\langle (1\mathrm{s})_\mathrm{A} | (1\mathrm{s})_\mathrm{A} \rangle}_{=1} + (Z-1) \left\langle (1\mathrm{s})_\mathrm{A} \left| \frac{1}{r_{\mathrm{A},\,1}} (1\mathrm{s})_\mathrm{A} \right.\right\rangle \\
= \;& -\frac{Z^2}{2} + (Z-1) \left\langle (1\mathrm{s})_\mathrm{A} \left| \frac{1}{r_{\mathrm{A},\,1}} (1\mathrm{s})_\mathrm{A} \right.\right\rangle . \qquad (8.19)
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\left\langle (1\mathrm{s})_\mathrm{A} \left| \hat{H}_{\mathrm{B},\,1} (1\mathrm{s})_\mathrm{B} \right.\right\rangle \\
= \;& \left\langle (1\mathrm{s})_\mathrm{A} \left| \left(-\frac{\nabla_1^2}{2} - \frac{Z}{r_{\mathrm{B},\,1}}\right)(1\mathrm{s})_\mathrm{B} + \frac{Z-1}{r_{\mathrm{B},\,1}}(1\mathrm{s})_\mathrm{B} \right.\right\rangle \\
= \;& \left\langle (1\mathrm{s})_\mathrm{A} \left| \left(-\frac{Z^2}{2}\right)(1\mathrm{s})_\mathrm{B} + \frac{Z-1}{r_{\mathrm{B},\,1}}(1\mathrm{s})_\mathrm{B} \right.\right\rangle \\
= \;& -\frac{Z^2}{2} \underbrace{\langle (1\mathrm{s})_\mathrm{A} | (1\mathrm{s})_\mathrm{B} \rangle}_{=u} + (Z-1) \left\langle (1\mathrm{s})_\mathrm{A} \left| \frac{1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_\mathrm{B} \right.\right\rangle \\
= \;& -\frac{Z^2}{2} u + (Z-1) \left\langle (1\mathrm{s})_\mathrm{A} \left| \frac{1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_\mathrm{B} \right.\right\rangle . \qquad (8.20)
\end{aligned}$$

The $1/r_{\mathrm{A},\,1}$ expectation value in Eq. 8.19 can be evaluated using spherical coordinates;

$$\begin{aligned}
&\left\langle (1\mathrm{s})_\mathrm{A} \left| \frac{1}{r_{\mathrm{A},\,1}} (1\mathrm{s})_\mathrm{A} \right.\right\rangle \\
= \;& Z^3 \pi^{-1} \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} \frac{1}{r_1} \exp\left(-2Zr_1\right) r_1^2 \sin\theta_1 \, dr_1 d\theta_1 d\phi_1 \\
= \;& 4Z^3 \int_0^{\infty} r \exp\left(-2Zr\right) \, dr \\
= \;& Z \int_0^{\infty} \rho \exp\left(-\rho\right) \, d\rho \\
= \;& Z \left[\left(-\rho - 1\right) \exp\left(-\rho\right)\right]_0^{\infty} \\
= \;& Z.
\end{aligned}$$

The remaining matrix elements are evaluated using elliptic coordinates. These coordinates include an angle, ϕ, that varies from 0 to 2π. A change in ϕ is a rotation about the z axis - the bond axis - . It is essentially the same as the angle ϕ of spherical coordinates. The other two coordinates have the following ranges: $1 < \xi < \infty$ and $-1 < \eta < 1$. The lines of constant ξ and ϕ are ellipses with focal points at the two nuclei. Lines of constant η and ϕ are hyperbolas with the same two focal points. We do not need all the details of the elliptic coordinates. We need only the volume element,

$$dxdydz = \frac{r_{\mathrm{A,B}}^3}{8}\left(\xi^2 - \eta^2\right) d\xi d\eta d\phi,$$

and the distances to the nuclei,

$$r_{\mathrm{A,\,1}} = \frac{r_{\mathrm{A,B}}}{2}\left(\xi - \eta\right)$$

and

$$r_{\mathrm{B,\,1}} = \frac{r_{\mathrm{A,B}}}{2}\left(\xi + \eta\right).$$

In terms of elliptic coordinates, the orbitals take the form,

$$(1\mathrm{s})_{\mathrm{A}}(1) = Z^{3/2}\pi^{-1/2}\exp\left(-\frac{Zr_{\mathrm{A,B}}}{2}\left(\xi_1 - \eta_1\right)\right)$$

and

$$(1\mathrm{s})_{\mathrm{B}}(1) = Z^{3/2}\pi^{-1/2}\exp\left(-\frac{Zr_{\mathrm{A,B}}}{2}\left(\xi_1 + \eta_1\right)\right).$$

Therefore,

$$(1\mathrm{s})_{\mathrm{A}}(1)\,(1\mathrm{s})_{\mathrm{B}}(1) = Z^3\pi^{-1}\exp\left(-Zr_{\mathrm{A,B}}\xi_1\right)$$

and

$$\begin{aligned}
u &= \langle (1\mathrm{s})_{\mathrm{A}} \,|(1\mathrm{s})_{\mathrm{B}}\rangle \\
&= Z^3\pi^{-1}\int_0^{2\pi}\int_{-1}^{1}\int_1^{\infty}\frac{r_{\mathrm{A,B}}^3}{8}\left(\xi_1^2 - \eta_1^2\right)\exp\left(-Zr_{\mathrm{A,B}}\xi_1\right) d\xi_1 d\eta_1 d\phi_1 \\
&= \frac{Z^3 r_{\mathrm{A,B}}^3}{4}\int_{-1}^{1}\int_1^{\infty}\left(\xi^2 - \eta^2\right)\exp\left(-Zr_{\mathrm{A,B}}\xi\right) d\xi d\eta \\
&= \frac{Z^3 r_{\mathrm{A,B}}^3}{4}\int_1^{\infty}\int_{-1}^{1}\left(\xi^2 - \eta^2\right) d\eta \exp\left(-Zr_{\mathrm{A,B}}\xi\right) d\xi \\
&= \frac{1}{4}\int_{Zr_{\mathrm{A,B}}}^{\infty}\int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}}\left(x^2 - y^2\right) dy \exp\left(-x\right) dx \\
&= \frac{1}{4}\int_{Zr_{\mathrm{A,B}}}^{\infty}\left(2Zr_{\mathrm{A,B}}x^2 - \frac{2Z^3 r_{\mathrm{A,B}}^3}{3}\right)\exp\left(-x\right) dx \\
&= \frac{Zr_{\mathrm{A,B}}}{2}\left[\left(\left(-x^2 - 2x - 2\right) + \frac{Z^2 r_{\mathrm{A,B}}^2}{3}\right)\exp\left(-x\right)\right]_{Zr_{\mathrm{A,B}}}^{\infty} \\
&= \left(\frac{Z^2 r_{\mathrm{A,B}}^2}{3} + Zr_{\mathrm{A,B}} + 1\right)\exp\left(-Zr_{\mathrm{A,B}}\right).
\end{aligned}$$

The remaining matrix elements are

$$\begin{aligned}
&\left\langle (1\mathrm{s})_{\mathrm{A}} \left| \frac{1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_{\mathrm{A}} \right. \right\rangle \\
= \;& Z^3\pi^{-1} \int_0^{2\pi} \int_{-1}^{1} \int_1^{\infty} \frac{r_{\mathrm{A,B}}^3}{8} \frac{\left(\xi_1^2 - \eta_1^2\right)}{\frac{r_{\mathrm{A,B}}}{2}\left(\xi_1 + \eta_1\right)} \exp\left(-Zr_{\mathrm{A,B}}\left(\xi_1 - \eta_1\right)\right) d\xi_1 d\eta_1 d\phi_1 \\
= \;& \frac{Z^3 r_{\mathrm{A,B}}^3}{2r_{\mathrm{A,B}}} \int_{-1}^{1} \int_1^{\infty} \left(\xi - \eta\right) \exp\left(-Zr_{\mathrm{A,B}}\left(\xi - \eta\right)\right) d\xi d\eta \\
= \;& \frac{1}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \int_{Zr_{\mathrm{A,B}}}^{\infty} \left(x - y\right) \exp\left(-x + y\right) dxdy \\
= \;& \frac{1}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \left[\left(-x - 1 + y\right) \exp\left(-x\right)\right]_{Zr_{\mathrm{A,B}}}^{\infty} \exp\left(y\right) dy \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \left(Zr_{\mathrm{A,B}} + 1 - y\right) \exp\left(y\right) dy \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} \left[\left(Zr_{\mathrm{A,B}} + 1 - y + 1\right) \exp\left(y\right)\right]_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} 2\left(\exp\left(Zr_{\mathrm{A,B}}\right) - \left(1 + Zr_{\mathrm{A,B}}\right) \exp\left(-Zr_{\mathrm{A,B}}\right)\right) \\
= \;& \frac{1}{r_{\mathrm{A,B}}} \left(1 - \left(1 + Zr_{\mathrm{A,B}}\right) \exp\left(-2Zr_{\mathrm{A,B}}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
&\left\langle (1\mathrm{s})_{\mathrm{B}} \left| \frac{1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_{\mathrm{A}} \right. \right\rangle = \left\langle (1\mathrm{s})_{\mathrm{A}} \left| \frac{1}{r_{\mathrm{B},\,1}} (1\mathrm{s})_{\mathrm{B}} \right. \right\rangle \\
= \;& Z^3\pi^{-1} \int_0^{2\pi} \int_{-1}^{1} \int_1^{\infty} \frac{r_{\mathrm{A,B}}^3}{8} \frac{\left(\xi_1^2 - \eta_1^2\right)}{\frac{r_{\mathrm{A,B}}}{2}\left(\xi_1 + \eta_1\right)} \exp\left(-Zr_{\mathrm{A,B}}\xi_1\right) d\xi_1 d\eta_1 d\phi_1 \\
= \;& \frac{Z^3 r_{\mathrm{A,B}}^3}{2r_{\mathrm{A,B}}} \int_{-1}^{1} \int_1^{\infty} \left(\xi - \eta\right) \exp\left(-Zr_{\mathrm{A,B}}\xi\right) d\xi d\eta \\
= \;& \frac{1}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \int_{Zr_{\mathrm{A,B}}}^{\infty} \left(x - y\right) \exp\left(-x\right) dxdy \\
= \;& \frac{1}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \left[\left(-x - 1 + y\right) \exp\left(-x\right)\right]_{Zr_{\mathrm{A,B}}}^{\infty} dy \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} \int_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \left(Zr_{\mathrm{A,B}} + 1 - y\right) dy \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} \left[\left(Zr_{\mathrm{A,B}} + 1\right) y - \frac{y^2}{2}\right]_{-Zr_{\mathrm{A,B}}}^{Zr_{\mathrm{A,B}}} \\
= \;& \frac{\exp\left(-Zr_{\mathrm{A,B}}\right)}{2r_{\mathrm{A,B}}} 2\left(Zr_{\mathrm{A,B}} + 1\right) Zr_{\mathrm{A,B}} \\
= \;& Z\left(1 + Zr_{\mathrm{A,B}}\right) \exp\left(-Zr_{\mathrm{A,B}}\right).
\end{aligned}$$

Altogether, with $r = r_{\mathrm{A,\ B}}$,

$$\begin{aligned}
& E_{\mathrm{H}_2\ \mathrm{g.s.}} \\
= \ & \frac{2}{1+u}\left(-\frac{Z^2}{2} + (Z-1)\,Z - \frac{1}{r}\,(1-(1+Zr)\exp(-2Zr))\right. \\
& \left. -\frac{Z^2}{2}u + (Z-2)\,Z\,(1+Zr)\exp(-Zr)\right) + \frac{1}{r} \\
= \ & -Z^2 + \frac{2}{1+u}\left(Z\,(Z-1) - \frac{1}{r}\,(1-(1+Zr)\exp(-2Zr))\right. \\
& \left. + Z\,(Z-2)\,(1+Zr)\exp(-Zr)\right) + \frac{1}{r} \\
= \ & -Z^2 + \frac{1}{r} \\
& +2\frac{Z\,(Z-1) - \frac{1}{r}\,(1-(1+Zr)\exp(-2Zr)) + Z\,(Z-2)\,(1+Zr)\exp(-Zr)}{1+\left(1+Zr+\frac{Z^2r^2}{3}\right)\exp(-Zr)}
\end{aligned}$$

and

Consider the two limiting cases, $r \to \infty$ and $r \to 0$ (the united atom limit). If the two atoms are far apart ($r \to \infty$), $\exp(-Zr) \cong 0$, and

$$\begin{aligned}
E_{\mathrm{H}_2\ \mathrm{g.s.}} &\cong -Z^2 + \frac{1}{r} + 2\left(Z\,(Z-1) - \frac{1}{r}\right) \\
&= Z^2 - 2Z - \frac{1}{r}.
\end{aligned}$$

The optimal Z value is the solution to

$$2Z - 2 = 0;$$

i.e., $Z = 1$. The optimal ground state energy estimate is

$$E_{\mathrm{H}_2\ \mathrm{g.s.}} \cong -1 - \frac{1}{r}.$$

When the electron repulsion is added, the net $-1/r$ attraction is canceled and the we get the energy of two hydrogen atoms.

If the two atoms are close together ($r \to 0$), $\exp(-Zr) \cong 1 - Zr$, positive powers of r can be neglected and

$$\begin{aligned}
& E_{\mathrm{H}_2\ \mathrm{g.s.}} \\
\cong \ & -Z^2 + \frac{1}{r} + 2\frac{Z\,(Z-1) - \frac{1}{r}\,(1-(1+Zr)\,(1-2Zr)) + Z\,(Z-2)\,(1+Zr)\,(1-Zr)}{2} \\
\cong \ & -Z^2 + \frac{1}{r} + Z\,(Z-1) - Z + Z\,(Z-2) \\
= \ & Z^2 - 4Z + \frac{1}{r}.
\end{aligned}$$

Here, the optimal Z value solves

$$2Z - 4 = 0;$$

i.e., $Z = 2$. The optimal ground state energy estimate is

$$E_{\mathrm{H}_2\ \mathrm{g.s.}} \cong -4 + \frac{1}{r}.$$

If, instead of two hydrogen atoms, we had two deuterium atoms, then when the nuclei come together they can form a helium nucleus. The $+1/r$ repulsion is part of the nuclear energy. It is more than canceled by the released binding energy resulting from the formation of helium from two deuterium nuclei. The remaining energy, -4, is the energy of helium with electron repulsion neglected.