Lecture Notes in Physical Chemistry
Semester 1: Quantum mechanics

George C. McBane
Department of Chemistry
Grand Valley State University

December 1, 2016

2

Chapter 1

Introduction and background

1.1 Manifestations of quantum mechanics

Chemistry is, at its heart, the study of electrons and nuclei in atoms and molecules. Physical chemistry is that branch of chemistry concerned with the physical basis and explanation of chemical behavior. It seeks a quantitative, predictive understanding of chemical phenomena.

Because chemistry is all about electrons and nuclei, to really understand chemistry we must understand the rules that govern their behavior.

Important Fact number 1: $F = ma$ does not apply to electrons or nuclei.

The simplest evidence is the emission spectrum of hydrogen atoms you learned about in freshman chemistry: a series of "lines" shown in Figure 1.1.

The separate lines indicate the electron can only have particular total energies. Contrast this result with the solar system, a mathematically identical but much larger system: asteroids live at many different distances between Mars and Jupiter.

Several kinds of experimental results, reported between ca. 1860 and 1925, made it clear that $F = ma$ really did not apply to the parts of atoms or to the atoms themselves.

1. Heat capacities of gases. You may recall that the heat capacity C of an object is the proportionality constant between the amount of energy added to it by heating, q, and its change in temperature, ΔT: $q = C\Delta T$. The heat capacity is somewhat different depending on whether you carry out the experiment with the sample at constant *pressure* or at constant *volume*; we call the resulting quantities C_p and C_V. For gases it is possible to measure the *heat capacity ratio* $\gamma = C_p/C_V$ accurately in relatively straightforward experiments. (One measures the *speed of sound* in the gas by measuring the wavelength λ and frequency f of a standing sound wave in a tube

Fig. 1. Emission Spectrum of the Hydrogen Atom in the Visible and Near Ultraviolet Region [Balmer series, Herzberg (41)]. H_∞ gives the theoretical position of the series limit.
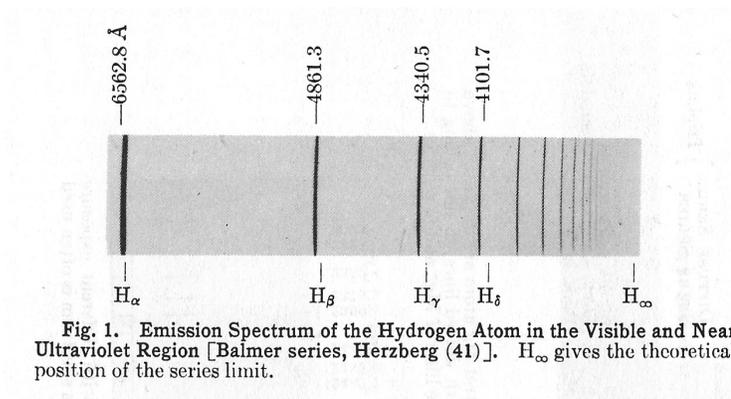
Figure 1.1: Emission spectrum of atomic H, from G. Herzberg, *Atomic Spectra and Atomic Structure* (Dover,1944); reprinted from *Annalen der Physik* (4), 84, 565 (1927). The dark lines appear where the emitted light, after dispersion by a grating, falls on photographic film.

containing the gas. The heat capacity ratio can then be extracted from $\lambda f = v_s = \sqrt{\gamma RT/M}$). Classical physics gives simple predictions for γ: $\gamma = {}^5\!/_3 \approx 1.667$ for monatomic gases, $\gamma = {}^9\!/_7 \approx 1.286$ for diatomic gases, with values that decrease toward 1 as the molecule becomes more complicated and are independent of temperature.

Table 1.1 gives measured values of γ for several gases. For gases other than the monatomics, they do not agree with the predictions at the temperatures shown; generally the observed values of γ are higher than those predicted by classical mechanics. In fact as the temperature becomes higher, the measured values decrease, and at very high temperatures, they do approach the classically predicted values. (Iodine is already near the classical expectation at 185 °C.)

Maxwell treated heat capacities of gases in his original 1859 work on the kinetic-molecular theory of gases, pointing out at the end of the paper that observed heat capacity ratios were inconsistent with a gas model of rigid elastic particles of any shape. In an 1875 lecture he described the disagreement with observed heat capacity ratios as "the greatest difficulty yet encountered by the molecular theory".[1] (He gave the expected value

---

[1] Both papers are included in *The Scientific Papers of James Clerk Maxwell*, ed. W. D. Niven, Cambridge Univ. Press 1890, reprinted by Dover in 1965. They are on p. 377 of Vol. I and p. 418 of Vol. II.

Table 1.1 Heat capacity ratios for gases, from Feynman, *Lectures on Physics*, vol. 1 p. 40-8. For NH_3 the classical expectation is $\gamma = {}^{10}/_9 \approx 1.111$; for C_2H_6 it is $\gamma = {}^{22}/_{21} \approx 1.048$.

Gas	He	Ar	Kr	H_2	O_2	HI	Br_2	I_2	NH_3	C_2H_6
T/°C	-180	15	19	100	100	100	300	185	15	15
γ	1.66	1.668	1.68	1.404	1.399	1.40	1.32	1.30	1.31	1.22

of $\gamma = 1.66$ for simple spherical particles, but no monatomic gases were known at the time and he said this value was "too great for any real gas.") The problem was not resolved until quantum mechanics explained the quantization of energy levels for rotation and vibration of molecules.

2. Black body radiation: A heated object emits light with a distribution of frequencies that had been measured carefully in the late 19th century. If you treat the object as a collection of electromagnetic oscillators and apply classical statistical mechanics (developed between 1840 and 1900), you get the prediction that even cool objects should emit bright light in the deep UV. This problem was described by Rayleigh in 1900. In the same year Planck showed that by making the assumption that oscillators could have only discrete amounts of energy, he could predict the experimentally observed frequency distribution. This development was the first real step toward quantum mechanics.

   Levine §18.1

3. The photoelectric effect: if you shine a bright red light at a piece of metal, nothing happens, but if instead you use a weak UV lamp such as a low-pressure mercury lamp, electrons start popping out of the metal surface. It was not clear why this should happen until Einstein's 1905 explanation that introduced the idea of a "photon". (Einstein called them "light quanta". The name *photon* was coined in 1926 by G. N. Lewis, better known for Lewis dot structures.) It extended Planck's 1900 hypothesis of quantization of oscillator energy levels to quantization of the electromagnetic field itself.

   Levine §18.2

4. Heat capacities of crystals. This problem is mathematically similar to that of black body radiation, except instead of electrons oscillating you have atoms. This problem was also solved by Einstein; his solution was later refined by Debye.

5. Spectra. Atomic and molecular spectra were well known in the second half of the nineteenth century, and were recognized as somehow related to internal motions of the particles. These cry out for quantization so loudly

they are impossible to ignore. One example from my own lab is shown in Figure 1.2. The figure shows a photoionization spectrum of CO after rotationless CO molecules collided with fast-moving He atoms. The vertical intensity is proportional to the CO ionization rate; the horizontal axis gives the wavelength of high-intensity light shone on the sample. The series of sharp lines indicates that the CO molecules cannot acquire just any rotational energy, but only a few select rotational energies.

Spectra are a reflection of *structure*. If you try to use classical mechanics to describe the behavior of even simple molecules (say, H_2 or CH_4) you fail miserably. Your molecules collapse, the electrons fall into the protons, and everything explodes.

The rules that govern electrons and nuclei are called *quantum mechanics* (QM). They were developed between 1910 and 1927, by several leaps from "classical" (Newtonian and extensions) mechanics.

Quantum mechanics has a reputation for being mathematical, abstract, puzzling, and difficult. In fact it is no more difficult to use than any other branch of physics. It is mathematical, but then it must be to be both general and quantitative, and its mathematics is not particularly hard. The precision it provides in prediction of measurable quantities is breathtaking. It is among the most successful theories of science; in the 90 years since its development, no experiments have been performed that show any flaws in its predictions. This is all the more remarkable given that QM was developed to describe atomic structure, but was then applied to the structure of the nucleus, and then again to the structure of the nucleons. Neither of these experimental leaps required any fundamental changes to quantum mechanics.

On the other hand, QM *is* puzzling in a philosophical sense. Taken as a set of instructions for predicting the outcome of experiments, it is fairly straightforward to apply and unerringly accurate. But some of the predictions it makes are altogether astonishing. In every case when such predictions have been tested, they have been correct, so we can only conclude that we live in an astonishing world. By the end of our study of QM we will have encountered some of the philosophically difficult questions, so you can appreciate why ninety years of thinking about quantum mechanics has still not provided very satisfactory explanations of why it works or exactly what some of its main features mean. But first, you must learn how to *do* quantum mechanics, that is, how it works and how to apply it to chemical problems. That is interesting, and fun, and not very hard; enjoy!
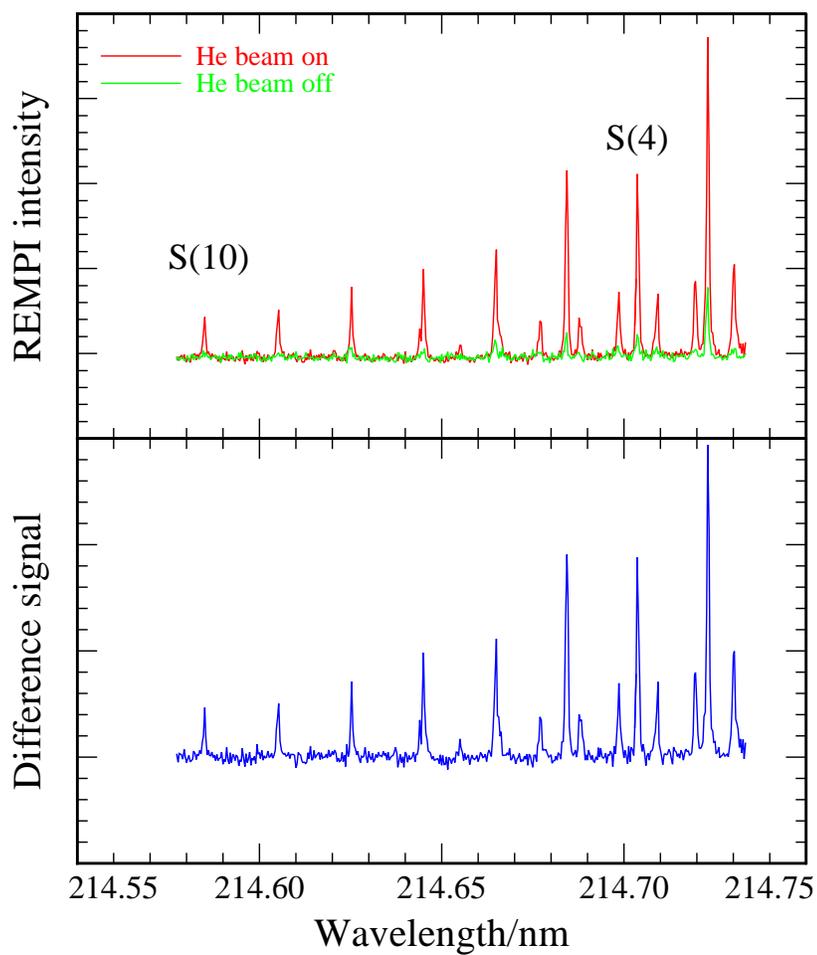
Figure 1.2: 2 + 1 resonance enhanced multiphoton ionization spectrum of CO rotationally excited by collisions with He atoms. The upper panel shows the spectra with and without the collisions with He atoms; the lower panel shows the difference.

1.2 Classical mechanics

I want to begin with some concepts from classical mechanics so you can appreciate the connections and differences with QM. I will do most of this in one dimension, studying a particle constrained to move along a line, to keep things simple.

In classical mechanics (CM) a particle's motion is described by a trajectory, a function of time: $x(t)$. At each instant of time, $x(t)$ is a number that specifies the particle's position on the x axis.

You generally don't start out knowing the entire $x(t)$ function. You start out knowing the particle's initial position $x(0)$ and initial velocity $v(0) = dx/dt(t = 0)$ (or momentum, $p(0) = mv(0)$). Then, if you know the *forces F* that act on the particle, you can use Newton's second law to write down the *equation of motion*:

$$F = ma = m\frac{d^2 x}{dt^2},\tag{1.1}$$

where m is the particle mass and a is the acceleration.

The forces typically depend on the position of the particle; occasionally they also depend on the velocity (friction is an example). The equation of motion is a differential equation, whose solutions are functions of time $x(t)$. The differential equation will have a whole family of solutions. Only one of them will match the initial conditions $x(0)$ and $v(0)$; that one is the particle's trajectory.

There is an alternative way to write the equations of motion. Instead of a single equation involving a second derivative, we can write two equations that involve only first derivatives. We do this by defining a new variable that is proportional to the first derivative of the original variable. In this case I will use the momentum $p = m\frac{dx}{dt}$. Then I have

$$\frac{dx}{dt} = \frac{p}{m}\tag{1.2}$$

$$\frac{dp}{dt} = F\tag{1.3}$$

where the first equation is the definition of p and the second comes from differentiating both sides of the first and then using Newton's second law. This pair of "first order" equations is exactly equivalent to the original equation of motion Eq. (1.1).

Another way to find the equations of motion, not usually taught in introductory mechanics courses, is applicable when the forces do not depend on the velocities. In this case you can write a function for the total energy in terms of

the position and momentum of the particle. For a single particle moving in a potential $V(x)$, that function, called the *classical Hamiltonian*, is

$$H(p, x) = T + V = \frac{p^2}{2m} + V(x),\tag{1.4}$$

where T is the usual notation for the kinetic energy. *Hamilton's equations* then give you a straightforward prescription for finding the equations of motion just by evaluating derivatives of $H(p, x)$ with respect to p and x.

In quantum mechanics, as you will see, the Hamiltonian plays a prominent role. Just as in the classical case, the equation of motion of a system—the equation that determines how it evolves in time—is determined by the Hamiltonian. In the quantum case, though, we cannot talk about a trajectory; the instantaneous description of the system is given not by a single set of coordinates but by a function that extends over all space. We will next examine why that must be so.

1.3 Wave nature of matter

1.3.1 Interference; Young's experiment

Two properties that were traditionally associated with *waves* are interference and diffraction. Interference is the tendency of two waves to add together constructively, to make a stronger wave, or destructively, to make a weaker one, depending on whether their peaks and troughs are aligned or not at the place of addition. Diffraction is the tendency of a single wave to "bend" around a corner, typically with some high-and-low-intensity variations in the bent part.

The experiment that (in hindsight) demonstrated wave properties of light was Thomas Young's 1801 "double slit" experiment. He used sunlight. The light passed through a small hole, then around a small card inserted edgewise into the beam. (A later version sent the light through a screen that has two parallel slits close together. That version is simpler to analyze and gives brighter fringes.) Beyond the slits was a screen. An alternating pattern of light and dark bands appears on the screen if both slits are open; see Figure 1.4.

The pattern arises from interference. If the light waves from the two slits arrive "in phase" at a particular spot (with their crests and troughs aligned), they add constructively and that spot is bright. If they arrive out of phase (one crest and one trough), they cancel and the spot is dark.

The difference has to do with the path length. For a spot on the screen displaced from the center line by an angle θ_m, the difference in distances traveled by waves from the two slits is $\Delta = a \sin \theta_m$, where a is the distance between the two slits. If that distance is exactly an integral number of wavelengths—that is, if

$a \sin \theta_m = n\lambda$, where n is an integer—the two waves will arrive at the screen in phase. If it is "$n + \frac{1}{2}$" wavelengths, they will arrive exactly out of phase and cancel.
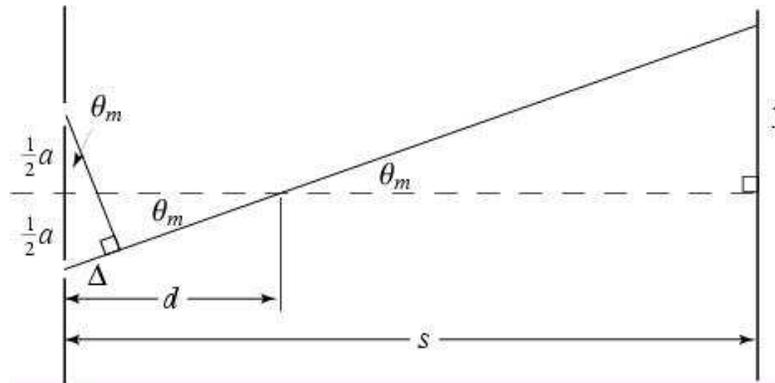

Figure 1.3: Variable definitions for Young's experiment (from `http://scienceworld.wolfram.com/physics/DoubleSlitInterference.html`).
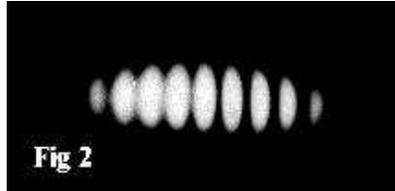
Figure 1.4: Fringe pattern produced by Young's experiment (from `http://www.cavendishscience.org/phys/tyoung/tyoung.htm`, originally published by Walter Scheider in *The Physics Teacher* **24**, 217-219 (1986).)

It is easy to estimate how far apart the fringes should be. If the distance to the screen is s, and we call the distance from the center line to the fringe in question y, then $\theta_m \approx \frac{y}{s}$. (See Figure 1.3.) For small θ_m, $\sin \theta_m \approx \theta_m$ (where θ_m must be measured in radians.) Then we have

$$\frac{y}{s} \approx \frac{n\lambda}{a} \tag{1.5}$$

$$y \approx \frac{n\lambda s}{a} \tag{1.6}$$

so the distance between fringes must be

$$(1.7)$$

$$y_{n+1} - y_n \approx \frac{(n+1)\lambda s}{a} - \frac{n\lambda s}{a} \tag{1.8}$$

$$= \frac{\lambda s}{a} \tag{1.9}$$

So, if you know a, this gives you a way to measure λ. Young estimated the average wavelength of sunlight this way and got very nearly the right answer.

1.3.2   Matter waves

Wave behavior for electrons   In 1924 de Broglie, trying to understand the origin   Levine §18.4
of Bohr's "quantization rules" that described the H atom, had suggested that
matter might have wavelike properties, and gave a formula for the wavelength:
$\lambda = \frac{h}{p}$, where h is the same Planck's constant that appears in the Einstein relation
$E = h\nu$. With this formula for the wavelength, and a very simple model, he was
able to reproduce Bohr's H-atom results.

In 1927, Davisson and Germer in the U.S. and G. P. Thomson in England
showed that electrons bounced off a crystal or a foil exhibited a diffraction pat-
tern. These experiments confirmed that material particles could have wavelike
properties. Figure 1.5 shows spots appearing on a phosphor screen when a beam
of electrons is sent at a glancing angle against a sapphire crystal, and the screen
is placed to intercept the reflected beam. (This technique, common in semicon-
ductor process control, is called "reflection high energy electron diffraction," or
RHEED.) The bright spots are a diffraction pattern produced by scattering of the
electrons from the regularly corrugated surface of the sapphire, and are a direct
consequence of the wave properties of electrons.

Wave behavior for helium atoms   In a beautiful experiment reported twenty
years ago (Schöllkopf and Toennies, *Science* **266**, 1345 (1994)), a group in Ger-
many sent a beam of helium atoms at a "diffraction grating" made with alter-
nating bars and slits each 100 nm wide. The de Broglie wavelength of atoms in
the beam from a 300 K source is 0.056 nm. Figure 1.6 describes the apparatus,
and Figure 1.7 shows the signal intensity as a function of deflection angle away
from the primary beam direction. As the mass spectrometer detector is moved
away from the straight-on position, a series of "bright" (many atoms detected per
second) and "dark" (almost no atoms detected per second) fringes appears. The
peaks appear just where a diffraction calculation would predict them. (Simple

particle behavior would give a single central peak that just decayed away rapidly at higher angles.) The bright fringes alternate between tall and short because of the way the slots and bars of the grating are spaced.

This is one of the best direct demonstrations of the wave behavior of neutral matter I know. (The authors were actually interested in the much weaker signals associated with the helium dimer, He_2, but the very careful apparatus design required to succeed at that task meant that their observations of the plain He atom diffraction are of excellent quality.)

Macroscopic behavior   Why don't we see interference and diffraction for ordinary objects, too? The answer lies in the values of the correponding de Broglie wavelengths. For example, let's calculate the de Broglie wavelength of a bug, mass 1 g, flying at 0.1 m/s. The momentum $p = mv = 10^{-4} \, \mathrm{kg\,m\,s^{-1}}$. The de Broglie wavelength is then $\lambda = h/p = 6.6 \times 10^{-34} \, \mathrm{J\,s} / 10^{-4} \, \mathrm{kg\,m\,s^{-1}} = 6.6 \times 10^{-30} \, \mathrm{m}$. That distance is small even compared to the sizes of atomic nuclei, so we cannot expect to observe wavelike properties for flying bugs. The same holds for other large things, where "large" applies to individual cells but not individual molecules. In general, you can expect to see wavelike effects if the de Broglie wavelength is comparable to or larger than the object itself or some feature of the apparatus used to measure its behavior. More fundamentally, we must begin to regard a particle not as a point mass located at a specific spot in space, but as a wave with crests and troughs, extended in space. That shift in description is the basis for the quantum mechanical *wavefunction*, which we will study next.
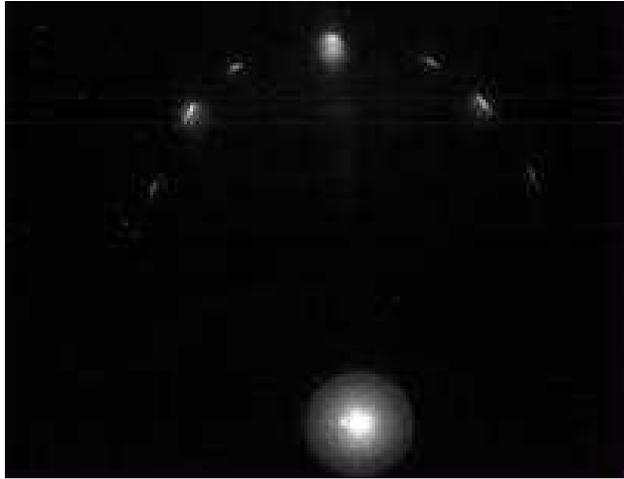
Figure 1.5: RHEED pattern from electrons on sapphire (from `http://mrlxp2.`
`mrl.uiuc.edu/~rheed/gallery/circles_rods.html`). The bright spot at the
bottom is the incident, unscattered beam; the spots at the top come from electron
diffraction.
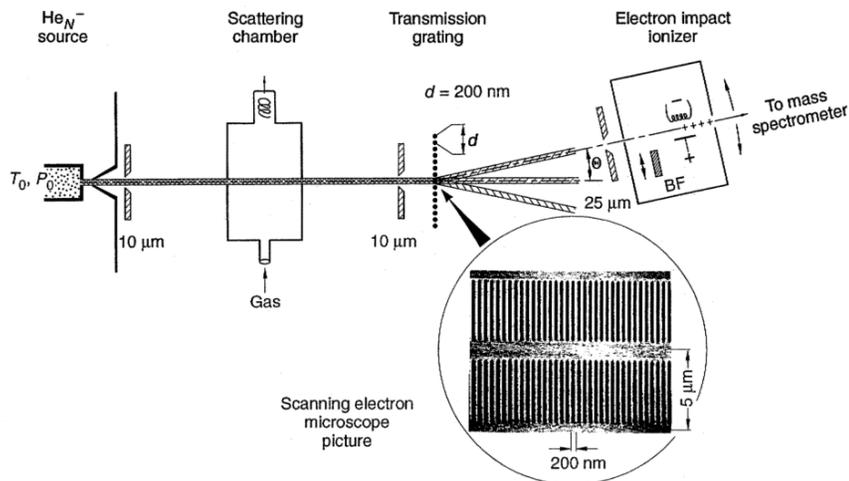
Fig. 1. Schematic diagram of the apparatus used for diffraction of the different components in a He
cluster beam. Only the vacuum system containing the ionizer is indicated. BF designates the beam flag
that intercepts the incident beam in the stagnation mode measurements. The inset (bottom right) shows
a scanning electron microscope picture of the grating used (10, 11).

Figure 1.6: Apparatus diagram from the Schöllkopf and Toennies paper.
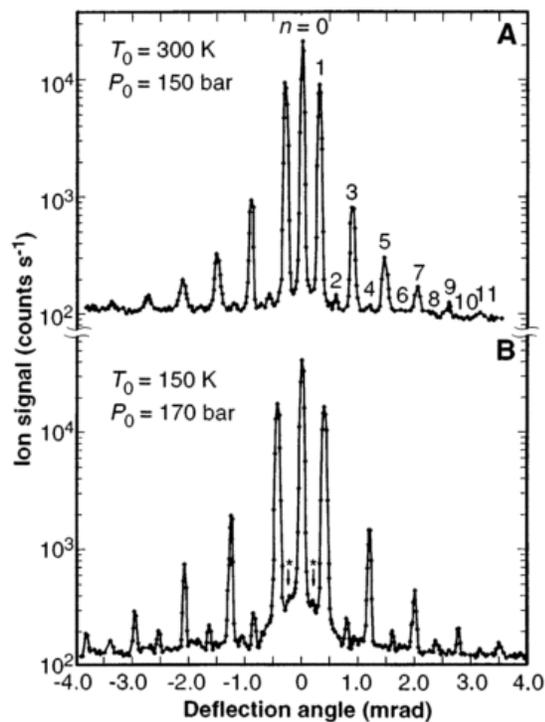
Fig. 2. Two angular distributions for He cluster beams, measured at $T_0 = 300$ K and $P_0 = 150$ bar (**A**) and at $T_0 = 150$ K and $P_0 = 170$ bar (**B**). Note that the logarithm of the He$^+$ signal measured in the direct in-flight mode is plotted versus the deflection angle. The numbers indicate the diffraction orders of the He atom diffraction peaks. The asterisks in (B) denote additional peaks attributable to first-order diffraction of He dimers.

Figure 1.7: Primary data figure from that paper. Because the He atoms were slower from the 150 K source, their de Broglie wavelength was longer and the diffraction peaks therefore are more widely spaced in the lower panel.

Chapter 2

Fundamentals of quantum mechanics

I now want to introduce the formal structure of quantum mechanics. I will do this by giving several *postulates.* These postulates have not been deduced from any earlier theory; they are the underlying tenents of the quantum theory and should be accepted only if predictions that follow from them agree with experimental observations. (So far, they have.) They play the role for QM that Newton's laws play for classical mechanics.

Two different notations can be used to state the rules of QM and do calculations. One uses differential equations and operators, functions of coordinates, and integrals; the other uses abstract operators, vectors, and inner products ("dot products"). The vector notation (called Dirac or bra-ket notation) is more compact and easier to manipulate, and several of the most important problems in QM can be solved easily with it. It is also more fundamental; some kinds of problems (notably problems involving *spin*) do not have an equivalent expression in function notation. The function notation is more familiar and is sometimes more convenient when it comes to actually computing a numerical result for a particular problem.

I will try to give the postulates and examples in both notations, and show you the relation between them. We will use both notations during the course. Often I will develop an idea or a particular problem in the vector notation and then at the end convert the vector expression into an integral to get a numerical result. Sometimes even this final conversion will be unnecessary.

In hopes of providing motivation, I will state each postulate first, in a precise language. Then, I will define and explain the unfamiliar terms that appear in the postulate, and try to show you where they appear in the theory.

2.1 State vectors and wavefunctions

Postulate 1 At a particular time t, the state of a physical system is described by a *wavefunction* $\psi(\mathbf{r})$, which is a function of the coordinates of all the particles in the system. $\psi(\mathbf{r})$ must be continuous and single-valued everywhere in space, and it must be square-integrable: $\int d\mathbf{r}\,\psi^*(\mathbf{r})\psi(\mathbf{r})$ must be finite and nonzero.

In Dirac notation, the state is described by a *ket* or *state vector*, written $|\psi\rangle$. The inner product of $|\psi\rangle$ with itself, $\langle\psi|\psi\rangle$, must be finite and nonzero.

2.1.1 Comments and definitions

Basic content of wavefunction As we discussed earlier, while a (classical) particle has a definite position at any specific time, a wave does not: it is spread out over space, though it might be small in some places and large in others. To describe a wave, then, we need a function of position. The wavefunction plays that role in quantum mechanics. Essentially it tells you where you might find the particle should you look for it. If the wavefunction is small in a particular spot, the particle probably won't be there; if it's large in another place, you're more likely to find the particle if you look in that place.

Table 2.1 shows some acceptable wavefunctions. All of them meet the criteria given in the postulate, but some have additional characteristics that could only appear for special kinds of problems. $\psi = $ constant is acceptable in the last example because of the finite limits of integration; if the limits extended to infinity, the integral described in the postulate would be infinite. In the second example, the wavefunction has "corners" (discontinuous derivatives) at 0 and l; that can happen only when the potential is infinite at those points. We will not worry too much about the additional conditions that apply in those cases.

Normalization Notice that if you have an acceptable wavefunction, you can multiply it by any nonzero constant and the result will also be acceptable. It turns out (as you will see in Postulate 3) that multiplying the wavefunction by a constant does not change its physical meaning; no predicted experimental observations change. So we can scale a given wavefunction to give it a convenient overall height or area. Often it is convenient to require that the area under the *square* of the wavefunction should be 1. That means

$$\langle\psi|\psi\rangle = \int d\mathbf{r}\left|\psi(\mathbf{r})\right|^2 = \int d\mathbf{r}\,\psi^*(\mathbf{r})\psi(\mathbf{r}) = 1. \tag{2.1}$$

When this is true we say that the wavefunction has been *normalized.* In the vector notation this requirement is equivalent to saying we require that the length of the state vector be 1.

Table 2.1 Some examples of wavefunctions.

One dimensional examples, on $-\infty \leq x \leq \infty$:

$$\psi(x) = e^{-ax^2}$$
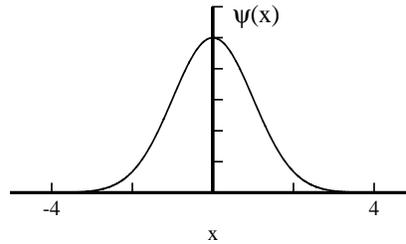
$$\psi(x) = \begin{cases} \sin\left(\frac{3\pi x}{l}\right) & \text{for } 0 \leq x \leq l, \\ 0 & \text{otherwise.} \end{cases}$$
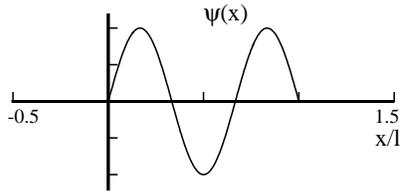
$$\psi(x) = \sin(5x)\,e^{-(x-3)^2/2}$$
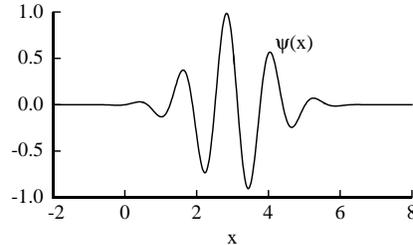
Two dimensional example, on $-\infty \leq x, y \leq \infty$:

$$\psi(x, y) = e^{-x^2 - y^2}$$
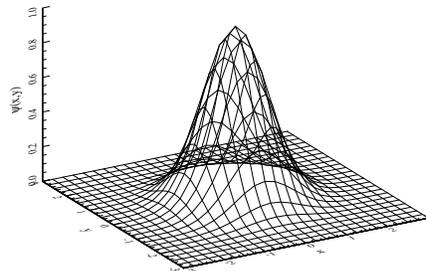
Two dimensional examples, on $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$:
$$\psi(\theta, \phi) = \tfrac{1}{2} e^{3i\phi}(3\cos^2\theta - 1)$$
$$\psi(\theta, \phi) = 1$$

If you already have a wavefunction $\psi(\mathbf{r})$, it is straightforward to make a new one $\psi_N(\mathbf{r})$ that is normalized. You must find a constant N such that $\psi_N(\mathbf{r}) = N\psi(\mathbf{r})$ is normalized. In a generic sense we can write

$$\psi_N(\mathbf{r}) = \frac{1}{\sqrt{\int d\mathbf{r}\,\psi^*(\mathbf{r})\psi(\mathbf{r})}}\,\psi(\mathbf{r}) \tag{2.2}$$

or

$$\left|\psi_N\right\rangle = \frac{1}{\sqrt{\langle\psi|\psi\rangle}}\,\left|\psi\right\rangle. \tag{2.3}$$

The fraction multiplying $\left|\psi\right\rangle$ is called a *normalization constant,* often written N. Sometimes it is convenient to use this formula for N; sometimes it is more straightforward to think about finding N directly as in the examples below.

Normalization examples

1. One physical system that is useful for thinking about conjugated ring systems is a "particle on a ring", a single particle (possibly an electron) that is free to move on a circle. If we use the polar angle θ to describe its position, then we must have $0 \le \theta \le 2\pi$, and that is the range of integration appropriate for normalization.

   Let's say we have an unnormalized wavefunction $\psi(\theta) = 3$. (This is about as simple a wavefunction as possible: it doesn't depend on θ at all!) We want to find a normalized version $\psi_N(\theta) = N\psi(\theta)$ of this wavefunction. We write

   $$\int_0^{2\pi} (N\psi(\theta))^2\, d\theta = 1. \tag{2.4}$$

   Since $\psi(\theta) = 3$ this reduces to

   $$\int_0^{2\pi} (3N)^2\, d\theta = 1 \tag{2.5}$$

and the constant $(3N)^2$ can be factored out of the integral. We then have

$$(3N)^2 \int_0^{2\pi} d\theta = 1 \tag{2.6}$$

$$(3N)^2 [\theta]_0^{2\pi} = 1 \tag{2.7}$$

$$9N^2 [2\pi - 0] = 1 \tag{2.8}$$

$$9N^2 \cdot 2\pi = 1 \tag{2.9}$$

$$N^2 = \frac{1}{18\pi} \tag{2.10}$$

$$N = \left(\frac{1}{18\pi}\right)^{\frac{1}{2}} = \frac{1}{3}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}. \tag{2.11}$$

We find our normalized wavefunction $\psi_N(\theta)$ by multiplying the original, unnormalized function by N, obtaining

$$\psi_N(\theta) = \frac{1}{3}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \cdot 3 = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}}. \tag{2.12}$$

2. For a more complicated example, let's say that a physically appropriate wavefunction for your problem has the form

$$\psi(x) = xe^{-\alpha x^2/2} \tag{2.13}$$

over the range $-\infty \leq x \leq \infty$. This wavefunction describes the nuclear motion of a diatomic molecule after it absorbs an infrared photon. It's shown in the margin, for the value $\alpha = 2$.

We find the normalization constant just as before, though our integral is more complicated.

$$\psi_N(x) = N\psi(x) = Nxe^{-\alpha x^2/2} \tag{2.14}$$

$$\int_{-\infty}^{\infty} \psi_N^*(x)\psi_N(x)\,dx = 1 \tag{2.15}$$

Since this wavefunction (like that of the previous example) is real, multiplying it by its complex conjugate is the same as just squaring it. We therefore have

$$N^2 \int_{-\infty}^{\infty} x^2 e^{-\alpha x^2}\,dx = 1. \tag{2.16}$$
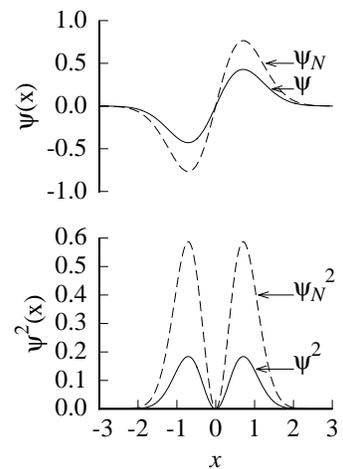
You can do that integral by parts, reducing it to a piece that looks like $\int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}$, but it's much easier to use integration by tables. In Table 15.1 of your textbook you'll find the standard definite integral

$$\int_0^\infty x^{2n} e^{-ax^2}\, dx = \frac{(2n)!\pi^{\frac{1}{2}}}{2^{2n+1} n! a^{n+\frac{1}{2}}}. \tag{2.17}$$

Since our integral goes from $-\infty$ to ∞ rather than from 0 to ∞, and the integrand is even, we need to multiply the standard integral by 2:

$$\int_{-\infty}^\infty x^{2n} e^{-ax^2}\, dx = 2\frac{(2n)!\pi^{\frac{1}{2}}}{2^{2n+1} n! a^{n+\frac{1}{2}}}. \tag{2.18}$$

Substituting that into Eq. (2.16) gives in our case ($n = 1$, $a = \alpha$)

$$2N^2 \frac{2!\pi^{\frac{1}{2}}}{8\alpha^{\frac{3}{2}}} = 1. \tag{2.19}$$

Now we can solve for N, being careful about exponents.

$$N^2 = \frac{2\alpha^{\frac{3}{2}}}{\pi^{\frac{1}{2}}} = \left(\frac{4\alpha^3}{\pi}\right)^{\frac{1}{2}} \tag{2.20}$$

$$N = \left(\frac{4\alpha^3}{\pi}\right)^{\frac{1}{4}} \tag{2.21}$$

and our normalized wavefunction, finally, is

$$\psi_N(x) = \left(\frac{4\alpha^3}{\pi}\right)^{\frac{1}{4}} x e^{-\alpha x^2}. \tag{2.22}$$

The area under $\psi_N^2(x)$ is now 1. You can see from the figure in the margin that the original wavefunction was "too small", so that N had to be bigger than 1 to normalize the function.

Usually, I will not use the subscript N to indicate that a particular wavefunction is normalized. I will just state whether it is or not. If I don't tell you, and you need to know, you can check by integrating $|\psi|^2$ to see whether the result is 1 or not.

Inner product in two notations   Notice how I wrote the denominator in the vector notation in Eq. (2.3): $\langle \psi | \psi \rangle$. That is an important special case of the *inner product* in the vector notation. In function notation the corresponding operation is defined by an integral:

$$\langle \psi | \phi \rangle = \int \psi^*(\mathbf{r}) \phi(\mathbf{r}) \, d\mathbf{r} \tag{2.23}$$

where the integration extends over all space. (For instance, for a particle constrained to move in one dimension, you have $\int_{-\infty}^{\infty} dx$). This definition gives the connection between the two notations.

In fact the symbol $\langle \phi |$, pronounced "bra phi", is an *operator* (a set of instructions) which means "multiply on the left by $\phi^*(x)$ and integrate over all space." Anything in $|\rangle$ is called a *ket*; anything in $\langle |$ is called a *bra*. Put together they make a *bracket*, a number that does not depend on the spatial variables.

Born interpretation   The wavefunction contains all the information that *can* Levine §18.6
be known about the physical system. Once you know it, you can predict the outcome of any experiment performed on that system. Further postulates will outline how to do that. But one particular case is important enough to mention right away: if the wavefunction is normalized, the probability density for the position of a particle is given by $P(x) = \psi^*(x)\psi(x)$. That means that if a particle's *normalized* wavefunction is $\psi(x)$, then if you measure the position of the particle, the probability of finding the particle between positions a and b is

$$P(a \le x \le b) = \int_a^b \psi^*(x)\psi(x) \, dx. \tag{2.24}$$

This rule is the precise statement of my earlier comment about "where the wavefunction is large, you have a large probability of finding the particle." The normalization requirement simply means that the probability of finding the particle *somewhere* in space must be 1: you are sure to find it if you look everywhere.

Figure 2.1 shows an example. In that case, I have used the wavefunction you tested in the first homework, and we normalized above: $\psi(x) = Nxe^{-\alpha x^2/2}$. The upper panel shows the wavefunction itself. The middle panel shows the square of the wavefunction, the probability density function. This function is normalized (that is, N was chosen by the prescription above), so that the total area under the probability density function is 1. The lowest panel gives the results of a series of experiments. Forty quantum mechanical particles were prepared with exactly the wavefunction $\psi(x)$. Then, on each system, I measured the position of the particle. Each measurement found the particle in a different position (even though the

particles were all *prepared* identically!) At each position where I found a particle, I plotted a dot. You can see that most of the dots appear where the probability density function is big, but it is possible to find a particle anywhere the probability density function is not zero. The only place it is truly zero is right at $x = 0$ where the wavefunction crosses the axis. It is extremely unlikely, though, that you will find the particle anywhere the probability density is very small.
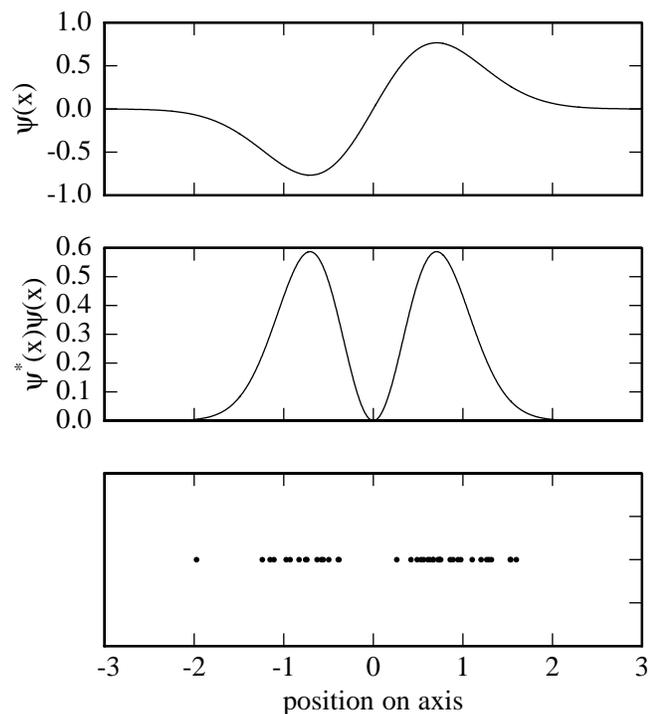
Figure 2.1: The Born interpretation of the wavefunction. Upper panel shows wavefunction; middle panel shows probability density function; bottom panel shows results of 40 measurements of position on particles prepared with this wavefunction.

Superpositions   Any linear combination of acceptable wavefunctions will also be an acceptable wavefunction. That is, if $\psi(x)$ and $\phi(x)$ are both acceptable wavefunctions, then $\lambda_1 \psi(x) + \lambda_2 \phi(x)$ will also be acceptable, where λ_1 and λ_2 are complex numbers. In Dirac notation, we say that the vectors $|\psi\rangle$ and $|\phi\rangle$ are members of a *vector space*.

This ability to form sums (or *superposition states*) of different wavefunctions provides some of the "mystery" of quantum mechanics, since you can add together wavefunctions that correspond to quite different states. The most famous example is "Schrödinger's cat", which describes a cat inside a sealed box whose wavefunction describes him as partially alive and partially dead. We will see several examples where superposition is a powerful physical tool, and a few where it is necessary to get correct predictions of experimental results.

2.2 Operators

Postulate 2 Every measurable physical quantity (position, momentum, energy, angular momentum, polarization, etc.) is described by an *operator*. This operator is Hermitian and its eigenfunctions form a complete set.

2.2.1 Comments and definitions

Physical quantities The word "every", which appears in this postulate in most QM texts, is misleading. Two important quantities that are definitely measurable in the lab, namely mass and time, do not have corresponding operators in ordinary nonrelativistic quantum mechanics. Instead they are regarded as "parameters", quantities whose values are specified as part of the problem to be solved. It might be better to write "Every measurable physical quantity that can be expressed in terms of positions and momenta". For our purposes that would work.

Definition of operators An *operator* is a mathematical object that specifies a list of instructions to be applied to a function; applying the operator (carrying out the instructions) usually turns one function into another. In the vector world, an operator is an object that turns one vector into another, often by multiplying the vector on the left by a matrix.

Some examples of operators are these:

Multiply by a constant:	$\hat{A}f(x) = 3f(x)$
Multiply by x:	$\hat{A}f(x) = xf(x)$
Differentiate:	$\hat{A}f(x) = \frac{d}{dx}f(x)$
"Parity":	$\hat{A}f(x) = f(-x)$

Properties of operator math Operators can be added together and multiplied by constants:

$$(\hat{A} + \hat{B})f(x) = \hat{A}f(x) + \hat{B}f(x) \tag{2.25}$$

$$(\lambda \hat{A})f(x) = \lambda(\hat{A}f(x)) \tag{2.26}$$

Addition of operators is commutative; that is, $\hat{A} + \hat{B} = \hat{B} + \hat{A}$. Also, operators important in quantum mechanics are always linear operators, which means that $\hat{A}(f(x) + g(x)) = \hat{A}f(x) + \hat{A}g(x)$ and $\hat{A}(\lambda f(x)) = \lambda \hat{A}f(x)$.

However, *the order of application of operators is important*; operating on a function first by $\hat{A}$ and then by $\hat{B}$ does not give the same result as operating first

by $\hat{B}$ and then by $\hat{A}$. For instance, consider

$$\hat{A} = x, \tag{2.27}$$

$$\hat{B} = \frac{d}{dx}. \tag{2.28}$$

Then

$$\hat{A}\hat{B}f(x) = x\frac{d}{dx}f(x). \tag{2.29}$$

But

$$\hat{B}\hat{A}f(x) = \frac{d}{dx}\left[xf(x)\right] \tag{2.30}$$

$$= x\frac{d}{dx}f(x) + f(x)\frac{d}{dx}x \tag{2.31}$$

$$= x\frac{d}{dx}f(x) + f(x), \tag{2.32}$$

which is not the same as the first case; it has an extra term, $f(x)$, added on. In other words, multiplication of operators is not commutative. Therefore, when you manipulate operators, you must be careful to preserve their orders. For example,

$$(\hat{A} + \hat{B})^2 = (\hat{A} + \hat{B})(\hat{A} + \hat{B}) \tag{2.33}$$

$$= \hat{A}^2 + \hat{A}\hat{B} + \hat{B}\hat{A} + \hat{B}^2 \tag{2.34}$$

$$\neq \hat{A}^2 + 2\hat{A}\hat{B} + \hat{B}^2 \tag{2.35}$$

Commutators   The lack of commutativity of operator multiplication is so im-   Levine §18.11
portant in quantum mechanics that there is a special term for the difference between the operator product taken in the two orders. The *commutator* of two operators $\hat{A}$ and $\hat{B}$ is written $[\hat{A}, \hat{B}]$ and is defined by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{2.36}$$

The commutator is itself an operator (often a very simple one, such as a constant.) When you want to work out the commutator of two operators, it usually helps to give them a "test function" $f(x)$, to help you keep track of things (though you never actually apply the operators to $f(x)$.) For example, with the two operators

$\hat{A} = x$, $\hat{B} = \frac{d}{dx}$ we used above, we would calculate $[\hat{A}, \hat{B}]$ as follows:

$$[\hat{A}, \hat{B}] f(x) = \hat{A}\hat{B} f(x) - \hat{B}\hat{A} f(x) \tag{2.37}$$

$$= x\frac{d}{dx} f(x) - \left(x\frac{d}{dx} f(x) + f(x) \right) \tag{2.38}$$

$$= (-1) f(x) \tag{2.39}$$

So we would say

$$[\hat{A}, \hat{B}] = -1. \tag{2.40}$$

Pairs of operators for whom order does not matter—that is, pairs of operators whose commutator is 0—are called "commuting operators"; we say that $\hat{A}$ and $\hat{B}$ *commute* if $[\hat{A}, \hat{B}] = 0$. The following properties are straightforward:

1. Any operator commutes with itself, and with powers of itself: $[\hat{A}, \hat{A}] = 0$.

2. Reversing the order of operators in a commutator changes the sign of the commutator: $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$.

3. Every operator commutes with any other operator that works on different variables. For example, $[\hat{x}, \frac{\hat{d}}{dy}] = 0$.

4. Linear operators (the only ones important in QM) commute with constants, as shown above. An example of an operator that is not linear, and therefore does not commute with constants, is the square root operator. $\sqrt{af(x)} = \sqrt{a}\sqrt{f(x)} \neq a\sqrt{f(x)}$.

Levine §18.11

The $\hat{x}$ and $\hat{p}$ operators   In QM the operator corresponding to the position of a particle (in one dimension) is the operator that multiplies by x:

$$\hat{x} f(x) = x f(x) \tag{2.41}$$

The operator corresponding to the momentum is $-i\hbar\frac{d}{dx}$:

$$\hat{p} f(x) = -i\hbar\frac{d}{dx} f(x). \tag{2.42}$$

The most important commutator in QM is

$$[\hat{x}, \hat{p}] = i\hbar \tag{2.43}$$

Let me show this explicitly:

$$[\hat{x}, \hat{p}]f(x) = (\hat{x}\hat{p} - \hat{p}\hat{x})f(x) \tag{2.44}$$

$$= x\left(-i\hbar\frac{d}{dx}f(x)\right) - \left(-i\hbar\frac{d}{dx}(xf(x))\right) \tag{2.45}$$

$$= -i\hbar x\frac{d}{dx}f(x) + i\hbar\left(x\frac{d}{dx}f(x) + f(x)\right) \tag{2.46}$$

$$= \left(-i\hbar x\frac{d}{dx} + i\hbar x\frac{d}{dx} + i\hbar\right)f(x) \tag{2.47}$$

$$= i\hbar f(x). \tag{2.48}$$

So, we write

$$[\hat{x}, \hat{p}] = i\hbar. \tag{2.49}$$

Eigenfunctions and eigenvalues   Usually, when you apply an operator to a function, you get some other function. But for some operators there is a special set of functions for which applying the operator just ends up multiplying the function by a constant. Such a function is called an *eigenfunction* of the operator, and the constant is the corresponding *eigenvalue*.

For instance, consider the operator $\hat{A} = \frac{d^2}{dx^2}$. Then

$$\hat{A}\left(3x^4\right) = \frac{d^2}{dx^2}\left(3x^4\right) \tag{2.50}$$

$$= \frac{d}{dx}\left(12x^3\right) \tag{2.51}$$

$$= 36x^2, \tag{2.52}$$

and

$$\hat{A}(e^{4x}) = \frac{d^2}{dx^2}\left(e^{4x}\right) \tag{2.53}$$

$$= \frac{d}{dx}\left(4e^{4x}\right) \tag{2.54}$$

$$= 16e^{4x}. \tag{2.55}$$

So $3x^4$ is not an eigenfunction of $\hat{A}$, but e^{4x} is an eigenfunction with eigenvalue 16.

The general definition of an eigenfunction of some operator $\hat{A}$ is any function $f(x)$ for which

$$\hat{A}f(x) = \lambda f(x) \tag{2.56}$$

where λ (the eigenvalue) is a constant.

If $\hat{A}$ is a linear operator, and $f(x)$ is an eigenfunction with eigenvalue λ, then any multiple of $f(x)$ is also an eigenfunction of $\hat{A}$ with the same eigenvalue. This statement is easy to prove, because linear operators by definition commute with constants. Let the constant multiplier be α:

$$\hat{A}(\alpha f(x)) = \alpha \hat{A} f(x) \tag{2.57}$$

$$= \alpha \lambda f(x) \tag{2.58}$$

$$= \lambda \alpha f(x) \tag{2.59}$$

$$= \lambda(\alpha f(x)) \qquad \text{QED.} \tag{2.60}$$

Degeneracy Operators usually have many different eigenfunctions; each one has its own eigenvalue. It sometimes happens that two or more different eigenfunctions will have the same eigenvalue; such eigenfunctions are called *degenerate*. The *degeneracy* is the number of eigenfunctions sharing the same eigenvalue. An eigenfunction whose eigenvalue is not shared by any other eigenfunction is called *nondegenerate*.

One-dimensional systems do not usually have any degeneracy. An operator that has degeneracy is the energy operator for the "two-dimensional particle in a box", which is

$$\hat{T} = -\frac{\hbar^2}{2\mu}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2}\right), \tag{2.61}$$

where μ is the mass of the particle and the boundary conditions require the eigenfunctions to be zero outside the square $0 \le x \le a$, $0 \le y \le a$. The eigenfunctions are

$$\phi_{m,n}(x,y) = \frac{2}{a}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{a}, \tag{2.62}$$

where n and m must be positive integers. The eigenvalues are (try it!)

$$E_{n,m} = \frac{(m^2 + n^2)\hbar^2}{4\mu a^2} \tag{2.63}$$

The lowest possible eigenvalue is the one with $m = n = 1$, whose value is $E_{1,1} = \frac{\hbar^2}{2\mu a^2}$. No other eigenfunctions have that same eigenvalue, so we say that that eigenvalue is nondegenerate. But the next higher eigenvalue, $\frac{5\hbar^2}{4\mu a^2}$, has two eigenfunctions that give it: one with $m = 1, n = 2$, and the other with $m = 2, n = 1$. That eigenvalue is therefore *doubly degenerate*. We might also say that its degeneracy is 2.

Hermitian operators $\hat{A}$ is a Hermitian operator if for any functions $\psi(x)$ and $\phi(x)$,

$$\int dx \phi(x)^* \hat{A} \psi(x) = \left(\int dx \psi(x)^* \hat{A} \phi(x) \right)^*,$$ (2.64)

or, in the more compact Dirac notation,   Levine §18.16

$$\langle \phi | A | \psi \rangle = \langle \psi | A | \phi \rangle^*.$$ (2.65)

Notice that if you let the two functions ϕ and ψ be the same in Eq. (2.65), you get

$$\langle \phi | \hat{A} | \phi \rangle = \langle \phi | \hat{A} | \phi \rangle^*.$$ (2.66)

A number can only be equal to its complex conjugate if it is real; therefore, for a Hermitian operator, expressions of the type $\langle \phi | \hat{A} | \phi \rangle$ are always real.

Hermitian operators are the analog of *symmetric* matrices in linear algebra, and share several of their important properties.

Real eigenvalues   The eigenvalues of a Hermitian operator are all real:

$$\hat{A} | \psi \rangle = \lambda | \psi \rangle$$ (2.67)

Act on left with $\langle \psi |$:

$$\langle \psi | \hat{A} | \psi \rangle = \langle \psi | \lambda | \psi \rangle$$ (2.68)

$$\langle \psi | \hat{A} | \psi \rangle = \lambda \langle \psi | \psi \rangle$$ (2.69)

Since $\langle \psi | \hat{A} | \psi \rangle$ and $\langle \psi | \psi \rangle$ are both real, λ must be real. QED.

It is important that Hermitian operators have real eigenvalues because (as you will see in Postulate 3) the results of measurements are always the eigenvalues of some Hermitian operator. A physical property measurement must always give a real number, and Hermitian operators are guaranteed to have real eigenvalues.

Acting to the left   In Dirac notation Hermitian operators have a handy property: they can act "backwards" on their eigenvectors. If

$$\hat{A} | \phi \rangle = \lambda | \phi \rangle,$$ (2.70)

then

$$\langle \phi | \hat{A} | \psi \rangle = \langle \psi | \hat{A} | \phi \rangle^*$$ (2.71)

$$= \langle \psi | \lambda | \phi \rangle^*$$ (2.72)

$$= \lambda^* \langle \psi | \phi \rangle^*$$ (2.73)

$$= \lambda \langle \psi | \phi \rangle^*$$ (2.74)

$$= \lambda \langle \phi | \psi \rangle$$ (2.75)

Orthogonal eigenvectors The word "orthogonal" in ordinary Euclidean space means perpendicular. The dot product of two orthogonal vectors is zero. In quantum mechanics the corresponding operation to the dot product is the "inner product" we have seen before. Two functions $\phi(x)$ and $\psi(x)$ are orthogonal if

$$\langle \phi | \psi \rangle = \int \phi^*(x)\psi(x)\,dx = 0. \tag{2.76}$$

The eigenvectors of a Hermitian operator corresponding to different eigenvalues are orthogonal. Let's prove that explicitly, not because the proof is particularly important but because it demonstrates techniques used often in quantum mechanical arguments. We have, by hypothesis,

$$\hat{A}|\phi_1\rangle = \lambda_1|\phi_1\rangle, \tag{2.77}$$

$$\hat{A}|\phi_2\rangle = \lambda_2|\phi_2\rangle, \tag{2.78}$$

with $\lambda_1 \neq \lambda_2$.

Now if we act on the left of Eq. (2.77) with $\langle \phi_2|$, we get

$$\langle \phi_2|\hat{A}|\phi_1\rangle = \langle \phi_2|\lambda_1|\phi_1\rangle = \lambda_1\langle \phi_2|\phi_1\rangle. \tag{2.79}$$

Since $\hat{A}$ is Hermitian, on the left side we can do

$$\langle \phi_2|\hat{A}|\phi_1\rangle = \lambda_2\langle \phi_2|\phi_1\rangle \tag{2.80}$$

to give

$$\lambda_2\langle \phi_2|\phi_1\rangle = \lambda_1\langle \phi_2|\phi_1\rangle. \tag{2.81}$$

Subtract the right-hand side from both sides:

$$\lambda_2\langle \phi_2|\phi_1\rangle - \lambda_1\langle \phi_2|\phi_1\rangle = 0 \tag{2.82}$$

Now factor out $\langle \phi_2|\phi_1\rangle$:

$$(\lambda_2 - \lambda_1)\langle \phi_2|\phi_1\rangle = 0 \tag{2.83}$$

Since $\lambda_1 \neq \lambda_2$, we can divide both sides by $(\lambda_2 - \lambda_1)$ to find

$$\langle \phi_2|\phi_1\rangle = 0 \tag{2.84}$$

so we have proved that the two eigenfunctions are orthogonal.

Completeness   A particular set of functions is called *complete* if it is possible to write any function depending on the same variables and satisfying the same boundary conditions as a linear combination of the functions in the set. A theorem from functional analysis (which I will not prove) says the eigenfunctions of a Hermitian operator with a finite number of linearly independent eigenfunctions always form such a set. When there is an infinite number of independent eigenfunctions (common in QM!), this is not always true mathematically; however, for operators corresponding to physical quantities, Postulate 2 says the eigenfunctions will form a complete set.

The idea of a complete set (or "basis") is easy to understand in ordinary 3-dimensional space. Any position can be represented as a vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, where the three components give the x, y, and z coordinates. All such vectors can be written as linear combinations of the three "basis vectors" $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. For example,

$$\begin{pmatrix} 18 \\ -4 \\ 6 \end{pmatrix} = 18 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + -4 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.85}$$

Note that these basis vectors are orthogonal,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0, \tag{2.86}$$

and normalized,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1. \tag{2.87}$$

We say they form an *orthonormal basis*.

With an orthonormal basis it is very easy to figure out what the linear coefficients are: just take the dot product of your vector with each of the basis vectors. $\begin{pmatrix} 18 \\ -4 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 18$, and so on.

Let's look at using these complete sets of functions in the more general, quantum mechanical case. We have a complete set of kets, $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and so on, that are orthogonal and normalized. That means we can write any ket $|\psi\rangle$ as a linear combination of the basis kets:

$$|\psi\rangle = c_1 |\phi_1\rangle + c_2 |\phi_2\rangle + c_3 |\phi_3\rangle + \cdots, \tag{2.88}$$

where the c_n are complex numbers. To find any one of the c_n (let's do c_2), act on the left of both sides of the equation with $\langle\phi_n|$:

$$\langle\phi_2|\psi\rangle = c_1 \langle\phi_2|\phi_1\rangle + c_2 \langle\phi_2|\phi_2\rangle + c_3 \langle\phi_2|\phi_3\rangle + \cdots \tag{2.89}$$

Now, because the different $|\phi_n\rangle$ are orthogonal, all the terms on the right are zero except for the one containing $\langle\phi_2|\phi_2\rangle$. And because $|\phi_2\rangle$ is normalized, $\langle\phi_2|\phi_2\rangle = 1$. So we end up with simply

$$c_2 = \langle\phi_2|\psi\rangle; \tag{2.90}$$

that is, we find the expansion coefficients exactly the same way as in the 3-dimensional case.

We're ready to move on to Postulate 3: measurements!

2.3 Measurements

The next three postulates concern the possible results of measurements, how you predict them given the wavefunction of the system, and how the measurements affect the system itself. Another essential difference between quantum and classical mechanics will appear here: in most cases, even if the state of the system is known as well as it can be, the outcome of a particular measurement cannot be predicted with certainty. QM provides only the list of possible outcomes and the probability of each. In other words, on two identically prepared systems, two identical measurements might not give the same result.

2.3.1 Possible measurement results

Postulate 3 The result of an accurate measurement of a physical property of a system is always one of the eigenvalues of the operator corresponding to that physical property. The only eigenvalues that appear are associated with eigenfunctions that are acceptable wavefunctions.   Levine §18.11

   This postulate shows why it's important that our physical property operators be Hermitian: if a measurement is going to give an eigenvalue as its result, that eigenvalue had better be a real number. Only Hermitian operators are guaranteed to have real eigenvalues.

Quantization

The requirement that observed eigenvalues correspond to acceptable wavefunctions imposes an important restriction. Many important operators—most notably those for energy and angular momentum— have eigenfunctions corresponding to any eigenvalue at all. But, only the eigenfunctions corresponding to particular values of the eigenvalues turn out to be acceptable (square-integrable, continuous, and so on). So only those *discrete* eigenvalues will ever be observed in an experiment, and values lying in between those special ones will never be observed. This is the origin of the idea of quantization.   Levine §18.7

   Let's use the one-dimensional harmonic oscillator as an example. For that problem the kinetic energy operator is

$$\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{2.91}$$

and the potential energy operator is

$$\hat{V} = \frac{1}{2}m\omega^2\hat{x}^2 = \frac{1}{2}m\omega^2 x^2 \tag{2.92}$$

so the total energy or Hamiltonian operator is

$$\hat{H} = \hat{T} + \hat{V} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \qquad (2.93)$$

Figure 2.2 shows three eigenfunctions of the harmonic oscillator Hamiltonian. The upper one has an energy just less than $4\frac{1}{2}\hbar\omega$; the second one has an energy just over $4\frac{1}{2}\hbar\omega$. You can see that neither of those two eigenfunctions is acceptable, because they both diverge (and keep going up or down to infinity) at the right side of the diagram; neither of them is square-integrable. The last eigenfunction has an eigenvalue of exactly $4\frac{1}{2}\hbar\omega$. It goes smoothly to zero at the right side of the diagram, and in fact approaches the x-axis exponentially as x increases. It is square-integrable.

For the harmonic oscillator, only those eigenfunctions whose eigenvalues are $(n+\frac{1}{2})\hbar\omega$, where n is a nonnegative integer, are square-integrable. Therefore, when the total energy of any physical harmonic oscillator is measured, the result is always one of those eigenvalues, and not ever anything in between. Thus we say that the energies of the harmonic oscillator are *quantized*, and that its allowed energy levels are described by the formula

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega. \qquad (2.94)$$

The whole set of permissible eigenvalues for an operator is called its *eigenvalue spectrum*.

In Problem Sets 2 and 3, you showed that several of the acceptable eigenfunctions of the harmonic oscillator Hamiltonian (energy operator) were orthogonal and normalized. (Those functions are the ones that contain the Gaussian function $e^{-\alpha x^2/2}$.)

Many books do not explicitly point out the important role of the requirement that eigenvalues that are measurement results must correspond to physically acceptable eigenfunctions. They just presume that we will consider only the acceptable ones, and never mention the existence of the others. But it is the combination of the operator itself, and the conditions for acceptability of its eigenfunctions, that limits many measurement results to a set of discrete values.

Continuous eigenvalue spectra

Not all physical quantities are quantized, so not all operators have discrete eigenvalues. Some operators have acceptable eigenfunctions for any eigenvalues at all. In particular, the position and momentum operators behave that way.
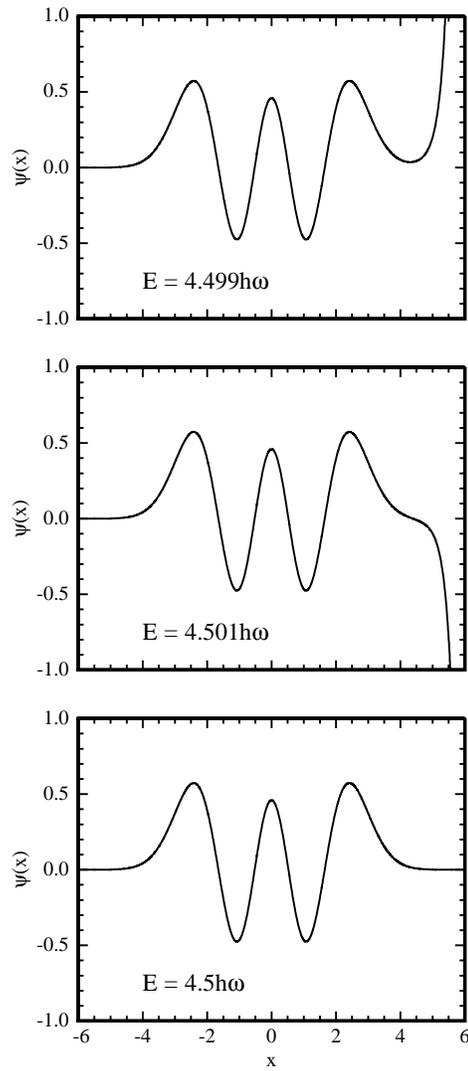
Figure 2.2: Eigenfunctions of the harmonic oscillator Hamiltonian. Only the one in the bottom panel is an acceptable wavefunction.

The math that is needed to handle continuous eigenvalue spectra is more difficult than that for discrete spectra, mostly because the normalization requirement is different. We will therefore not spend much time on continuous spectra. However, we should look at the important special cases of position and momentum.

Eigenvalues of $\hat{p}$ An example of an operator with continuous eigenvalues is the momentum operator $\hat{p}$, whose eigenvalue equation (in one dimension) is

$$-i\hbar\frac{d}{dx}\phi(x) = p\phi(x), \tag{2.95}$$

where p is a number, the momentum. The eigenfunctions of $\hat{p}$ are the functions

$$\phi_p(x) = e^{\frac{ipx}{\hbar}} \tag{2.96}$$

as we can see by applying $\hat{p}$ directly:

$$-i\hbar\frac{d}{dx}e^{\frac{ipx}{\hbar}} = -i\hbar(\frac{ip}{\hbar})e^{\frac{ipx}{\hbar}} \tag{2.97}$$

$$= pe^{\frac{ipx}{\hbar}} \tag{2.98}$$

Any value of p works in this formula. Therefore the eigenvalues of $\hat{p}$ are a continuous set (the real numbers). A measurement of the momentum of one particle moving in one dimension can give any result at all. (Which results are more probable depends on the wavefunction of the particle, as we will see in the next postulate.)

The functions $e^{\frac{ipx}{\hbar}}$ are called *plane waves*, and correspond to particles whose wavefunctions cover all of space but whose momenta p are well-defined. Those functions do not meet the usual square-integrability requirement, so real particles cannot have them as their wavefunctions. However, they can be combined in a way that gives a square-integrable sum. Real particles can have these combinations (called "wavepackets") as their wavefunctions.

Eigenvalues of $\hat{x}$ The eigenfunctions of $\hat{x}$ are strange beasts called "delta functions": they are sharp spikes located at various positions along the axis. Such a function exists for any value of the coordinate x, so again a measurement of particle position can give any result. So both $\hat{p}$ and $\hat{x}$ have continuous eigenvalue spectra.

2.3.2 Predicting measurement results

Nondegenerate eigenfunctions

I want to concentrate first on predicting the results of measurements whose operators have nondegenerate eigenfunctions. The rules for what to do when there are degenerate eigenfunctions are straightforward extensions, and putting them off will not hurt the physical picture and will avoid a bit of extra complication.

Postulate 4 (Discrete, nondegenerate case) If a property whose operator $\hat{A}$ has a discrete, nondegenerate eigenvalue spectrum is measured for a system with the normalized wavefunction $|\psi\rangle$, the probability $P(a_n)$ of obtaining the eigenvalue a_n is $|\langle\phi_n|\psi\rangle|^2$, where $|\phi_n\rangle$ is the normalized eigenfunction of $\hat{A}$ associated with eigenvalue a_n.

The probability in this case is a simple number: since there is a discrete set of eigenvalues, we can give a finite probability that each one of them will appear.

Example Imagine a harmonic oscillator whose normalized wavefunction is

$$|\psi\rangle = \sqrt{\frac{1}{3}}\,|\phi_1\rangle + \sqrt{\frac{2}{3}}\,|\phi_3\rangle, \tag{2.99}$$

where $|\phi_1\rangle$ and $|\phi_3\rangle$ are the normalized harmonic oscillator eigenfunctions with $n = 1$ and 3. Then, if we make a measurement of the total energy of the oscillator, the probability of obtaining the measurement result $E = 3\frac{1}{2}\hbar\omega$ (the eigenvalue corresponding to $|\phi_3\rangle$), is

$$P\left(3\frac{1}{2}\hbar\omega\right) = \left|\langle\phi_3|\psi\rangle\right|^2 \tag{2.100}$$

$$= \left|\langle\phi_3|\left(\sqrt{\frac{1}{3}}\,|\phi_1\rangle + \sqrt{\frac{2}{3}}\,|\phi_3\rangle\right)\right|^2 \tag{2.101}$$

$$= \left|\sqrt{\frac{1}{3}}\,\langle\phi_3|\phi_1\rangle + \sqrt{\frac{2}{3}}\,\langle\phi_3|\phi_3\rangle\right|^2 \tag{2.102}$$

$$= \left|\sqrt{\frac{1}{3}}\,(0) + \sqrt{\frac{2}{3}}\,(1)\right|^2 \tag{2.103}$$

$$= \left|\sqrt{\frac{2}{3}}\right|^2 \tag{2.104}$$

$$= \frac{2}{3} \tag{2.105}$$

So, if we prepared many identical oscillators with the wavefunction $|\psi\rangle$, and measured the total energy of each, we would expect to get the result $E = 3\frac{1}{2}\hbar\omega$ two-thirds of the time.

The connection between the coefficients in the example wavefunction and the resulting probability leads to an alternative statement of this postulate, as shown in the next section.

Predictions in terms of expansion coefficients   Recall that the eigenfunctions of an operator corresponding to a physically measurable quantity form a complete set according to postulate 2. Therefore, it is always possible to write the wavefunction $|\psi\rangle$ as a linear combination of those eigenfunctions:

$$|\psi\rangle = c_0 |\phi_0\rangle + c_1 |\phi_1\rangle + c_2 |\phi_2\rangle + \cdots, \tag{2.106}$$

and we have already worked out a formula for the expansion coefficients c_n,

$$c_n = \langle \phi_n | \psi \rangle. \tag{2.107}$$

This lets us state Postulate 4 in an alternate way: the probability of finding any particular eigenvalue a_n is given by the square of the coefficient of the corresponding eigenfunction in the expansion of the normalized system wavefunction.

Wavefunction already an eigenfunction of $\hat{A}$ We also encounter here the one case where the result of a measurement *can* be predicted exactly: what happens when the system wavefunction is already an eigenfunction of the operator? Recall that the eigenfunctions corresponding to different eigenvalues are orthogonal. Therefore, if the system wavefunction is proportional to one of the eigenfunctions, it will be orthogonal to all the others. All the coefficients in the expansion will be zero except for one, and that one will be 1: there is unit probability (that is, certainty) of obtaining that one eigenvalue as the measurement result.

Example   Imagine that a harmonic oscillator has wavefunction $|\phi_1\rangle$, the eigenfunction of the harmonic oscillator Hamiltonian with quantum number $n = 1$. (In other words, when we expand the system wavefunction in terms of the eigenfunctions of $\hat{H}$, only one term appears, and it's $|\phi_1\rangle$; all the other $|\phi_n\rangle$ have coefficients of zero.) As you showed in your homework, the eigenvalue of $\hat{H}$ corresponding to this function is $E_1 = 1\frac{1}{2}\hbar\omega$. If you measure the energy of this system, you are guaranteed to get the result $E = 1\frac{1}{2}\hbar\omega$.

Postulate 4   (Continuous, nondegenerate case) If a property whose operator $\hat{A}$ has a continuous, nondegenerate eigenvalue spectrum is measured for a system

with the normalized wavefunction $|\psi\rangle$, the probability density $P(a)$ of obtaining a result between a and $a + da$ is $\left|\langle\phi_a|\psi\rangle\right|^2$, where $|\phi_a\rangle$ is the normalized eigenfunction associated with eigenvalue a.

In the case of a continuous spectrum of eigenvalues, it is not possible to give finite probabilities for individual results. The corresponding information is the probability *density*, which is a function that lets you calculate the probability of finding a result within any specified interval of values. Given a probability density function $P(a)$, the probability of finding a result between, say, 2 and 4 is given by

$$P(2 \leq a \leq 4) = \int_2^4 P(a)\, da, \tag{2.108}$$

as we have seen before.

Born interpretation An important application of this postulate is in the so-called Born interpretation of the wavefunction: the probability of finding a particle at a particular location. Applying the postulate directly is a little tricky, because of the unfamiliar delta-function eigenfunctions of $\hat{x}$. But the result is easy to state: the probability density for finding a particle at position x is $|\psi(x)|^2$.

Since the wavefunction is normalized,

$$\int_{-\infty}^{\infty} |\psi(x)|^2\, dx = \langle\psi|\psi\rangle = 1; \tag{2.109}$$

that says that the probability of finding the particle *somewhere* is 1 (certainty). Integrals over smaller regions give the probability of finding the particle in smaller regions; these must be less than or equal to 1.

Expectation values When (as is usually the case) the wavefunction $|\psi\rangle$ is not Levine §18.11
an eigenfunction of the operator $\hat{A}$ corresponding to a measurement, we cannot predict the measurement result with certainty. As you have seen, we can predict the probability of any particular possible result. We can also predict a cruder quantity, the *expectation value* or average expected value. If we make many measurements on identically prepared systems in state $|\psi\rangle$, we expect the average of the results to be

$$\langle a\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{2.110}$$

This quantity must be real because $\hat{A}$ is Hermitian. $\langle a\rangle$ will not usually be equal to any of the eigenvalues a_n of $\hat{A}$, so usually no individual measurement will yield $\langle a\rangle$ as its result. In your homework next week you will prove that

$$\langle\psi|\hat{A}|\psi\rangle = \sum_n P(a_n) a_n, \tag{2.111}$$

as you would expect from the ordinary definition of the average. To show this, expand $\langle\psi|$ in terms of linear combinations of the eigenfunctions of $\hat{A}$, and use the hermititicy of $\hat{A}$, Eq. (2.107), and the expression for $P(a_n)$ from Postulate 4.

There is a widely used shorthand notation for expectation values. In many contexts where the system wavefunction $|\psi\rangle$ has already been specified, people will leave it out of the two sides of the bracket expression, so instead of $\langle\psi|\hat{A}|\psi\rangle$ they will write simply $\langle\hat{A}\rangle$.

Degenerate spectra

Postulate 4 gets extended in a simple way when two or more eigenfunctions of the relevant operator are degenerate. Different eigenfunctions corresponding to the same eigenvalue are not necessarily orthogonal; however, it is always possible to "mix" them in a way that you get a set of orthogonal eigenfunctions that all have the same eigenvalue. (The simplest prescription for doing this mixing is called "Gram-Schmidt orthogonalization" in linear algebra books.) To calculate the probability of obtaining a degenerate eigenvalue a_n in a measurement, you first construct this set of orthogonal eigenfunctions, expand the wavefunction in terms of them, and then add up the squares of the expansion coefficients corresponding to each of the degenerate eigenfunctions. We will put off a detailed study of this procedure until we need it for some problem.

2.3.3   Effect of a measurement on the system

Postulate 5 (Discrete, nondegenerate case) If the measurement of a physical property corresponding to operator $\hat{A}$ results in the eigenvalue a_n, the wavefunction of the system immediately after the measurement is $|\phi_n\rangle$, the eigenfunction of $\hat{A}$ associated with eigenvalue a_n.

A measurement has a drastic effect on the physical system. Though the wavefunction before the measurement gave us probabilities for obtaining several different possible answers, the measurement forces the system to "take a stand". Therefore, if you make another measurement of property $\hat{A}$ immediately after the first, you will get the same result.

This drastic change in the wavefunction upon a measurement goes by the name of "wavefunction collapse". How it happens is not understood in any fundamental way. "Orthodox" quantum mechanics (the kind I am teaching you now) essentially says not to worry about it. It just happens when somebody measures something. The wavefunction collapse is one of the kinds of "quantum jumps" for which the theory is famous.

There's an old joke that illustrates the importance of the wavefunction collapse. Three baseball umpires are bragging about the quality of their work:

Lenny: "I calls'em like I sees'em."

Joe: "I calls'em like they are."

Arthur: "They ain't *nothin'* 'till I calls'em!"

State preparation

The wavefunction collapse upon measurement is a powerful tool for the experimentalist. It provides a way to prepare a system whose wavefunction is known exactly: just measure some property of the system at the beginning of the experiment. Immediately after the measurement, you know exactly what the wavefunction is: it is the eigenfunction of the measurement operator that corresponds to the eigenvalue you just found. After that, it will evolve smoothly (as described in Postulate 6) until you make another measurement.

A good example of the use of this trick is in the spectroscopic technique called "double resonance" (or, more prosaically, "pump and dump.") From a sample of molecules, a laser is used to measure the energy content. The molecules that appear to have a particular energy (determined by the laser wavelength) are then left in one of the eigenfunctions of $\hat{H}$, and the experimenter knows which one. She then allows the molecules to collide with other molecules, move through space, and so on, and then finally measures their energy content again with a second laser. From the result of the second measurement, she tries to work out what happened in between the two.

Uncertainty principle

In quantum mechanics, if two operators corresponding to observable properties   Levine §18.5 do not commute, it is not possible to know both properties of a system precisely and simultaneously. On the other hand, if two operators *do* commute, it *is* possible. Let's examine that case first.

Commuting operators   We need another mathematical fact: if two commuting operators $\hat{A}$ and $\hat{B}$ have discrete, nondegenerate eigenvalue spectra, then the eigenfunctions of $\hat{A}$ are also eigenfunctions of $\hat{B}$ and vice versa. In other words, $\hat{A}$ and $\hat{B}$ have a common set of eigenfunctions. The eigenvalue of a particular function with respect to the two operators need not be the same. (For an accessible proof of this point, see Cohen-Tannoudji, Diu, and Laloë, *Quantum Mechanics* (Wiley, 1977), chapter II, section 3a.)

If we measure the property $\hat{A}$, and find the eigenvalue a_n, the system is left with the wavefunction $|\phi_n\rangle$. Now, if property $\hat{B}$ is measured immediately afterwards, we know that $|\phi_n\rangle$ is also an eigenfunction of $\hat{B}$, with eigenvalue b_n. After the measurement of $\hat{B}$, the system wavefunction is *unchanged*; it is still $|\phi_n\rangle$. We could measure $\hat{B}$ again and be sure of getting result b_n, or we could measure $\hat{A}$ again and be sure of getting result a_n. That is what we mean when we say that both properties can be measured simultaneously and precisely.

Noncommuting operators   Now, what happens when $\hat{A}$ and $\hat{B}$ do not commute? The first measurement, of property $\hat{A}$, has the same effect as before: it finds eigenvalue a_n and leaves the system in state $|\phi_n\rangle$, an eigenfunction of $\hat{A}$. But now, since the two operators do not commute, it is quite unlikely that $|\phi_n\rangle$ will also be an eigenfunction of $\hat{B}$. So if $\hat{B}$ is measured next, we don't know for sure what result we will get. When the $\hat{B}$ measurement is made, the system wavefunction will change again, this time to one of the eigenfunctions of $\hat{B}$. Now, if we measure $\hat{B}$ again, we will get that same eigenvalue. But if we measure $\hat{A}$, once again we do not know what result we will get. It is not guaranteed that we will get a_n again, because the system is no longer in the state $|\phi_n\rangle$; the wavefunction of the system was changed by the measurement of $\hat{B}$. The second measurement has modified the system so that the result of the first measurement no longer applies.

It turns out that the extent of "disturbance" of the system by the second measurement is related to the commutator of the two observables. If the commutator is zero, there is no disturbance, and you can measure the two properties in any order you like without changing the system at all. But if the commutator is not zero, the second measurement modifies the system, introducing an uncertainty into our knowledge of the property we measured first. The general rule (which I will not prove here; see Cohen-Tannoudji et al., chapter 3, complement C) is

$$(\Delta A)(\Delta B) \geq \frac{1}{2}|\langle[\hat{A}, \hat{B}]\rangle|, \tag{2.112}$$

where ΔA and ΔB are the uncertainties (measured as standard deviations) in the quantities A and B.

This "incompatibility" of properties $\hat{A}$ and $\hat{B}$ is at the heart of the famous Uncertainty Principle of Heisenberg. Its most widely known form is that for the $\hat{x}$

and $\hat{p}$ operators, where Eq. (2.112) reduces to

$$\Delta X \Delta P \geq \frac{1}{2} |\langle [\hat{X}, \hat{P}] \rangle| \tag{2.113}$$

$$\geq \frac{1}{2} \geq \frac{1}{2} |\langle i\hbar \rangle| \tag{2.114}$$

$$\geq \frac{1}{2} |i\hbar| \tag{2.115}$$

$$\Delta X \Delta P \geq \frac{1}{2} \hbar \tag{2.116}$$

This version of the Uncertainty Principle says that if you measure a particle's momentum precisely (with zero ΔP), you will then have no idea where it is (that is, all possible results for subsequent measurements of its position are equally likely.)

Hidden variables

It is exceedingly tempting to think about a physical system carrying along a "program" that tells it how to respond to particular measurements. Then the wavefunction, which can tell us only how probable particular outcomes are, is consistent with this more detailed program but does not really carry all the information about the system. In fact the existence of such programs (at least, of the sort that might be carried along with an individual particle) has been conclusively eliminated by a series of experiments. Discussions of this problem appear in most popular books on quantum theory. The sort of experiment required was first proposed by Einstein, Podolsky, and Rosen in 1948, and nowadays are referred to as "EPR experiments" (not to be confused with the unrelated electron paramagnetic resonance experiments, also called EPR.) E, P, and R regarded their experiment more or less purely as a thought experiment; they argued that the result predicted by QM made it clear that such "programs" must be present in physical systems. In 1964 John Bell published a famous theorem (now called, naturally, Bell's theorem) that showed that there would be a measureable difference in experimental results of the EPR experiment if the "programs" existed. This theorem spurred efforts to actually carry out EPR experiments. There has been a steady stream of such experiments, with increasing sophistication, since 1972. The overwhelming result of this experimental program has been that QM is right and E, P, and R were wrong; there can be no program, no "local hidden variables".

2.4 Time evolution

Postulate 6 In the absence of measurements, the wavefunction $|\psi\rangle$ evolves according to the time dependent Schrödinger equation

$$i\hbar\frac{\partial |\psi\rangle}{\partial t} = \hat{H}|\psi\rangle \tag{2.117}$$

where $\hat{H}$ is the Hamiltonian operator for the system.

$\hat{H}$ in quantum mechanics is constructed by taking the classical Hamiltonian and replacing the variables x and p by the operators $\hat{x}$ and $\hat{p} = -i\hbar\frac{d}{dx}$ respectively. Sometimes, $\hat{H}$ depends on time explicitly; for instance, if you shine light on a molecule, you put it in a time-varying electric field and since the molecule contains charged particles (nuclei and electrons) its potential energy will also change with time. But for ordinary, isolated atoms and molecules $\hat{H}$ is independent of time.

Consider a Hamiltonian $\hat{H}$ that does not depend on time. Then, since $\hat{H}$ represents a measureable quantity (the total energy), there is a complete set of eigenfunctions of $\hat{H}$:

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle. \tag{2.118}$$

and the E_n and $|\phi_n\rangle$ are also independent of time.

The system wavefunction $|\psi(t)\rangle$, which probably does depend on time (I will write the time dependence explicitly now), can always be written as a linear combination of the eigenfunctions of $\hat{H}$:

$$|\psi(t)\rangle = \sum_n c_n(t)|\phi_n\rangle, \tag{2.119}$$

where

$$c_n(t) = \langle\phi_n|\psi(t)\rangle. \tag{2.120}$$

Since the eigenfunctions of $\hat{H}$ are time-independent, all the time dependence of $|\psi(t)\rangle$ is contained in the $c_n(t)$. You can think of the time-varying $|\psi(t)\rangle$ as requiring a different linear combination of the $|\phi_n\rangle$ to describe it at each moment.

Let us determine how the $c_n(t)$ vary in time. First we write down the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{2.121}$$

Act on the left with $\langle \phi_n |$:

$$i\hbar \frac{d}{dt}\langle \phi_n | \psi(t) \rangle = \langle \phi_n | \hat{H} | \psi(t) \rangle \tag{2.122}$$

Note that I could move $\langle \phi_n |$ inside the derivative on the left side because it is indepedent of t. Now $\hat{H}$ is Hermitian, so it can operate to the left on its eigenfunction:

$$i\hbar \frac{d}{dt}\langle \phi_n | \psi(t) \rangle = \langle \phi_n | E_n | \psi(t) \rangle \tag{2.123}$$

$$i\hbar \frac{d}{dt}\langle \phi_n | \psi(t) \rangle = E_n \langle \phi_n | \psi(t) \rangle \tag{2.124}$$

Because $c_n(t) = \langle \phi_n | \psi(t) \rangle$, I can write this as

$$i\hbar \frac{d}{dt}c_n(t) = E_n c_n(t) \tag{2.125}$$

This is a first order, linear differential equation for c_n. It requires a single initial condition, the value of c_n at $t = 0$, to completely specify $c_n(t)$. The solution to this equation is

$$c_n(t) = c_n(0)e^{-\frac{iE_n t}{\hbar}}, \tag{2.126}$$

as we can verify by substitution:

$$i\hbar \frac{d}{dt}\left(c_n(0)e^{-\frac{iE_n t}{\hbar}}\right) = E_n c_n(0)e^{-\frac{iE_n t}{\hbar}} \tag{2.127}$$

$$i\hbar \left(\frac{-iE_n}{\hbar}\right)c_n(0)e^{-\frac{iE_n t}{\hbar}} = E_n c_n(0)e^{-\frac{iE_n t}{\hbar}} \tag{2.128}$$

$$E_n c_n(0)e^{-\frac{iE_n t}{\hbar}} = E_n c_n(0)e^{-\frac{iE_n t}{\hbar}} \tag{2.129}$$

So, if we can write at $t = 0$

$$\left| \psi(0) \right\rangle = \sum_n c_n(0)\left| \phi_n \right\rangle, \tag{2.130}$$

then we can find the system wavefunction for all future times, in the absence of any measurements:

$$\left| \psi(t) \right\rangle = \sum_n c_n(0)e^{\frac{-iE_n t}{\hbar}}\left| \phi_n \right\rangle \tag{2.131}$$

Notice that the evolution of $|\psi\rangle$ in time is smooth and deterministic. Once we know $|\psi\rangle$ at some time, the wavefunction at any later time is completely determined. The only time the wavefunction changes in an unpredictable way is when a measurement is made; then, the wavefunction changes instantly to one of the eigenfunctions of the operator corresponding to the measurement, as described by Postulate 5.

Stationary states

What happens if $|\psi(0)\rangle$ is one of the eigenfunctions of $\hat{H}$? Then we have

$$|\psi(0)\rangle = |\phi_n\rangle \tag{2.132}$$

$$|\psi(t)\rangle = e^{\frac{-iE_n t}{\hbar}} |\phi_n\rangle \tag{2.133}$$

Consider some observable property whose operator is $\hat{A}$. $\hat{A}$ has its own set of eigenfunctions $|v_m\rangle$ with eigenvalues b_m:

$$\hat{A}|v_m\rangle = b_m |v_m\rangle \tag{2.134}$$

When we measure property $\hat{A}$ on the system in state $|\psi(t)\rangle$, we must get one of the b_m as the result. The probability of getting one particular result is

$$P(b_m) = \left|\langle v_m|\psi(t)\rangle\right|^2 \tag{2.135}$$

$$= \left|\left\langle v_m \left| e^{\frac{-iE_n t}{\hbar}} \right| \phi_n \right\rangle\right|^2 \tag{2.136}$$

$$= \left| e^{\frac{-iE_n t}{\hbar}} \langle v_m|\phi_n\rangle \right|^2 \tag{2.137}$$

$$= \left| e^{\frac{-iE_n t}{\hbar}} \right|^2 \left|\langle v_m|\phi_n\rangle\right|^2 \tag{2.138}$$

$$= \left|\langle v_m|\phi_n\rangle\right|^2 \tag{2.139}$$

which is *independent of time*. So if the wavefunction of the system at $t = 0$ is one of the eigenfunctions of $\hat{H}$, all the measureable properties of the system remain constant. Such states are called *stationary states*.

In fact we have just proved a general mathematical property of wavefunctions: you can multiply any wavefunction by a number of the form $e^{i\theta}$, where θ is any real number, and none of the physical properties of the system will change. Note, however, that if your wavefunction is made of a sum of terms, multiplying any one of the terms by a number $e^{i\theta}$ *will* change the physical properties; the so-called "phase factors" are unimportant only if they multiply the entire wavefunction and not just a part of it.

Time independent Schrödinger equation

We have seen that the eigenfunctions of the Hamiltonian play an important role in two ways:

1. If an arbitrary wavefunction $|\psi(0)\rangle$ is written as a linear combination of eigenfunctions of $\hat{H}$, then it is very easy to predict the wavefunction at any later time. You just multiply each of the expansion coefficients c_n by $\exp(-iE_n t/\hbar)$, where E_n is the energy (eigenvalue of $\hat{H}$) that goes with that particular eigenfunction.

2. If at some time a system has one of the eigenfunctions of $\hat{H}$ as its wavefunction, then none of the properties of the system will ever change, unless someone makes a measurement or changes $\hat{H}$. The system is in a "stationary state".

For these reasons, finding the eigenfunctions and eigenvalues of $\hat{H}$ is one of the most important tasks in the quantum mechanical approach to any problem. The eigenvalue equation for $\hat{H}$,

$$\hat{H}|\psi\rangle = E|\psi\rangle, \tag{2.140}$$

is called the *time independent Schrödinger equation.* It can be derived from the regular (time dependent) Schrödinger equation when $\hat{H}$ is independent of time by a standard technique called separation of variables.

Evolution of expectation values; Ehrenfest's theorem

I will state, without proof (the proof is not hard) the following property of expectation values: if $\hat{A}$ is the operator for some observable property, and has no explicit time dependence (the usual case), and $|\psi\rangle$ is the normalized wavefunction of the system, then the expectation value of $\hat{A}$ evolves in time according to

$$\frac{d}{dt}\langle\psi|\hat{A}|\psi\rangle = \frac{1}{i\hbar}\langle\psi|[\hat{A},\hat{H}]|\psi\rangle \tag{2.141}$$

That is, the time derivative of the expectation value of a property is proportional to the commutator of the corresponding operator with the Hamiltonian. If an operator commutes with the Hamiltonian, its expectation value will not change with time, even though the wavefunction does change with time; it is called a "constant of the motion".

Applying Eq. (2.141) to the $\hat{x}$ and $\hat{p}$ operators, for a system whose Hamiltonian is given by the usual

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{2.142}$$

yields the results

$$\frac{d}{dt}\langle \hat{x} \rangle = \frac{1}{m}\langle \hat{p} \rangle \tag{2.143}$$

$$\frac{d}{dt}\langle \hat{p} \rangle = \langle -\nabla V(\hat{x}) \rangle \tag{2.144}$$

This result is known as Ehrenfest's theorem. (Keep in mind that the notation $\langle \hat{A} \rangle$ is shorthand for $\langle \psi | \hat{A} | \psi \rangle$.) If you think of a particle moving along as a "wavepacket", it shows that the expectation value of the position (essentially, the "center" of the wavepacket) moves classically according to the average value of the momentum, and the expectation value of the momentum responds classically to the force *averaged over the spatial extent of the wavepacket.* So, in some sense, the average values of position and momentum follow classical mechanics even though the wavefunction itself does not.

Chapter 3

Spectroscopic applications

> The theory of the possible vibrations of a molecule has not yet been studied as it ought, with the help of a continual comparison between the dynamical theory and the evidence of the spectroscope. An intelligent student, armed with the calculus and the spectroscope, can hardly fail to to discover some important fact about the internal constitution of a molecule.
>
> — J. Clerk Maxwell, lecture to the Chemical Society, 18 February 1875

3.1   Harmonic oscillator as model of diatomic molecule vibrations

3.1.1   Born-Oppenheimer approximation and the effective nuclear potential

The *Born-Oppenheimer approximation*, which we will study explicitly in a few weeks, suggests that it is reasonable to approach molecular problems by assuming that the electrons can respond very quickly to changes in nuclear position. We can therefore study the motion of nuclei in molecules without worrying about the details of the electrons' motions, by using the following tactic. We will find the eigenvalues of the electronic Schrödinger equation with the nuclei in fixed positions. At the moment, we will not worry about the details of how the electronic eigenvalues might be found (students in 353 have been doing just that this week, and later in 356 we will talk about the available methods.) The total electronic energy, plus the internuclear Coulomb repulsion, at one internuclear distance will be used as an effective potential energy for the motion of the nuclei.

Levine §20.2

Repeating this procedure for many different internuclear separations will allow us to map out an effective potential energy curve for the nuclei. We can then study the motion of the nuclei in this effective potential, without any longer worrying about what the electrons are doing.

A typical effective potential curve looks like that in Figure 3.1. It rises steeply at small internuclear separations, because of the overlap of the inner-shell electrons of the two atoms and (less importantly) because of the Coulomb repulsion of the two nuclei. At intermediate distances it becomes negative, reflecting the bonding interaction. At long distances it weakens and asymptotically becomes independent of separation; once the two electron clouds are far apart, they do not interact and one atom no longer cares where the other is.
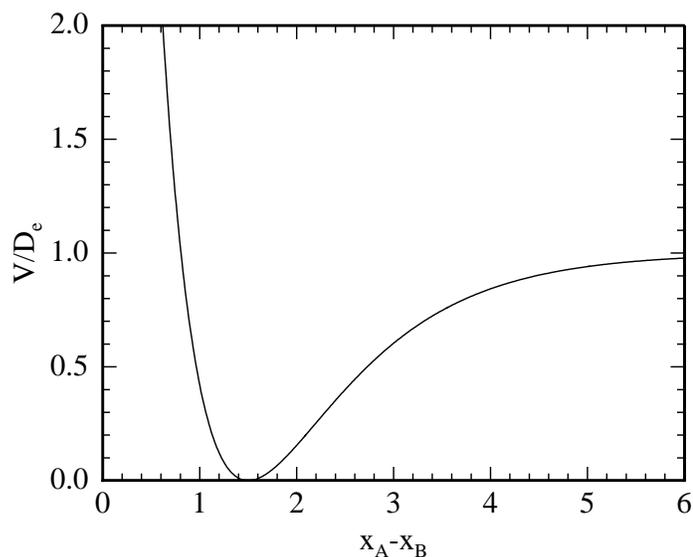
Figure 3.1: Typical diatomic molecule effective potential curve.

In Figure 3.1 I have set the zero of potential energy (arbitrarily) at the minimum, because that corresponds best to our study of the harmonic oscillator problem. It is also common to see the zero of potential energy set at the long-range dissociation asymptote. In electronic structure calculations, total energies generally have the zero of potential set at completely separated nuclei and electrons, and the calculated total energies are therefore large and negative. Note that this convention of a zero of energy really is our own choice: adding a constant to both the Hamiltonian and the energy eigenvalue in the time-independent

Schrödinger equation leaves the wavefunctions the same as before.

3.1.2   Reduction to one-particle problem

Given this effective potential energy that depends on the distance between the   Levine §18.13
two nuclei, the next task is to write down the complete Hamiltonian for the
problem. The classical Hamiltonian is

$$H = \frac{p_A^2}{2m_A} + \frac{p_B^2}{2m_B} + V(x_A - x_B) \tag{3.1}$$

Because the potential depends only on the separation between the two atoms,
and not on where either one is in the universe, it is possible to separate the overall
or center-of-mass motion from the relative motion of the two atoms. To do this
we introduce two new coordinates, the *center of mass position*

$$X = \frac{m_A x_A + m_B x_B}{m_A + m_B} \tag{3.2}$$

and the *relative separation* (or interparticle distance)

$$x = x_A - x_B, \tag{3.3}$$

in terms of which the old variables are

$$x_A = X + \frac{m_B}{m_A + m_B} x \tag{3.4}$$

$$x_B = X - \frac{m_A}{m_A + m_B} x \tag{3.5}$$

Now since

$$p_A = m_A \frac{dx_A}{dt}, \tag{3.6}$$

and similarly for p_B, by differentiating Eq. (3.4) and Eq. (3.5) you can find a new
form of the classical Hamiltonian written in terms of X and x. When you do
this you find that all the terms involving products of time derivatives of X and x
cancel, leaving you with the following simple expression:

$$H = \frac{P^2}{2M} + \frac{p^2}{2\mu} + V(x), \tag{3.7}$$

where the total mass is

$$M = m_A + m_B, \tag{3.8}$$

the *reduced mass,* or effective one-particle mass, is

$$\mu = \frac{m_A m_B}{m_A + m_B}, \tag{3.9}$$

the momentum of the center of mass is

$$P = M\frac{dX}{dt}, \tag{3.10}$$

and the relative momentum is

$$p = \mu\frac{dx}{dt}. \tag{3.11}$$

$$\tag{3.12}$$

So now, we have a new Hamiltonian that looks like the sum of two independent parts: one, a particle of mass M moving freely (not subject to any forces, and therefore having only kinetic energy), and a second one of mass μ subject to the force $V(x)$. The quantum Hamiltonian is obtained by replacing the P, p, and x variables with their respective operators.

In general, when you have a quantum mechanical operator that is a sum of two parts that depend on different variables, the corresponding eigenfunctions are *products* of two functions, one for each variable, and the corresponding eigenvalues are similarly sums of two values. Let's see this by trying in the eigenvalue equation of $\hat{H}$ the wavefunction $\Phi(X)\phi(x)$:

$$\left(\frac{\hat{P}^2}{2M} + \frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\Phi(X)\phi(x) = E\Phi(X)\phi(x) \tag{3.13}$$

Now since $\hat{P}$ does not care about x, and vice versa, we have

$$\phi(x)\frac{\hat{P}^2}{2M}\Phi(X) + \Phi(X)\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E\Phi(X)\phi(x) \tag{3.14}$$

Divide both sides by $\Phi(X)\phi(x)$ to get

$$\frac{1}{\Phi(X)}\frac{\hat{P}^2}{2M}\Phi(X) + \frac{1}{\phi(x)}\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E \tag{3.15}$$

Now here comes a subtle bit of argument. On the left side of this equation is a sum of two functions. One is a function of X only; the other is a function of x only. The equation says that the sum of these two functions must be constant.

But I could always vary X while holding x fixed (that is, move the molecule in space without stretching or compressing the bond), and I would still have to satisfy the equation; that means that just the X part must itself be constant. A similar argument holds for x. So in order for this equation to be satisfied, *both* the following equations must be satisfied:

$$\frac{\hat{P}^2}{2M}\Phi(X) = E_M\Phi(X) \tag{3.16}$$

$$\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E_\mu\phi(x) \tag{3.17}$$

where we have

$$E_M + E_\mu = E \tag{3.18}$$

In other words, our quantum problem has separated into two completely independent ones; we have separate time-independent Schrödinger equations for the center of mass motion and for the relative motion. The capital-lettered problem represents the molecule, mass M, moving through space; the lowercase-lettered problem represents the relative motion of the two nuclei with the motion of a single hypothetical particle of mass μ moving in the potential $V(x)$. When the Hamiltonian is the sum of two independent pieces, you get two independent problems you can solve separately.

3.1.3 Approximation of potential curve by parabola

Now that we have separated the center-of-mass and relative motions, let us concentrate on the solution of the relative motion problem. We have a single, hypothetical "particle" of mass μ moving in a potential $V(x)$. That potential is, for real molecules, definitely not a parabola; at large separation, real molecules dissociate. But if for the moment we confine ourselves to thinking about molecules that do not contain nearly enough vibrational energy to dissociate, we can make good progress by approximating the interatomic potential by a parabola.

Which parabola? Use a Taylor series, expanding about the minimum of the potential (called x_e):

$$V(x) = V(x_e) + \left(\frac{dV(x)}{dx}\right)_{x_e}(x - x_e) + \frac{1}{2}\left(\frac{d^2V(x)}{dx^2}\right)_{x_e}(x - x_e)^2 + \cdots \tag{3.19}$$

The first term is a constant, and sets the zero of potential energy for the system. The second term is always zero, because x_e is the minumum of the interatomic potential, and the way you find a minimum is to take the derivative and set it to zero. So the first interesting term is the second-derivative term. Comparing it to

the normal harmonic oscillator potential $V(x) = \frac{1}{2}kx^2$, it is easy to see that the force constant k is equal to the second derivative of the intermolecular potential evaluated at the potential minimum.

To complete our model we must make a new variable $y = x - x_e$ and change to that variable; because x_e is a constant, $\frac{d^2}{dy^2} = \frac{d^2}{dx^2}$ (that is, the kinetic energy operator does not change) and we have

$$\left(-\frac{\hbar^2}{2\mu}\frac{d^2}{dy^2} + \frac{1}{2}ky^2\right)\phi_n(y) = E_n\phi_n(y) \tag{3.20}$$

which is exactly the time-independent Schrödinger equation for the harmonic oscillator.

For normal molecules held together by chemical bonds, this is not too bad a model for the low vibrational levels. Recall that in the harmonic oscillator, the energy eigenstates are all equally spaced with spacing $\hbar\omega$. That equal spacing is a special property of the harmonic oscillator problem with its parabolic potential; no other potential shape has equally spaced levels. In real molecules, the potential is not parabolic, but goes asymptotically to a dissociation limit as in Figure 3.1. In that case, the energy levels usually become closer together as the excitation increases. However, near the bottom of the potential well, the parabola is quite a good approximation to the true curve and the error in spacings predicted by the harmonic oscillator model is only a few percent.

Example: CO vibrational energy levels   The energy levels of carbon monoxide are very precisely known from spectroscopy, so it makes a good test case for the harmonic potential approximation. The ground state of CO has an energy $1081.77097\,\mathrm{cm}^{-1}$ above the bottom of the potential well. (An energy quoted in "wavenumbers", cm^{-1}, is really the energy (in joules) divided by hc.) If we take that value as $\frac{1}{2}\hbar\omega$ for CO, then we have

$$\frac{1}{2}\hbar\omega = hc(1081.77097\,\mathrm{cm}^{-1}) \tag{3.21}$$

$$\omega = 4\pi c(1081.77097\,\mathrm{cm}^{-1}) \tag{3.22}$$

$$= 4.075462 \times 10^{14}\,\mathrm{s}^{-1} \tag{3.23}$$

$$k = \mu\omega^2 \tag{3.24}$$

$$= \frac{12 \cdot 16\,\mathrm{amu}}{12 + 16}(1.660538 \times 10^{-27}\,\mathrm{kg/amu}) * (4.075462 \times 10^{14}\,\mathrm{s}^{-1})^2 \tag{3.25}$$

$$k = 1891.237\,\mathrm{kg\,s}^{-2} = 1891.237\,\mathrm{J\,m}^{-2} \tag{3.26}$$

$$= 9.52071 \times 10^5\,\mathrm{cm}^{-1}\mathrm{\AA}^{-2} \tag{3.27}$$

where I made the last conversion by dividing by hc to convert to cm^{-1} and converting from meters to angstroms, Å. ($1\,\text{Å} = 0.1\,\text{nm} = 1.0 \times 10^{-10}\,\text{m}$. The Å is slowly being replaced by the nm and pm, but is still very widely used, and is still "officially tolerated" by the international committee responsible for the metric system.)

Figure 3.2 shows the accurately known potential energy curve for CO and its energy levels, and the harmonic potential with the force constant just calculated with its energy levels. For the first few levels, the harmonic approximation is pretty good, but as v increases, the spacing of the levels becomes smaller while the harmonic spacing remains the same. By $v = 9$, the accumulated errors are about half a vibrational spacing.
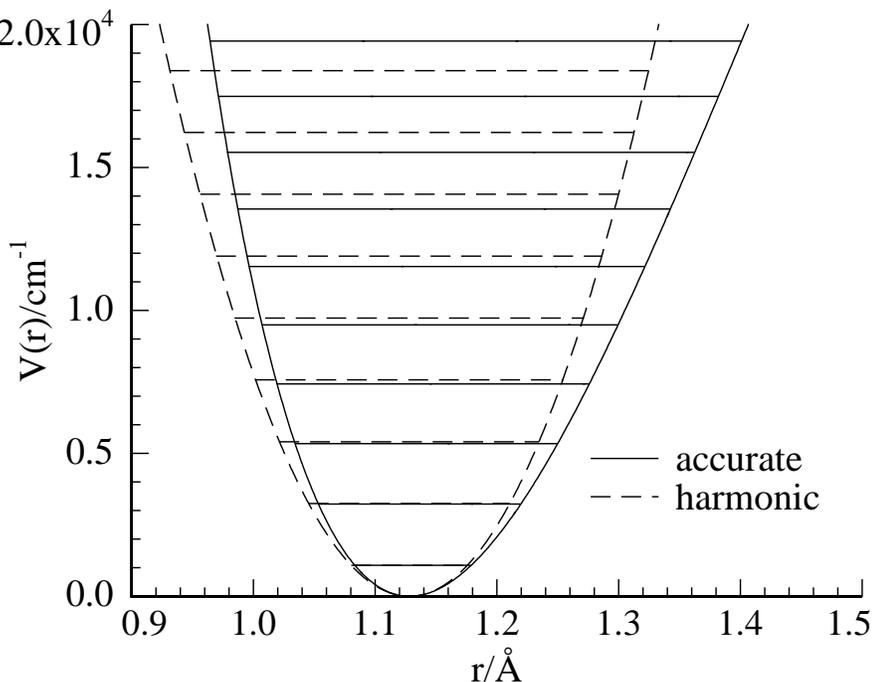

Figure 3.2: Low-energy part of the potential curve for CO, and energy levels for the normal isotope. The harmonic approximation to the potential, with its energy levels, is also shown.

The harmonic oscillator model can also be used to make predictions about

transitions that the molecule can make when it is disturbed by some outside force. That is our next topic.

3.2 Transitions

We have already talked about how wavefunctions change with time under the influence of a constant Hamiltonian: if

$$\left|\psi(t=0)\right\rangle = \sum_n c_n \left|\phi_n\right\rangle, \tag{3.28}$$

where the $\left|\phi_n\right\rangle$ are the eigenfunctions of $\hat{H}$ and have associated eigenvalues E_n, then at later times,

$$\left|\psi(t)\right\rangle = \sum_n c_n e^{-\frac{iE_n t}{\hbar}} \left|\phi_n\right\rangle \tag{3.29}$$

The quotient $E_n/\hbar$ in the exponent has units of a frequency and is often written ω_n; the expression is then written

$$\left|\psi(t)\right\rangle = \sum_n c_n e^{-i\omega_n t} \left|\phi_n\right\rangle. \tag{3.30}$$

If the sum has only one term—that is, if $\left|\psi(t=0)\right\rangle$ is already an eigenfunction of $\hat{H}$—then the single phase factor $e^{-i\omega_n t}$ is unimportant and none of the observable properties of the system will change with time. On the other hand, if more than one of the $c_n(0)$ is nonzero, then the wavefunction will change in time in a nontrivial way. In that case, only properties whose operators commute with $\hat{H}$ will be constant; other properties will have expectation values that change with time.

3.2.1 Temporary change of $\hat{H}$

Now consider what happens if the system starts out in a stationary state (the wavefunction is an eigenfunction of $\hat{H}$), but then the Hamiltonian changes to something different. (For example, we might suddenly apply an electric field to the molecules.) If $\hat{H}$ changes, then the original ket $\left|\phi_k\right\rangle$ will probably not be an eigenfunction of the new Hamiltonian. Its expansion in terms of the eigenfunctions of the *new* Hamiltonian will have more than one term; it will therefore evolve in time according to Eq. (3.30), where now the $\left|\phi_n\right\rangle$ and E_n must represent the eigenfunctions and eigenvalues of the new Hamiltonian. If, after some time, the Hamiltonian reverts to the old one (for example, we turn off

the electric field), the wavefunction will not be the same one we started with. During the time when the Hamiltonian was different, the wavefunction will have changed, and when the Hamiltonian returns to the original one the wavefunction will have a new expansion with (probably) more than one nonzero term. If we measure the total energy at that point, we might find any of the eigenvalues whose eigenfunctions are part of the new wavefunction. If the system started out in state $|\phi_k\rangle$, and after changing the Hamiltonian to something else and then back again we measure a total energy E_m, then we say that the system has made a *transition* from state $|\phi_k\rangle$ to the state $|\phi_m\rangle$.

The transition probabilities can be difficult to calculate. (I do these calculations sometimes, where $\hat{H}$ is changed by bringing another atom close to the molecule being studied, and I often run the calculations on supercomputers.) But in the common case where the new Hamiltonian $\hat{H}_b$ is not too different from the old one, and is not applied for very long, there is an elegant theory that gives good approximations to the transition probabilities with simple formulas. We discuss it next.

3.2.2   Time dependent perturbation theory

There is a very useful theory that lets you solve a quantum mechanical problem in an approximate way if your problem is similar to another problem that has already been solved. The theory is called perturbation theory, and it is useful in several different situations.

In one version, called time independent perturbation theory, you can solve a problem whose Hamiltonian (still not a function of time) can be obtained from another Hamiltonian $\hat{H}_0$ by adding a small piece, $\hat{H}_1$.

For instance, if you want to make a better approximation to the vibrational energies of a diatomic molecule, you could take the next term in the Taylor expansion of the potential, proportional to x^3, and treat it as $\hat{H}_1$. Then time independent perturbation theory gives you simple formulas for the changes in the harmonic oscillator energy levels and wavefunctions (which you already know) that appear when the cubic term is added to the potential. We'll discuss this procedure in more detail in about two weeks.

The version I am interested in at this point is called time dependent perturbation theory, and it handles the case where $\hat{H}_1$ might itself be a function of time. The most important such problem for chemists is spectroscopy: we have atoms and molecules with their own Hamiltonians $\hat{H}_0$, and then we shine light on them. The light contains electric and magnetic fields that change with time sinusoidally: in a continuous laser beam, for example, the electric field intensity E varies as $E = E_0 \cos \omega_l t$, where ω_l is the angular frequency of the light. Since atoms and

molecules contain charged particles (electrons and nuclei), they feel forces due to those electric and magnetic fields, and those forces modify the potential energy parts of their Hamiltonians.

I will not develop time dependent perturbation theory in detail, though the derivation is not difficult (we could do it in maybe two lectures). I will simply give its most important result: the probability of making a transition from stationary state $|\phi_k\rangle$ to state $|\phi_m\rangle$ under the influence of a small, short-lived perturbing Hamiltonian $\hat{H}_1$ is proportional to $\langle\phi_m|\hat{H}_1|\phi_k\rangle$.

3.2.3   Selection rules

Levine §21.2

The expression just given is the basis for spectroscopic *selection rules*, rules that tell us what kinds of spectroscopic transitions are possible and what kinds are not. Whenever the matrix element $\langle\phi_m|\hat{H}_1|\phi_k\rangle$ is known to be zero, and the perturbing Hamiltonian is weak enough and applied for a short enough time, transitions from $|\phi_k\rangle$ to $|\phi_m\rangle$ will not happen.

As our first example, consider a heteronuclear diatomic molecule whose motion we will model with a harmonic oscillator. Heteronuclear molecules have dipole moments that change as the internuclear distance changes. We can use a Taylor expansion to approximate the change in dipole moment for small displacements from the equilibrium internuclear separation:

$$\mu(x) = \mu(0) + \mu'(0)x + \frac{\mu''(0)}{2}x^2 + \cdots \tag{3.31}$$

where $\mu(0)$ is the molecule's permanent dipole moment, $\mu'(0)$ is the slope of the dipole vs. x graph at $x = 0$, and so on. The energy of interaction between an electric field $\mathbf{E}$ and a dipole moment $\boldsymbol{\mu}$ is proportional to $\mathbf{E}\cdot\boldsymbol{\mu}$. Therefore,

$$\hat{H}_1 \propto E_x(\mu(0) + \mu'(0)\hat{x}) \tag{3.32}$$

if we include terms only up to first order in the dipole moment function; E_x is the component of the electric field along the internuclear axis of the molecule. The perturbing Hamiltonian therefore looks like

$$\hat{H}_1 \propto a + b\hat{x} = a + b'(\hat{a}^\dagger + \hat{a}). \tag{3.33}$$

To decide which transitions are possible, we must consider matrix elements of the type $\langle\phi_m|\hat{H}_1|\phi_k\rangle$. We will therefore have transitions only when

$$\left\langle\phi_m\left|a + b'(\hat{a}^\dagger + \hat{a})\right|\phi_k\right\rangle \neq 0. \tag{3.34}$$

Brief inspection should convince you that this can only be true if $m = k$ or $m = k \pm 1$. The first possibility, $m = k$, can produce transitions but they do not correspond to changes in vibrational state (they are responsible for "pure rotational" spectroscopy, carried out primarily in the microwave and radio parts of the spectrum). If $m = k + 1$ the vibrational quantum number has increased by one, and we have an absorption transition; this is what the laboratory experiment on HCl absorption studies. If $m = k - 1$, the molecule has *emitted* a photon and dropped to a vibrational state one lower. The vibrational transitions of most molecules occur in the infrared region of the electromagnetic spectrum, between wavelengths of 1 and 10 μm.

This analysis has given us our first example of a selection rule: for a harmonic oscillator perturbed by electric dipole radiation, $\Delta v = \pm 1$. Selection rules have enormous importance in spectroscopy and we shall see many more examples.

3.3 Particle in a Box

Before we move on to angular momentum (our next big topic), let's examine another important one-dimensional problem. Many books use it as the main supply of basic quantum ideas, much as I have used the harmonic oscillator.

The "particle in a box" describes a particle constrained to move only within a certain part of the x axis by infinitely high potential walls at its boundaries. Within the "box", the particle feels no forces at all. The Hamiltonian is therefore

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) \tag{3.35}$$

with

$$V(x) = \begin{cases} 0 & \text{if } 0 \le x \le L, \\ \infty & \text{otherwise.} \end{cases} \tag{3.36}$$

As we have seen, if you know the eigenfunctions and eigenvalues of the Hamiltonian for a quantum mechanical system, you can describe any possible wavefunction in terms of them. Let us look for the eigenfunctions and eigenvalues of the particle-in-a-box Hamiltonian.

3.3.1 Eigenfunctions of $\hat{H}$

Writing $\hat{H}$ in terms of derivatives, the time-independent Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \phi(x) = E\phi(x) \tag{3.37}$$

Outside the box the potential is infinite. Any $\phi(x)$ other than $\phi(x) = 0$ will therefore require infinite energy, so we need $\phi(x) = 0$ outside the box.

Inside the box we need to satisfy

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \right]\phi(x) = E\phi(x) \tag{3.38}$$

and, as always, $\phi(x)$ must be continuous, finite everywhere, and square-integrable. We are looking for eigenfunctions of the second derivative operator that satisfy those requiremements. Finding them is very similar to the classical treatment of a harmonic oscillator we used at the beginning of the course.

We could write the solution of Eq. (3.38) in terms of either exponentials or trig functions; the latter course is a little more direct and I will adopt it here. Because Eq. (3.38) is a second-order equation, for the most general solution we need to consider a sum of two *linearly independent* solutions; sin and cos functions will do. By educated guessing I propose that a solution of Eq. (3.38) is therefore

$$\phi(x) = A\cos ax + B\sin ax, \tag{3.39}$$

and we can demonstrate that this is in fact a solution because the second derivative of $\phi(x)$ is

$$\frac{d^2}{dx^2}\phi(x) = -a^2(A\cos ax + B\sin ax). \tag{3.40}$$

Note that we must use the same "wavenumber", a, in both the sin and cos terms because otherwise $\phi(x)$ will not be an eigenfunction of $\hat{H}$; we won't be able to factor a single value out of the second derivative as we did here.

A, B, and a will be determined by the conditions on $\phi(x)$; note that *any* values of A, B, and a satisfy the time-independent Schrödinger equation for some E. Just as in the harmonic oscillator case, it is the combination of the Schrödinger equation and the conditions on $\phi(x)$ that restrict the permissible set of eigenfunctions.

$\phi(x)$ must be continuous; therefore we must have $\phi(0) = 0$ and $\phi(L) = 0$ to match up with the solution outside the box. Setting $x = 0$ in Eq. (3.39) and requiring $\phi(0) = 0$ gives immediately $A = 0$ since $\cos(0) \neq 0$. We therefore have

$$\phi(x) = B\sin ax. \tag{3.41}$$

B cannot be 0; otherwise we would have no particle. However, we must have $\phi(L) = 0$. Therefore the sin function must vanish at $x = L$, and it vanishes only for arguments that are integer multiples of π. So,

$$aL = n\pi \tag{3.42}$$

which gives

$$a = \frac{n\pi}{L}, \tag{3.43}$$

where $n = 1, 2, 3, \cdots$. ($n = 0$ satisfies the boundary condition but makes $\phi(x)$ zero everywhere: again, this is no good because we have no particle. Negative integers also work, but because $\sin(-x) = -\sin(x)$ they amount to choosing a different value of B rather than providing independent eigenfunctions.).

We now have

$$\phi_n(x) = B \sin \frac{n\pi x}{L}. \tag{3.44}$$

How do we find B? We normalize!

$$1 = \int_0^L |\phi(x)|^2 \, dx \tag{3.45}$$

$$= \int_0^L B^2 \sin^2 \frac{n\pi x}{L} \, dx \tag{3.46}$$

$$= B^2 \left[\frac{x}{2} - \frac{1}{4} \frac{L}{n\pi} \sin \frac{2n\pi x}{L} \right]_0^L \tag{3.47}$$

$$= B^2 \frac{L}{2} \tag{3.48}$$

so

$$B = \left(\frac{2}{L} \right)^{\frac{1}{2}}. \tag{3.49}$$

We have now arrived at the normalized, acceptable eigenfunctions of the Hamiltonian:

$$\phi_n(x) = \left(\frac{2}{L} \right)^{\frac{1}{2}} \sin \frac{n\pi x}{L} \tag{3.50}$$

with n a positive integer. The first few of these functions are shown in Figure 3.3 below. As expected, they are zero at the edges of the box, have increasing numbers of nodes (with the lowest one having no nodes), and have a definite symmetry (either even or odd) with respect to the center of the box.

3.3.2   Eigenvalues of the Hamiltonian

We have not yet worked out the energies corresponding to each of the eigenfunctions of the Hamiltonian. Inserting the eigenfunctions Eq. (3.50) into the eigenvalue equation Eq. (3.38) and applying the second derivative formula Eq. (3.40)
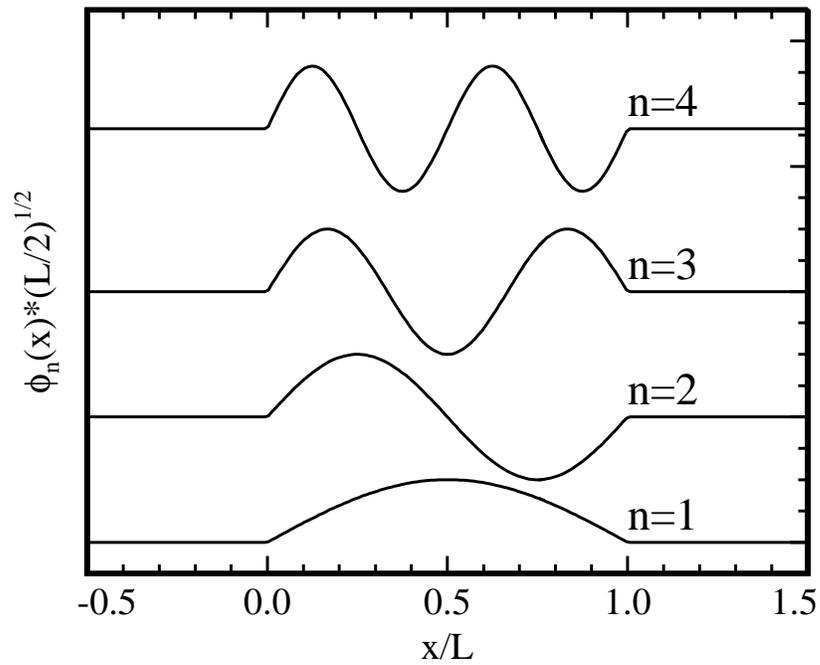
Figure 3.3: Stationary state wavefunctions for the particle in a one-dimensional box. The functions are offset vertically for clarity; all of them have value zero outside the range $0 \leq \frac{x}{L} \leq 1$. The vertical offsets do *not* scale with the energy differences between the functions. (In this potential, the energies go up quadratically with the quantum number n.)

gives us

$$-\frac{\hbar^2}{2m}\left[-\left(\frac{n\pi}{L}\right)^2\right]\sin\frac{n\pi x}{L} = E_n\sin\frac{n\pi x}{L} \tag{3.51}$$

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2} \tag{3.52}$$

$$= \frac{h^2 n^2 \pi^2}{(2\pi)^2 2mL^2} \tag{3.53}$$

$$= \frac{n^2 h^2}{8mL^2} \tag{3.54}$$

So, the energies go up *quadratically* with n, in contrast to the harmonic oscillator where they are equally spaced. The zero point energy is that of the lowest level, $n = 1$, and is $\frac{h^2}{8mL^2}$. Notice that the shorter the box, the higher the ZPE, and the lighter the particle, the higher the ZPE.

The PIB eigenfunctions form a complete set; any function of x that is continuous, single-valued, and zero for $x \le 0$ and $x \ge L$ can be expanded as a weighted sum of them. Eq. (2.90) gives a prescription for finding the expansion coefficients just as it did in the harmonic oscillator case.

Chapter 4

Angular Momentum

The angular momentum vector **L** in classical mechanics is defined as the cross product of the momentum and position vectors:[1]

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{4.1}$$

where **r** is the position vector of the particle and **p** is its momentum vector. From the definition of the cross product, we find formulas for the three components of the angular momentum vector, and moving to the quantum problem just requires the placements of hats on all the variables:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \tag{4.2}$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \tag{4.3}$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{4.4}$$

Note how each of the above lines follows from the previous by the replacements, $x \to y$, $y \to z$, $z \to x$. This replacement is called a "cyclic permutation" of x, y and z.

The angular momentum is a vector, with three components; usually I will consider each of the components separately, but occasionally I will write $\hat{\mathbf{L}}$ when I can make some statement about all three components at once. You should realize that any equation in which $\hat{\mathbf{L}}$ appears is really three equations, one for each of the three components.

4.1   Approach to the problem

You can find the harmonic oscillator eigenvalues and eigenfunctions in either of two equivalent ways: either by working directly with the raising and lowering

---

[1]Much of the substance of this section is due to Prof. Sherwin Singer of The Ohio State University.

operators and their commutators, or by solving the differential equation that appears when you write out the eigenvalue equation of the Hamiltonian in terms of the momentum operator $\hat{p} = -i\hbar\frac{d}{dx}$. Both approaches give the same result. We have not done either in detail; I have showed you solutions to the HO problem and asked you to verify some of their properties.

In the angular momentum problem, which in its simplest physical incarnation corresponds to a particle constrained to move around a circle (in two dimensions) or on the surface of a sphere (in three dimensions), the algebraic and differential approaches are *not* equivalent. The differential approach, which seeks eigenvalues and eigenfunctions of the differential angular momentum operators (replacing every $\hat{p}$ above with its corresponding differential form: $\hat{p}_z = -i\hbar\frac{d}{dz}$, for example), finds only *half* the possible solutions. The half it finds is the description of "orbital" or "rotational" angular momentum: the motion of an electron about a nucleus in an atom, or the tumbling of a linear molecule in space. The other half are responsible for everything having to do with "spin" in atoms and molecules; they simply do not appear!

I will therefore follow the algebraic path that leads to the eigenvalues of the angular momentum operators and their corresponding kets. At the end, I will point out what the operators look like in terms of derivatives, because we'll need to recognize them later, but that's a pretty messy problem and we'll avoid it when we can.

4.1.1 Commutation relations among the components of $\hat{\mathbf{L}}$

First let's see if the x, y and z components of the angular momentum operator commute with each other. Start by evaluating the commutator of $\hat{L}_x$ and $\hat{L}_y$.

$$[\hat{L}_x, \hat{L}_y] = [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \tag{4.5}$$

Writing the commutator out and collecting terms gives

$$= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] \tag{4.6}$$

The last two commutators are zero, because operators depending on different variables commute. So we have

$$= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \tag{4.7}$$

and writing out those two and factoring (separately) we get

$$= \hat{y}\hat{p}_x \left[\hat{p}_z, \hat{z}\right] + \hat{x}\hat{p}_y \left[\hat{z}, \hat{p}_z\right] \tag{4.8}$$

$$= \hat{y}\hat{p}_x \left(-i\hbar\right) + \hat{x}\hat{p}_y \left(i\hbar\right) \tag{4.9}$$

$$= i\hbar \hat{L}_z \tag{4.10}$$

In a similar way (by cyclic permutation of x, y and z) you can evaluate both the other commutators of the components of angular momentum to find that

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar \hat{L}_z \tag{4.11}$$

$$\left[\hat{L}_y, \hat{L}_z\right] = i\hbar \hat{L}_x \tag{4.12}$$

$$\left[\hat{L}_z, \hat{L}_x\right] = i\hbar \hat{L}_y \tag{4.13}$$

Hence the commutators among different components of the angular momentum operator $\hat{\mathbf{L}}$ are *not* zero. A non-zero commutator between position and momentum operators guaranteed that there was an incompatibility in quantum mechanics between knowing both those quantities with precision. The same is true for the components of angular momentum: *We cannot measure more than one component of angular momentum with precision.* Put another way, functions that are simultaneously eigenfunctions of two different components of $\hat{\mathbf{L}}$, say, $\hat{L}_x$ and $\hat{L}_y$, do not exist.

4.1.2   Commutators with $\hat{\mathbf{L}}^2$

$\hat{\mathbf{L}}$ is a vector operator with three components. We can define a "dot product of vector operators" in the natural way: the dot product is the sum of the products of corresponding components. It is useful to define $\hat{\mathbf{L}}^2$ in that way: $\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$. (Just as the dot product of two vectors is a scalar, the dot product of two vector operators is a scalar operator: it represents one quantity, not three.) Now we show that each of the components of $\hat{\mathbf{L}}$ commutes with $\hat{\mathbf{L}}^2$. First examine the x-component.

$$\left[\hat{\mathbf{L}}^2, \hat{L}_x\right] = \left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right)\hat{L}_x - \hat{L}_x\left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right) \tag{4.14}$$

$$= \left(\hat{L}_y^2 + \hat{L}_z^2\right)\hat{L}_x - \hat{L}_x\left(\hat{L}_y^2 + \hat{L}_z^2\right) \tag{4.15}$$

$$= \hat{L}_y\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\hat{L}_z \tag{4.16}$$

$$= \hat{L}_y(\hat{L}_y\hat{L}_x) - (\hat{L}_x\hat{L}_y)\hat{L}_y + \hat{L}_z(\hat{L}_z\hat{L}_x) - (\hat{L}_x\hat{L}_z)\hat{L}_z \tag{4.17}$$

Eq. (4.17) differs from Eq. (4.16) only by parentheses added. I added the parentheses to highlight pairs of operators that I will replace using the commutation

relations proved in the previous section.

$$\begin{aligned}
\left[\hat{\mathbf{L}}^2, \hat{L}_x\right] &= \hat{L}_y(\hat{L}_x\hat{L}_y - i\hbar\hat{L}_z) - (\hat{L}_y\hat{L}_x + i\hbar\hat{L}_z)\hat{L}_y \\
&\quad + \hat{L}_z(\hat{L}_x\hat{L}_z + i\hbar\hat{L}_y) - (\hat{L}_z\hat{L}_x - i\hbar\hat{L}_y)\hat{L}_z \\
&= 0
\end{aligned} \tag{4.18}$$

By very similar means, we can show that $\hat{\mathbf{L}}^2$ also commutes with $\hat{L}_y$ and $\hat{L}_z$.

$$\left[\hat{\mathbf{L}}^2, \hat{L}_y\right] = \left[\hat{\mathbf{L}}^2, \hat{L}_z\right] = 0 \tag{4.19}$$

From the fact that $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do not commute with each other [Eqs.(4.11-4.13)] but they all separately commute with $\hat{\mathbf{L}}^2$, we conclude that in quantum mechanics *we can know the value of $\hat{\mathbf{L}}^2$ and only one of the components of $\hat{\mathbf{L}}$ simultaneously with precision.* By convention, the z-component of angular momentum is usually chosen as the special one paired with $\hat{\mathbf{L}}^2$, but this is only convention.

4.1.3   Eigenvalues of angular momentum

(This section, not covered in lecture, is here for background. You should be able to follow the arguments, but are not responsible for their workings. You need to be able to use the results, which are summarized in Section 4.1.4.)

Here we deduce the eigenvalues of $\hat{\mathbf{L}}^2$ and one of the components of $\hat{\mathbf{L}}$, which we take to be $\hat{L}_z$.

There must be some set of kets that are simultaneously eigenfunctions of both $\hat{\mathbf{L}}^2$ and $\hat{L}_z$. Their eigenvalues with respect to the two operators will probably be different; I don't know much about them, so I will simply write them as $|c\,b\rangle$. They must satisfy

$$\hat{\mathbf{L}}^2 |c\,b\rangle = c\,|c\,b\rangle \tag{4.20}$$

$$\hat{L}_z |c\,b\rangle = b\,|c\,b\rangle . \tag{4.21}$$

At this point, the eigenvalues c and b are unknown. However, we know that the magnitude of a vector must always be greater than the magnitude of any one of its components. Hence,

$$\left\langle \hat{\mathbf{L}}^2 \right\rangle \geq \left\langle \hat{L}_z \right\rangle^2 , \tag{4.22}$$

and therefore

$$c \geq 0 \text{ and } c \geq b^2 . \tag{4.23}$$

Raising and lowering operators for $\hat{L}_z$ eigenvalues

Now I will take a tack very similar to that I used in the harmonic oscillator problem. Define the raising operator,

$$\hat{L}_+ \equiv \hat{L}_x + i\hat{L}_y. \tag{4.24}$$

I can show that $\hat{L}_+ |c\, b\rangle$ is also an eigenstate of $\hat{L}_z$ with eigenvalue $b + \hbar$:

$$\hat{L}_z\left(\hat{L}_+ |c\, b\rangle\right) = \hat{L}_z\left(\hat{L}_x + i\hat{L}_y\right)|c\, b\rangle \tag{4.25}$$

Now apply commutation relation (4.13) in the form $\hat{L}_z\hat{L}_x = i\hbar\hat{L}_y + \hat{L}_x\hat{L}_z$, and commutation relation (4.12) in the form $\hat{L}_z\hat{L}_y = -i\hbar\hat{L}_x + \hat{L}_y\hat{L}_z$:

$$= \left(i\hbar\hat{L}_y + \hat{L}_x\hat{L}_z\right)|c\, b\rangle + i\left(-i\hbar\hat{L}_x + \hat{L}_y\hat{L}_z\right)|c\, b\rangle \tag{4.26}$$

Now apply $\hat{L}_z |c\, b\rangle = b|c\, b\rangle$ and collect terms:

$$= \hbar\left(\hat{L}_x + i\hat{L}_y\right)|c\, b\rangle + \left(\hat{L}_x + i\hat{L}_y\right)b|c\, b\rangle \tag{4.27}$$

and now recognize the definition of $\hat{L}_+$ to get

$$= (b + \hbar)\left(\hat{L}_+ |c\, b\rangle\right) \tag{4.28}$$

so we have proved that $\hat{L}_+$ is a raising operator: it raises the eigenvalue of $|c\, b\rangle$ with respect to $\hat{L}_z$ by $\hbar$.

The lowering operator with respect to the $\hat{L}_z$ eigenvalue is

$$\hat{L}_- \equiv \hat{L}_x - i\hat{L}_y. \tag{4.29}$$

By very similar manipulations to Eqs.(4.25-4.28), you can show that, if $\hat{L}_z |c\, b\rangle = b|c\, b\rangle$, then

$$\hat{L}_z\left(\hat{L}_- |c\, b\rangle\right) = (b - \hbar)\left(\hat{L}_- |c\, b\rangle\right), \tag{4.30}$$

confirming that $\hat{L}_-$ is a lowering operator.

Since

$$\left[\hat{\mathbf{L}}^2, \hat{L}_\pm\right] = \left[\hat{\mathbf{L}}^2, \hat{L}_x\right] \pm i\left[\hat{\mathbf{L}}^2, \hat{L}_y\right] = 0, \tag{4.31}$$

the raised or lowered states, $\hat{L}_\pm |c\, b\rangle$, are eigenfunctions of $\hat{\mathbf{L}}^2$ with the same $\hat{\mathbf{L}}^2$-eigenvalue as the original state $|c\, b\rangle$. (You can show that: write down $\hat{\mathbf{L}}^2\hat{L}_+ |c\, b\rangle$, switch the two operators (they commute), operate with $\hat{\mathbf{L}}^2$, and then factor out the eigenvalue.)

The eigenvalues of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ from commutation relations

Since $|b| < \sqrt{c}$, we cannot apply $\hat{L}_+$ or $\hat{L}_-$ indefinitely to obtain states with ever increasing or decreasing $\hat{L}_z$ eigenvalue. For a given c, there has to be a $b_{\min}$ and a $b_{\max}$. In the case of the maximum,

$$\hat{L}_+ |c\, b_{\max}\rangle = 0. \tag{4.32}$$

Operate on the left with $\hat{L}_-$,

$$\hat{L}_- \hat{L}_+ |c\, b_{\max}\rangle = 0. \tag{4.33}$$

Follow some algebra to obtain a relation between c and $b_{\max}$:

$$0 = \hat{L}_- \hat{L}_+ |c\, b_{\max}\rangle = \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right)|c\, b_{\max}\rangle \tag{4.34}$$

$$= \left\{\hat{L}_x^2 + \hat{L}_y^2 + i\left(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x\right)\right\}|c\, b_{\max}\rangle \tag{4.35}$$

$$= \left\{\hat{L}_x^2 + \hat{L}_y^2 - \hbar\hat{L}_z\right\}|c\, b_{\max}\rangle \tag{4.36}$$

$$= \left\{\hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z\right\}|c\, b_{\max}\rangle \tag{4.37}$$

$$= \left\{c - b_{\max}^2 - \hbar b_{\max}\right\}|c\, b_{\max}\rangle \tag{4.38}$$

Therefore

$$c = b_{\max}^2 + \hbar b_{\max}. \tag{4.39}$$

In a similar fashion, it follows from $\hat{L}_- |c\, b_{\min}\rangle = 0$ that

$$c = b_{\min}^2 - \hbar b_{\min}. \tag{4.40}$$

Subtract Eq.(4.40) from Eq.(4.39).

$$b_{\max}^2 + \hbar b_{\max} - b_{\min}^2 + \hbar b_{\min} = 0 \tag{4.41}$$

Regard this equation as a quadratic equation in $b_{\max}$ for a given $b_{\min}$ (or the other way around - it doesn't matter). The quadratic equation has two solutions, $b_{\max} = -b_{\min}$ and $b_{\max} = b_{\min} - \hbar$. The second solution is not relevant, since $b_{\max}$ is supposed to be bigger than $b_{\min}$. Therefore, the first solution is the one we want:

$$b_{\max} = -b_{\min}. \tag{4.42}$$

We should be able to turn $|c\, b_{\min}\rangle$ into $|c\, b_{\max}\rangle$ (within a normalization constant) by repeatedly applying the raising operator $\hat{L}_+$. Let's say, for a given value of c, it takes a total of n applications of $\hat{L}_+$ to turn $|c\, b_{\min}\rangle$ into $|c\, b_{\max}\rangle$.

$$(\hat{L}_+)^n |c\, b_{\min}\rangle \propto |c\, b_{\max}\rangle \tag{4.43}$$

Each time $\hat{L}_+$ is applied, the $\hat{L}_z$ eigenvalue is increased by $\hbar$. Therefore

$$b_{\text{max}} - b_{\text{min}} = n\hbar, \quad n = 0, 1, 2, \ldots, \tag{4.44}$$

or by substituting $b_{\text{max}} = -b_{\text{min}}$ into this equation,

$$b_{\text{max}} = \frac{1}{2} n\hbar, \quad n = 0, 1, 2, \ldots \tag{4.45}$$

$$\equiv j\hbar, \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots, \tag{4.46}$$

where $\left(j = \frac{n}{2} \right)$.

And since $b_{\text{max}} = -b_{\text{min}}$,

$$b_{\text{min}} = -j\hbar, \; j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{4.47}$$

Finally, upon substitution of $b_{\text{max}} = j\hbar$ into Eq.(4.39), we obtain

$$c = b_{\text{max}}^2 + \hbar b_{\text{max}} = j^2 \hbar^2 + j\hbar^2 = j(j+1)\hbar^2 \tag{4.48}$$

4.1.4   Summary

We have shown that the commutation relations among $\hat{\mathbf{L}}^2, \hat{L}_x, \hat{L}_y$, and $\hat{L}_z$ imply the existence of eigenstates $\left| j\, m \right\rangle$ with the properties

$$\hat{\mathbf{L}}^2 \left| j\, m \right\rangle = \hbar^2 j(j+1) \left| j\, m \right\rangle \tag{4.49}$$

$$\hat{L}_z \left| j\, m \right\rangle = \hbar m \left| j\, m \right\rangle \tag{4.50}$$

where

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{4.51}$$

$$m = -j, -j+1, \ldots j. \tag{4.52}$$

The raising and lowering operators are

$$\hat{L}_+ \equiv \hat{L}_x + i\hat{L}_y \tag{4.53}$$

$$\hat{L}_- \equiv \hat{L}_x - i\hat{L}_y \tag{4.54}$$

and, inverting those,

$$\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-) \tag{4.55}$$

$$\hat{L}_y = \frac{1}{2i}(\hat{L}_+ - \hat{L}_-) \tag{4.56}$$

The action of $\hat{L}_+$ and $\hat{L}_-$ on the $\left| j\, m \right\rangle$ is

$$\hat{L}_+ \left| j\, m \right\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\left| j\, m+1 \right\rangle \tag{4.57}$$

$$\hat{L}_- \left| j\, m \right\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\left| j\, m-1 \right\rangle \tag{4.58}$$

So, for example,

$$\hat{\mathbf{L}}^2 \left| 3, -1 \right\rangle = 12\hbar^2 \left| 3, -1 \right\rangle \tag{4.59}$$

$$\hat{L}_z \left| 3, -1 \right\rangle = -\hbar \left| 3, -1 \right\rangle \tag{4.60}$$

$$\hat{L}_+ \left| 3, -1 \right\rangle = \hbar\sqrt{12} \left| 3, 0 \right\rangle \tag{4.61}$$

$$\hat{L}_x \left| 3, -1 \right\rangle = \frac{\hbar}{2}\left(\sqrt{12}\left| 3, 0 \right\rangle + \sqrt{10}\left| 3, -2 \right\rangle\right) \tag{4.62}$$

Table 4.1 Angular momentum operators.

Type	Operator	Quantum numbers
General (unspecified)	$\hat{\mathbf{J}}$	j, m
Molecular rotation	$\hat{\mathbf{J}}$	j, m
Electronic orbital motion	$\hat{\mathbf{L}}$	l, m_l
Electron spin	$\hat{\mathbf{S}}$	s, m_s
Nuclear spin	$\hat{\mathbf{I}}$	I, m_I

4.1.5 Types of angular momentum

Many forms of angular momentum appear in quantum mechanics, and there are some loose conventions about the letters used to identify each; the most common ones are listed in Table 4.1. Each of these has its own set of raising and lowering operators and so on, and all of them behave in exactly the same way.

The spin quantum number s or I is fixed for a particular kind of particle; electrons are always $s = \frac{1}{2}$, protons always $I = \frac{1}{2}$ (sometimes written s), deuterons always $I = 1$, and so on.

For rotational motions (electron orbital motion, molecular rotation) the j or l quantum number is always an integer. Spin quantum numbers can be either integer or half integer.

4.1.6 Properties of Rotational Angular Momentum States

When we applied the harmonic oscillator model to molecular vibrations, we Levine §18.14
transformed the Hamiltonian from one set of coordinates (x_A and x_B) to another set (x and X). In problems that have spherical symmetry (including free rotations of molecules, and motions of electrons in atoms) it is useful to use spherical polar coordinates. Such coordinates often permit the separation of a complicated quantum problem into several simpler ones, just as happened in the center-of-mass/relative coordinate transformation. There is probably a section in your text that describes spherical polar coordinates. (There is probably also a section in your calculus book, but note that mathematicians use a different convention from physicists and chemists for the names of the angles. For us, θ is the polar angle giving the angle between a vector and the z axis, while φ is the aximuthal angle describing the displacement away from the x axis in the x–y plane; math books usually reverse the definition.)

Converting differential operators into spherical coordinates is a straightfor-

ward but tedious exercise; I'll just give the results here. The $\hat{\mathbf{L}}^2$ operator becomes

$$\hat{\mathbf{L}}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) \tag{4.63}$$

This operator looks fearsome but once again is really just a prescription for a series of multiplications and differentiations. Note that it depends only on the angles θ and ϕ, the normal spherical polar angles. The polar radius r does not appear.

$\hat{L}_z$ in spherical polar coordinates is

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial\phi}. \tag{4.64}$$

Notice that it contains only ϕ. The component of the angular momentum around the z axis just refers to rotation about that axis, so ϕ is the only relevant coordinate.

The eigenvalues of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ were deduced from commutation relations above. For rotational motion, we can also find them by finding the eigenvalues and eigenfunctions (which will be functions of θ and ϕ) of the $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ operators. That means finding functions $Y_{lm}(\theta,\phi)$ that solve the differential equations

$$\hat{\mathbf{L}}^2 Y_{lm}(\theta,\phi) = -\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) Y_{lm}(\theta,\phi)$$
$$= (\text{eigenvalue of } \hat{\mathbf{L}}^2) \, Y_{lm}(\theta,\phi) \tag{4.65}$$
$$\hat{L}_z Y_{lm}(\theta,\phi) = \frac{\hbar}{i} \frac{\partial}{\partial\phi} Y_{lm}(\theta,\phi)$$
$$= (\text{eigenvalue of } \hat{L}_z) \, Y_{lm}(\theta,\phi) \tag{4.66}$$

with the boundary conditions that $Y_{lm}(\theta,\phi)$ is normalizable,

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta \, \left| Y_{lm}(\theta,\phi) \right|^2 = 1 \,, \tag{4.67}$$

and single-valued. The latter means that if you change θ and ϕ so that you travel in a circle back to your starting point, $Y_{lm}(\theta,\phi)$ remains the same. For example,

$$Y_{lm}(\theta,\phi+2\pi) = Y_{lm}(\theta,\phi) \,. \tag{4.68}$$

This is the condition that knocks out the half-integral values of j for rotational angular momentum. Recall that both half-integral and integral values of j are allowed for spin angular momentum; that is why the commutator-style analysis is the more general one.

The Y_{lm}

The functions $Y_{lm}(\theta, \phi)$ are well known. They are called the *spherical harmonics*. For each l, m pair the spherical harmonic is the product of a complex-exponential function of ϕ with a real function of θ. There is a funny notation for the different angular momentum states, which has its origins in nineteenth-century atomic spectroscopy: angular momentum states with $l = 0, 1, 2, 3, 4, \ldots$ are traditionally called $s, p, d, f, g, \ldots$ states. They are often written with a subscript specifying the value of m, so, for example, the "spectroscopic notation' for Y_{21} is d_1. These names are the ones you know and love for the H-atom orbitals; the angular part of each H-atom wavefunction is just one of (or a linear combination of) the Y_{lm}. Let's examine the first few functions.

Since $-l \le m \le l$, there is only one s-state (with $m = 0$).

$$Y_{00}(\theta, \phi) = \sqrt{\frac{1}{4\pi}} \tag{4.69}$$

This function is a constant; it looks the same in every direction from the center (it is *spherically symmetric*). It has no nodes. In general we will find that the angular function Y_{lm} has a total of l angular nodes.

For $l = 1$, m can be -1, 0, or 1, so there are three p-states:

$$Y_{11}(\theta, \phi) = -\sqrt{\frac{3}{4\pi}} \sqrt{\frac{1}{2}} \sin\theta \, e^{i\phi}, \tag{4.70}$$

$$Y_{10}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos\theta. \tag{4.71}$$

$$Y_{1,-1}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \sqrt{\frac{1}{2}} \sin\theta \, e^{-i\phi}. \tag{4.72}$$

Representations of these functions are shown in Figure 4.1.

The p_0 state is maximum at the "poles", zero around the "equator", positive in the northern hemisphere, and negative in the southern hemisphere. The other p states are zero at the poles (because of the $\sin\theta$ term), but both their real and imaginary parts oscillate around the z axis. The combination of real and imaginary parts corresponds to counterclockwise rotation around the z axis for Y_{11} and to clockwise rotation around z for Y_{1-1}. (Recall that in one dimension we found that to have nonzero momentum, a wavefunction had to have both a real and an imaginary part. The same is true here, for *angular* momentum around z.)
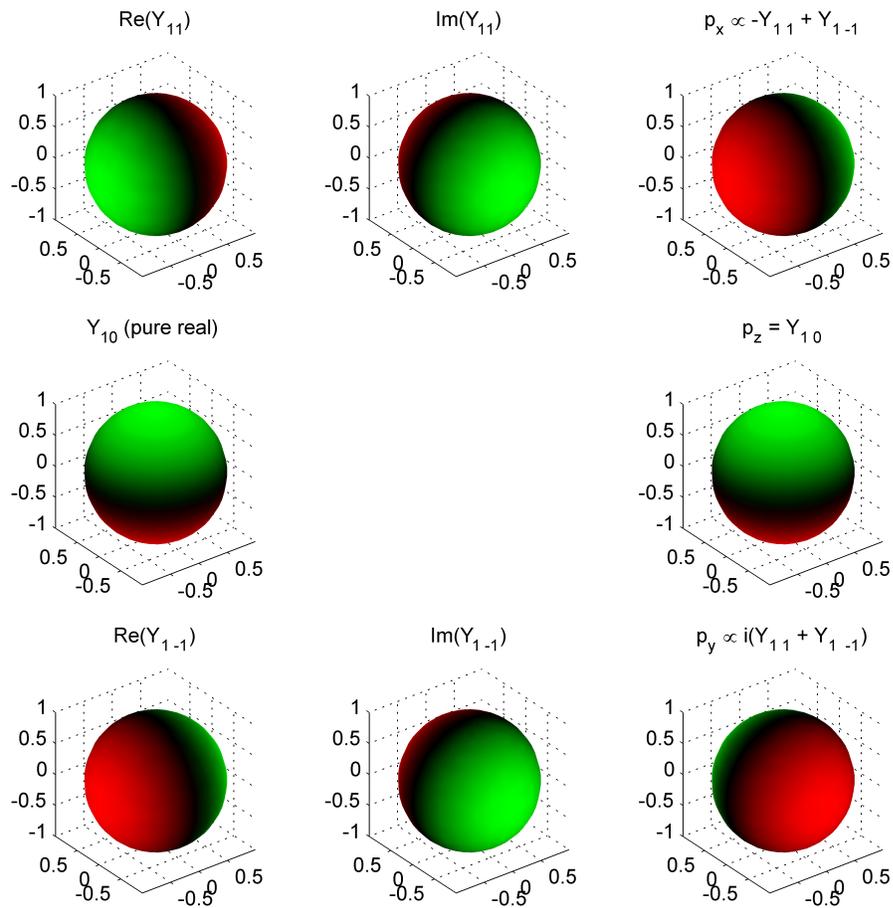
Figure 4.1: Color-on-sphere representations of the $l = 1$ spherical harmonics. The left column gives the real part and the center column gives the imaginary part of each, with green representing positive values, black zero, and red negative. The right column shows the real linear combinations of the Y_{lm} often used to depict p orbitals. These functions, which are eigenfunctions of $\hat{L}^2$ but not $\hat{L}_z$, are purely real.

You are used to seeing linear combinations of the $m = \pm 1$ states that are real:

$$p_x = \sqrt{\frac{1}{2}}\left[-Y_{11} + Y_{1,-1}\right] = \sqrt{\frac{3}{4\pi}}\sin\theta\,\cos\phi = \sqrt{\frac{3}{4\pi}}\frac{x}{r} \tag{4.73}$$

$$p_y = i\sqrt{\frac{1}{2}}\left[Y_{11} + Y_{1,-1}\right] = \sqrt{\frac{3}{4\pi}}\sin\theta\,\sin\phi = \sqrt{\frac{3}{4\pi}}\frac{y}{r} \tag{4.74}$$

The Y_{10} state is the p_z without modification.

$$p_z = Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r}. \tag{4.75}$$

You can verify that the p_x and p_y states are still eigenfunctions of $\hat{\mathbf{L}}^2$, but are not eigenfunctions of $\hat{L}_z$. (That is much easier in Dirac notation than in function notation!) The p-states each have one angular node. For example, the state $Y_{10} = p_z$ is zero when $\theta = \frac{\pi}{2}$, that is, when $z = 0$ so that you are looking in the xy plane.

The five d-states are

$$Y_{20}(\theta,\phi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta - 1) \tag{4.76}$$

$$Y_{2\pm 1}(\theta,\phi) = \mp\frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta\,\cos\theta\,e^{\pm i\phi} \tag{4.77}$$

$$Y_{2\pm 2}(\theta,\phi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\,e^{\pm 2i\phi} \tag{4.78}$$

Except for Y_{20}, the d-states are complex. It is traditional to form linear combinations of d-states that are real: $d_{z^2}, d_{xy}, d_{yz}, d_{x^2-y^2}$ and d_{xy}. The $l = 2$ spherical harmonics and the familiar real combinations are shown in Figure 4.2. Each $l = 2$ function has two angular nodes.

Orthonormality of the Y_{lm}

The spherical harmonics are orthogonal, as they must be (they are special cases, for integral j, of the kets $\left|j\,m\right\rangle$.) They are also normalized. Since they are defined in terms of spherical polar coordinates, we must use the spherical polar "area element" $\sin\theta\,d\theta$ in doing the integrals. Orthonormality therefore means, in the function notation,

$$\int_0^{2\pi} d\phi \int_0^{\pi}\sin\theta\,d\theta\,Y_{l'm'}^*(\theta,\phi)Y_{lm}(\theta,\phi) = \begin{cases} 1 & \text{if } l = l' \text{ and } m = m', \\ 0 & \text{otherwise.} \end{cases} \tag{4.79}$$

In Dirac notation the angular momentum states behave just as you would expect: $\left\langle j\,m | j'\,m'\right\rangle = 1$ if $j = j'$ and $m = m'$, and 0 if either of the two quantum numbers doesn't match.
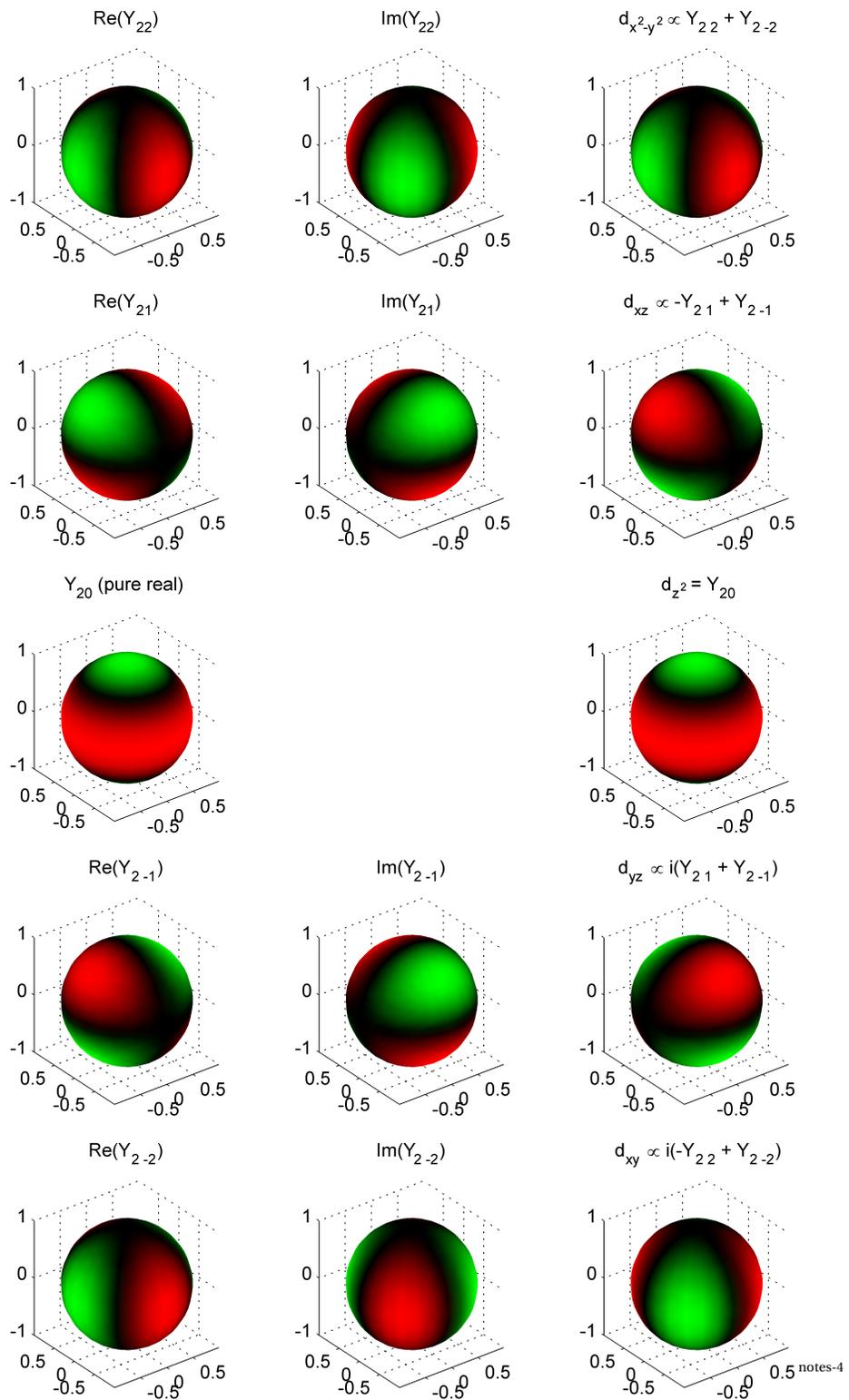
Figure 4.2: Color-on-sphere representations of the $l = 2$ spherical harmonics, as in Figure 4.1.

4.1.7   The rigid rotor

Let's consider a simple model for *rotation* of a linear molecule. This might be   Levine §18.14
either a diatomic molecule or a linear polyatomic like CO_2, OCS, or C_2H_2. In this
model we will ignore the vibration of the molecule, treating the chemical bonds
as stiff, massless rods. In this case there is no potential energy to worry about
since the relative distance between the atoms cannot change, so we need worry
only about kinetic energies. The atoms will rotate about their common center
of mass. For diatomic problems, the same center-of-mass transformation we
used for converting the vibrational problem to an effective one-particle problem
works here, with the same reduced mass μ.

Recall that in classical mechanics we could write the kinetic energy of linear
motion either in terms of velocity, $T = \frac{1}{2}mv^2$, or in terms of momentum, $T = \frac{p^2}{2m}$.
Similar formulas apply to rotational motion, where the role of the mass is played
by the *moment of inertia, I*. For a diatomic molecule $I = \mu r_e^2$ where r_e (a constant)
is the separation between the two atoms; for linear polyatomics the formula is
$I = \sum_i m_i r_i^2$ where the sum is over all atoms and r_i is the distance of atom i from
the molecule's center of mass. We can write the rotational kinetic energy either
as $T_R = \frac{1}{2}I\omega^2$, where ω is the rotational speed in radians/s, or as $T_R = \frac{\mathbf{L}^2}{2I}$, where
$\mathbf{L}$ is the angular momentum. As usual, we will convert to quantum mechanical
operators using the momentum form. Since there is no potential energy, the
Hamiltonian is just the kinetic energy operator and is given by

$$\hat{H} = \frac{\hat{\mathbf{L}}^2}{2I} \tag{4.80}$$

The Schrödinger equation becomes

$$\frac{\hat{\mathbf{L}}^2}{2I}\psi(\theta,\phi) = E\psi(\theta,\phi) \tag{4.81}$$

But this Hamiltonian is just a constant times $\hat{\mathbf{L}}^2$, and we know the eigenfunctions
of $\hat{\mathbf{L}}^2$; they are the $Y_{lm}(\theta,\phi)$. So the energy eigenfunctions of two particles con-
nected by a massless, stiff rod (a rigid rotor) are exactly the spherical harmonics,
the $Y_{lm}(\theta,\phi)$! This is the simplest model of the rotational motion of a linear
molecule. In this application, the angles θ and ϕ describe the orientation of the
molecule in space.

What are the eigenvalues? We have

$$\hat{H}Y_{lm}(\theta,\phi) = \frac{1}{2I}\hat{\mathbf{L}}^2 Y_{lm}(\theta,\phi) \tag{4.82}$$

$$= \frac{\hbar^2}{2I}l(l+1)Y_{lm}(\theta,\phi) \tag{4.83}$$

so the energy eigenvalues are simply $E_l = \frac{\hbar^2 l(l+1)}{2I}$. In spectroscopy it is traditional to use J instead of l and to define

$$B = \frac{\hbar^2}{2I},$$
(4.84)

the *rotational constant*. The version of this expression specifically for diatomic molecules is, since $I = \mu r_e^2$,

$$B = \frac{\hbar^2}{2\mu r_e^2}.$$
(4.85)

The energies of the rotating molecule are then

$$E_J = BJ(J+1)$$
(4.86)

Typically this equation is divided by hc on both sides, and both the rotational energy E_J and the rotational constant B are reported in cm^{-1} instead of in joules. Pedants use the symbol $\tilde{B}$ for B in cm^{-1} (since it is really B/hc), but most spectroscopists just call it B and expect you to know they are using cm^{-1} as a unit of energy.

Note that these energies are *not* equally spaced; the spacings increase as J gets larger. For a molecule with $B = 2\,$cm^{-1}, the energies for $J = 0$–4 are 0, 4, 12, 24, and 40 cm^{-1}.

It's worthwhile to stop and think for a bit about what the eigenfunctions for a rigid rotor mean. A molecule in the Y_{00} state has no rotational energy ($BJ(J+1) = 0$), so classically it is not rotating. On the other hand, its wavefunction has the same value in every direction: it has no well defined "orientation" in space. A $J = 0$ molecule is a fuzzy thing, having no rotational energy but also not being stationary in the classical sense.

There's no way in quantum mechanics to measure directly how a molecule is oriented, because there is no angle operator that is equivalent to the position operator, for example. What we have to do is measure the components of some vector whose orientation we know with respect to the internuclear axis. For instance, in linear molecules the dipole moment (actually its expectation value) must lie along the molecular axis, and there are experiments involving electric fields that measure one component of the dipole moment.

Selection rules for the rigid rotor

Levine §21.4   A heteronuclear diatomic molecule (modeled for the moment as a rigid rotor) can have a permanent dipole moment μ_0, which must lie along the internuclear axis.

If we apply an electric field to the molecule, the Hamiltonian of the interaction is the same one we encountered before:

$$\hat{H}_1 = \mathbf{E} \cdot \mu \tag{4.87}$$

To find what transitions between different angular momentum states can be induced by the applied electric field, we must find the nonzero matrix elements of $\hat{H}_1$. That is, if we are to observe a transition from $|j\,m\rangle$ to $|j'\,m'\rangle$, then $\langle j'\,m'|\hat{H}_1|j\,m\rangle$ must not be zero.

To write the dipole moment (which must lie along the internuclear axis) in terms of the angles θ and ϕ, we need the transformation between spherical polar and Cartesian coordinates:

$$x = r\sin\theta\cos\phi \tag{4.88}$$

$$y = r\sin\theta\sin\phi \tag{4.89}$$

$$z = r\cos\theta \tag{4.90}$$

The components of the dipole moment are therefore $\mu_x = \mu_0\sin\theta\cos\phi$, and so on.

In terms of the angles θ and ϕ, the kets $|j\,m\rangle$ are just the spherical harmonics $Y_{jm}(\theta,\phi)$. We can therefore have transitions if

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta\, Y_{j'm'}^*(\theta,\phi) E_z \mu_0 \cos\theta\, Y_{jm}(\theta,\phi) \neq 0 \tag{4.91}$$

and similarly for the x and y components of the electric field, using the expressions above for the different components of μ.

The function $\cos\theta$ acts like a combination of raising and lowering operators on j; that is,

$$\cos\theta\, Y_{jm}(\theta,\phi) = f(j,m)\, Y_{j+1\,m}(\theta,\phi) + g(j,m)\, Y_{j-1\,m}(\theta,\phi) \tag{4.92}$$

where f and g are expressions involving j and m but not the angles. (For a proof of this statement, see Cohen-Tannoudji et al, *Quantum Mechanics*, p. 683.) Since the Y_{jm} are orthogonal, this expression immediately gives us the selection rules (for the z-component of the electric field) $\Delta j = \pm 1$, $\Delta m = 0$. The x and y components behave similarly but m gets raised and lowered in addition to j. So for a rigid rotor with a permanent dipole moment excited by an electric field, the overall selection rules are

$$\Delta j = \pm 1 \tag{4.93}$$

$$\Delta m = 0, \pm 1 \tag{4.94}$$

Transitions between angular momentum states of molecules are called "pure rotational" transitions. The discussion above shows that for such transitions to occur the molecules must have permanent dipole moments. Homonuclear diatomic molecules, and in fact any molecules that have centers of symmetry (such as CO_2), therefore do not have pure rotational transitions.

Typical rotational constants B for molecules range from $60\,\text{cm}^{-1}$ for H_2 to small fractions of cm^{-1} for large molecules. The energy difference between two rotational states of a molecule with quantum numbers J and $J+1$ is (in cm^{-1})

$$\Delta E = B(J+1)(J+2) - B(J)(J+1) = 2B(J+1) \tag{4.95}$$

For $B = 2\,\text{cm}^{-1}$, typical of second-row diatomics like CO and NO, and $J = 9$, we expect light of wavenumber $40\,\text{cm}^{-1}$; the wavelength is $1/40 = 0.025\,\text{cm}$ and the frequency is $1.2\times10^{12}\,\text{Hz}$. This light is in the microwave frequency region, and microwave spectroscopy of gases is among the most precise tools in science.

4.2   Rotation: the central field problem

In applying the harmonic oscillator model to the molecular vibration problem, you performed a coordinate transformation from coordinates of individual particles to center-of-mass and relative coordinates. That same transformation, still for two particles, works exactly the same in three dimensions as it does in one, and if the potential energy depends (as before) only on the distance between the two particles, we can separate the Hamiltonian into a sum of a center-of-mass and a relative part:

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(r) \tag{4.96}$$

where

$$\hat{\mathbf{p}}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right). \tag{4.97}$$

and similarly for $\hat{\mathbf{P}}^2$. The sum of second partial derivative operators is called the Laplacian.

Just as in the one-dimensional case, with a Hamiltonian that is a sum of two independent pieces (with different coordinates), the wavefunction is a product of two pieces, one describing the overall motion of the two-body system through space and the other describing the relative motion. Let me now concentrate on the relative motion.

Because the potential $V(r)$ depends only on the distance r (which equals $\sqrt{x^2 + y^2 + z^2}$) between the two particles, it is helpful to switch to spherical polar coordinates. I described this procedure in general terms before. The Hamiltonian of the relative motion becomes

$$\hat{H} = \left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2\mu r^2} \hat{\mathbf{L}}^2 + V(r) \right], \qquad (4.98)$$

where

$$\hat{\mathbf{L}}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot\theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right) \qquad (4.99)$$

is the same differential operator $\hat{\mathbf{L}}^2$ we studied as part of the treatment of angular momentum.

All the angular dependence of the Hamiltonian appears in $\hat{\mathbf{L}}^2$, and we know what its eigenfunctions are already: they are the $Y_{lm}(\theta, \phi)$. We might be able to separate the overall wavefunction of the system into a product of pieces that depend only on r and only on angles. The angular part will be just the $Y_{lm}(\theta, \phi)$. Therefore, let's try

$$\psi(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi). \qquad (4.100)$$

The eigenvalue equation for $\hat{H}$ (the time-independent Schrödinger equation) is then

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2\mu r^2} \hat{\mathbf{L}}^2 + V(r) \right] R(r) Y_{lm}(\theta, \phi) = E R(r) Y_{lm}(\theta, \phi) \quad (4.101)$$

Since $Y_{lm}(\theta, \phi)$ is an eigenfunction of $\hat{\mathbf{L}}^2$ with eigenvalue $\hbar^2 l(l+1)$, and is not affected by any of the r operations, this immediately becomes

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] R(r) Y_{lm}(\theta, \phi) = E R(r) Y_{lm}(\theta, \phi) \quad (4.102)$$

and now Y_{lm} divides out on both sides to give us a differential equation for the $R(r)$ function:

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] R_{nl}(r) = E R_{nl}(r) \qquad (4.103)$$

Notice that I have put quantum numbers n and l on the R function. The n number is the one that will describe quantization in r, and the l number is necessary because l appears in Eq. (4.103) so that we must find a different $R(r)$ for every l.

This "radial equation" appears in every two-body potential problem. With $V(r) = \frac{1}{2}k(r - r_e)^2$, it describes the vibration of a diatomic molecule (we'll see how to get there shortly). With $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$, it describes the hydrogen atom. With other forms of V it describes the behavior of each electron in a multielectron atom in the Hartree-Fock approximation. Before examining these cases let us see how we can describe simple molecular rotation with only the angular part of the problem.

4.2.1 Vibrational motion in 3D problems

I claimed that the radial equation, Eq. (4.103), describes vibrational motion of molecules if the vibrational potential is inserted for $V(r)$. On the other hand, that equation does not look very much like the simple one-dimensional Schrödinger equation we solved during our study of harmonic oscillators. There are a couple of additional steps required to go from the radial equation to the simple harmonic description of molecular vibration.

First, let's apply a mathematical educated guess, and search for radial wave-functions $R(r)$ of the form

$$R_{nl}(r) = \frac{1}{r}u_{nl}(r) \tag{4.104}$$

In your homework, you will show that substituting that form of $R(r)$ into the radial equation gives

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u_{nl}(r) = E_{nl}u_{nl}(r). \tag{4.105}$$

This is now a one-dimensional equation for $u(r)$. If $l = 0$ (that is, the molecule is not rotating), it looks exactly like the harmonic oscillator problem if $V(r) = \frac{1}{2}k(r - r_e)^2$. You can therefore use all the intuition and tools you have developed for one-dimensional potential problems (notably, the harmonic oscillator and particle in a box) and apply them to the radial parts of three-dimensional problems.

If $l \neq 0$, then we end up with a one-dimensional problem that has an effective potential $V_{\text{eff}}(r)$ that looks like the real potential $V(r)$ plus a "centrifugal potential" piece:

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2} \tag{4.106}$$

The centrifugal term, $\frac{\hbar^2 l(l+1)}{2\mu r^2}$, represents the rotational kinetic energy, and you can see that if r was fixed at r_e it would just add the usual rigid-rotor rotational energy $Bl(l + 1)$ to the Hamiltonian, where B was defined in Eq. (4.85). That idea leads to the "rigid rotor-harmonic oscillator approximation", where we solve

the vibrational problem while ignoring the centrifugal term (yielding the usual vibrational energies $\left(n + \frac{1}{2}\right)\hbar\omega$ if we take $V(r) = \frac{1}{2}k(r - r_e)^2$), and then adding the rotational energy $Bl(l+1)$ corresponding to a *fixed* internuclear distance r_e to the vibrational energy to find the total. In that approximation we have

$$\frac{E(n, j)}{hc} = \left(n + \frac{1}{2}\right)\omega_e + BJ(J + 1), \tag{4.107}$$

where the usual "harmonic frequency" $\omega_e = \frac{\hbar\omega}{hc}$ and $B = \frac{\hbar^2}{2\mu r_e^2 hc}$, both with units of cm^{-1}. In spectroscopic discussions the vibrational quantum number n is often written v.

In real molecules r is *not* fixed. The centrifugal term adds an r-dependent piece to the vibrational potential, making it a little steeper at small r. This change in potential corresponds to a "centrifugal force"; as the rotational quantum number increases, the atoms appear to be thrown apart, just as two balls connected by a spring will spread apart if they are spun. That means the vibrational wavefunctions will no longer be exactly the harmonic oscillator functions, and they will depend on the rotational quantum number l (traditionally called J in this context.). However, for most molecules, the variation in the centrifugal term with r is small compared to the regular vibrational potential energy, and the change in the vibrational wavefunctions is not very big. There is no analytic expression for the vibrational wavefunctions $u_{nl}(r)$ or the energies E_{nl} when $l > 0$ for most $V(r)$), but they can be computed numerically to very high accuracy without difficulty. One widely used program for this purpose is the LEVEL program of R. J. Le Roy (`http://scienide2.uwaterloo.ca/~rleroy/level/`).

To construct an example, I selected a molecule with a weaker bond than is typical (I_2), and computed the intermolecular and centrifugal potentials and accurate $n = 0$ vibrational wavefunctions for both $J = 0$ and $J = 100$. Rather than a harmonic potential for $V(r)$, I used a more realistic model function called a "Morse potential" that has the correct qualitative shape. The potentials and the two wavefunctions are shown in Figure 4.3. You can see that the effect of the centrifugal term is to shift the vibrational potential upward (by roughly $BJ(J + 1)$) and slightly outward. The vibrational wavefunction therefore shifts slightly outward as well. The average bond distance $\langle\phi_0|r|\phi_0\rangle$ therefore gets a little larger, and the effective rotational constant B a little smaller, so the rotational levels are not quite as far apart at high J as you would expect from the low-J rotational constant. This small change in B as J increases is usually accounted for by including a "centrifugal distortion term" in the expression for rotational energy levels. Rather than the rigid rotor expression $E_J/hc = BJ(J + 1)$, people
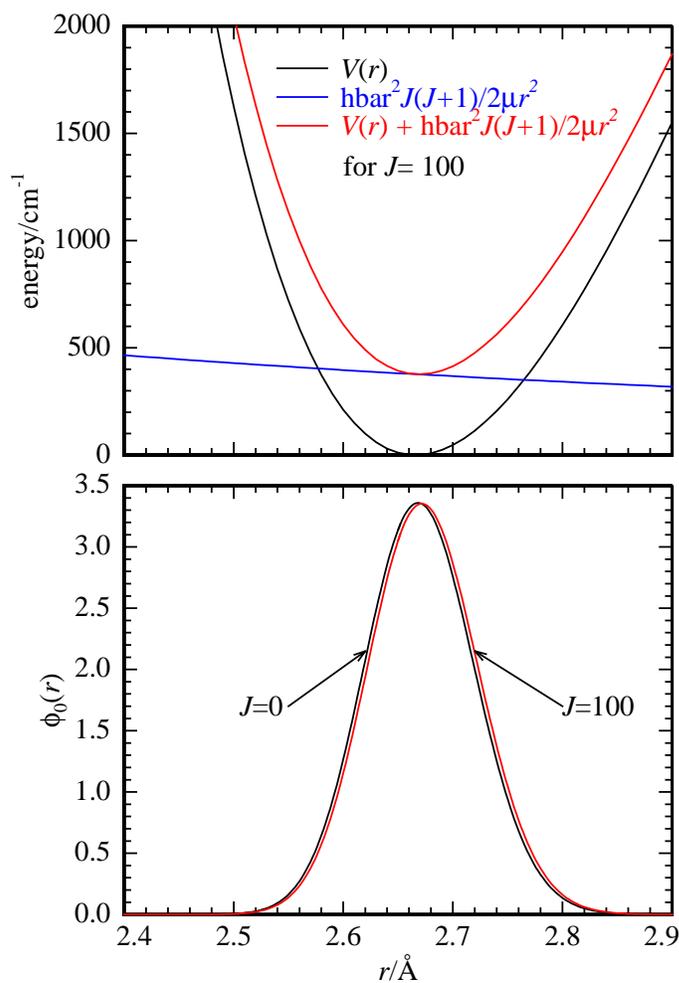
Figure 4.3: Intermolecular, centrifugal, and effective potentials (upper panel) and vibrational wavefunctions (lower panel) for I_2 in its ground vibrational state with rotational quantum numbers $J = 0$ and $J = 100$. The wavefunctions were computed using the LEVEL program of R. J. Le Roy. The vibrational potential $V(r)$ was a Morse function with $r_e = 2.666$Å, dissociation energy $12\,542\,\text{cm}^{-1}$, and Morse frequency parameter $\beta = 1.852\text{Å}^{-1}$. The harmonic spacing $\frac{\hbar\omega}{hc} = \omega_e$ of I_2 is about $214\,\text{cm}^{-1}$.

use the more accurate formula

$$\frac{E_J}{hc} = BJ(J+1) - D[J(J+1)]^2 \tag{4.108}$$

where D, the *centrifugal distortion constant*, has units of cm^{-1} and is usually provided in tables of molecular constants like those you have been using. For I_2, $B = 0.03737\,cm^{-1}$ and $D = 4.25 \times 10^{-9}\,cm^{-1}$. The rigid-rotor energy at $J = 100$ is therefore $E_J/hc = BJ(J+1) = 377.4\,cm^{-1}$, and the centrifugal correction is $D[J(J+1)]^2 = 0.433\,cm^{-1}$.

Actually, this trick of looking for $u_{nl}(r)$ instead of $R_{nl}(r)$ to get a one-dimensional Schrödinger equation with a centrifugal term works for all kinds of radial potentials. Let us turn next to the hydrogen atom and use it there.

4.3   Hydrogen atom

Eq. (4.103) describes hydrogenlike atoms if we use $\hspace{4cm}$ Levine §19.3

$$V(r) = \frac{1}{4\pi\epsilon_0} \frac{q_e q_N}{r} \tag{4.109}$$

$$= \frac{-Ze^2}{4\pi\epsilon_0} \frac{1}{r}, \tag{4.110}$$

where q_e and q_N are the charges on the electron and nucleus, Z is the atomic number, e is the unit of charge (1.6×10^{-19} Coulombs), and ϵ_0 is the permittivity of free space; $4\pi\epsilon_0 = 1.11265 \times 10^{-10}\,C^2\,J^{-1}\,m^{-1}$. (A hydrogenlike atom is any atom with only a single electron: H, He^+, Li^{2+}, etc.)

Making the substitution $R_{nl}(r) = \frac{1}{r} u_{nl}(r)$, and plugging that into Eq. (4.103), we find after a little work (which you do in your homework)

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \right] u_{nl}(r) = E_{nl} u_{nl}(r) \tag{4.111}$$

This looks like a one-dimensional Schrödinger equation for a particle of mass $\mu = \frac{m_e m_p}{m_e + m_p} \approx m_e$ in a potential

$$V_{\text{eff}}(r) = \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \tag{4.112}$$

If $l \neq 0$, the first (centrifugal) term gives a positive potential that becomes large as $r \to 0$; it is therefore difficult for the electron to approach close to the nucleus if $l \neq 0$ (that is, if it is in anything other than an s orbital.) The centrifugal term

is much more important in the H atom problem than it was in the rotation-vibration problem we just studied, because the reduced mass of the electron that appears in the denominator is much smaller than the reduced mass of a nucleus, and because the normal vibrational potential keeps the atoms apart from one another anyway so the region of small r where the centrifugal term diverges is not accessible. The second (Coulomb) term is negative and becomes more negative as $r \to 0$; that provides a "binding energy", that it, it holds the electron near the nucleus so long as the overall electron energy is small enough that it cannot escape.

These effective radial potentials are shown for $l = 0$, 1, and 2 in Figure 4.4. You can see that the $l = 0$ potential is qualitatively different from the others; it diverges toward negative values as the electron approaches the nucleus because of the $-1/r$ term in Eq. (4.112), while when $l > 0$ the curves diverge toward positive values at small r because of the $1/r^2$ term. The first few energy levels are also shown in each potential (though I have not yet shown you how those were obtained.) Notice that the *second* energy level for the $l = 0$ curve lines up with the *first* level for the $l = 1$ curve. This pattern continues as you go up in energy; the third level for $l = 0$ lines up with the second for $l = 1$ and the first for $l = 2$. We therefore use the same quantum number n to label all the levels with the same energy, though when $l = 1$ the lowest n we can have is $n = 2$, and so on. This lining up of the energy levels in the various different potentials is something of a freak of nature, present only for the special $V(r) \propto 1/r$ potential, and is called an "accidental degeneracy."

Eq. (4.111) can be solved by the standard method of power series expansion, taught in courses on differential equations. You may recall that in the harmonic oscillator case, we found solutions that looked like a polynomial in x times e^{-cx^2}, and the polynomials were called Hermite polynomials. The H-atom case is similar; the solutions are polynomials in r times e^{-cr}. (Note that the exponential function is a simple decaying exponential, not a Gaussian.) The polynomials for the H-atom problem are related to "associated Laguerre polynomials".

Each H-atom stationary state wavefunction is given by a normalization constant, times a polynomial in r with maximum power $n-1$, times $e^{-Zr/na}$, times a spherical harmonic function of θ and ϕ. We write

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi). \tag{4.113}$$

The functions $R_{nl}(r)$ for $n = 1$, 2, and 3 are listed in Table 4.2. They are written in terms of the quantity

$$a = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2}, \tag{4.114}$$
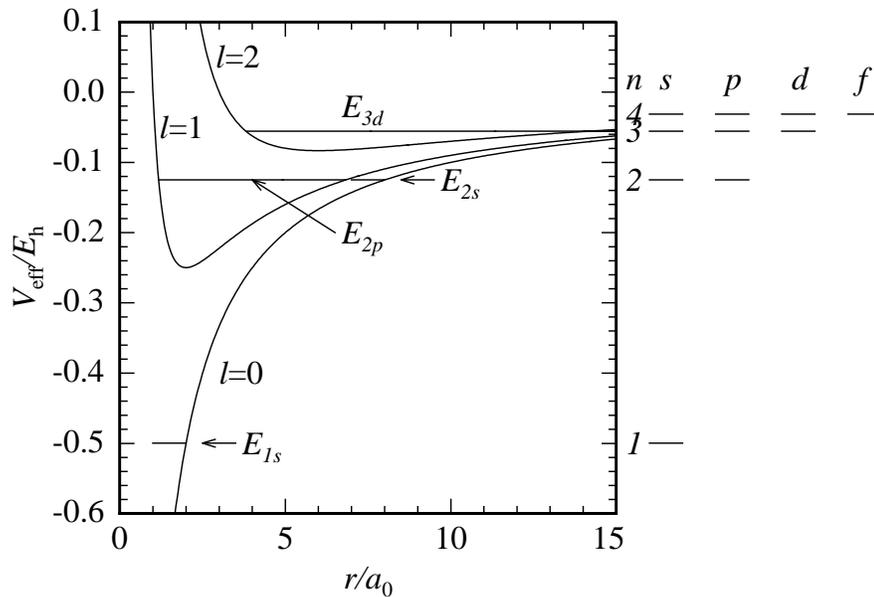
Figure 4.4: Effective radial potentials given by Eq. (4.112) for $l = 0$, 1, and 2. The first few energy levels in each potential are shown, and are plotted more clearly on the right. The scaling factors E_h and a_0 used on the two axes are defined in Section 4.3.2.

called the Bohr radius; it is 2/3 the average distance from the nucleus to an electron in the ground state of a hydrogen atom, and has the value 0.0529 nm (that is, just over half an Å).

(In fact there are many different sets of "orthogonal polynomials", mostly named after mathematicians: Hermite, Legendre, Laguerre, Chebychev, Jacobi, ... Each one has its own "weight function"; Gaussian for Hermite, exponential for Laguerre, constant for Legendre. You have now met the first three sets. The polynomials in $\cos\theta$ that make up the θ part of the Y_{lm} are Legendre polynomials.)

The quantum number n can have only positive integer values. The number l, as we saw in the section on angular momentum, can have only nonnegative integer values; in the H-atom case, acceptable wavefunctions are only possible if $n > l$, so for a given n, l can have the values $0, 1, \ldots n - 1$. In ordinary chemical notation we identify a particular radial wavefunction with the number n and the

Table 4.2 Radial functions $R_{nl}(r)$ for the first few energy levels of hydrogenlike atoms with atomic number Z. The length a is defined in Eq. (4.114).

n	l	$R_{nl}(r)$
1	0	$\left(\frac{Z}{a}\right)^{3/2} 2 e^{-Zr/a}$
2	0	$\left(\frac{Z}{a}\right)^{3/2} \frac{1}{\sqrt{2}} \left(1 - \frac{1}{2}\frac{Zr}{a}\right) e^{-Zr/2a}$
2	1	$\left(\frac{Z}{a}\right)^{3/2} \frac{1}{2\sqrt{6}} \frac{Zr}{a} e^{-Zr/2a}$
3	0	$\left(\frac{Z}{a}\right)^{3/2} \frac{2}{3\sqrt{3}} \left[1 - \frac{2}{3}\frac{Zr}{a} + \frac{2}{27}\left(\frac{Zr}{a}\right)^2\right] e^{-Zr/3a}$
3	1	$\left(\frac{Z}{a}\right)^{3/2} \frac{8}{27\sqrt{6}} \left(\frac{Zr}{a}\right) \left(1 - \frac{1}{6}\frac{Zr}{a}\right) e^{-Zr/3a}$
3	2	$\left(\frac{Z}{a}\right)^{3/2} \frac{4}{81\sqrt{30}} \left(\frac{Zr}{a}\right)^2 e^{-Zr/3a}$

letter corresponding to l; for instance, $1s$ is $n = 1$, $l = 0$; $2p$ is $n = 2$, $l = 1$, and so on.

Figure 4.5 gives plots of the radial functions $R_{nl}(r)$ and of the radial probability densities $r^2 R_{nl}^2(r)$ for $1 \le n \le 3$ and all the permissible values of l. The area under the probability density curve between $r = a$ and $r = b$ gives the probability of finding an electron at a distance $a \le r \le b$ of the nucleus. The r^2 factor in the probability density is proportional to the area of a spherical shell of radius r, since the probability of finding the electron between $r = a$ and $r = b$ really involves an integral over the entire volume between radii a and b. This is the same r^2 factor that appears in the "spherical polar volume element" $dV = r^2 \sin\theta\, dr\, d\theta\, d\phi$.

The total number of nodes in an H-atom wavefunction is $n - 1$. The number of angular nodes in Y_{lm} is l, so the number of radial nodes in $R_{nl}(r)$ is $n - l - 1$. So the $1s$ radial function has no nodes, the $2s$ has one node, the $2p$ has no nodes, and so on. These nodes are apparent in the upper half of Figure 4.5.

Only the radial functions with $l = 0$, that is, the functions associated with s orbitals, are nonzero at $r = 0$. When $l > 0$, the first term in the effective potential given by Eq. (4.112) becomes infinite at $r = 0$, and this "centrifugal potential" keeps the electron away from the nucleus. This same centrifugal potential is what makes it difficult to walk to the center of a spinning merry-go-round.

4.3.1 Energies

If we set the zero of energy as corresponding to an electron and a proton infinitely far apart and having no kinetic energy, then the energies of the individual H atom levels are

$$E_n = -\frac{R_{\mathrm{H}}}{n^2} = -\frac{1}{n^2}\frac{\mu Z^2 e^4}{32\pi^2 \epsilon_0^2 \hbar^2}, \tag{4.115}$$
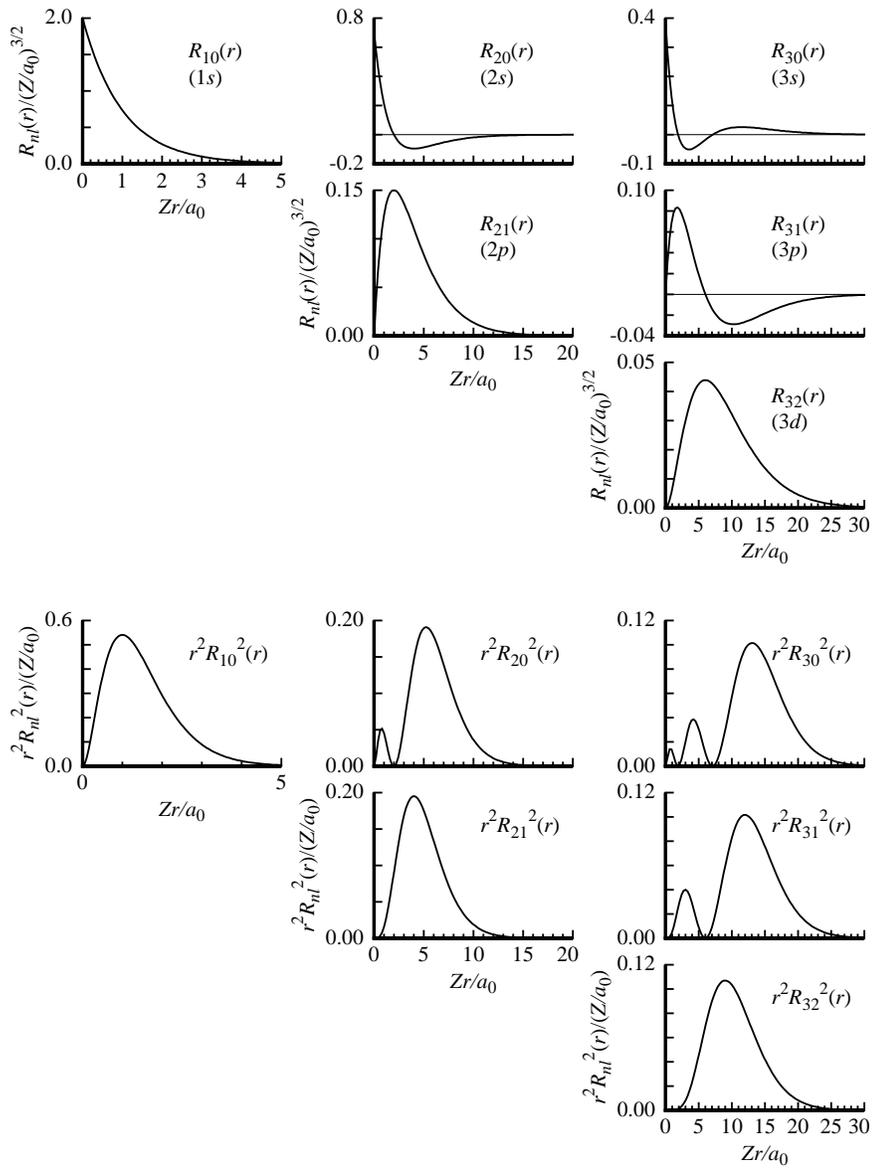
Figure 4.5: Radial functions $R_{nl}(r)$ and radial probability densities $r^2 R_{nl}^2(r)$ for hydrogenlike atoms with atomic number Z.

where R_H, the *Rydberg constant*, is defined by the collection of constants on the right side of the formula. It has units of energy, and is about $R_\text{H} = 2.18 \times 10^{-18}$ J = 13.6 eV. The energy levels go up as $\frac{1}{n^2}$; at $n = \infty$ we have $E_n = 0$; the lowest level has some energy (-13.6 eV); the next level up is three-fourths of the way to the ionization limit, and so on. There are an infinite number of electronic states whose energies approach zero asymptotically from below.

The minimum energy needed to remove the electron from a hydrogenlike atom in a state with principal quantum number n is $E_\infty - E_n = -E_n$; this is called the *ionization energy*. For the ground state of H the ionization energy is therefore 13.6 eV (a number that most chemists know).

4.3.2   Introduction to atomic units

There is a unit of energy called a "hartree" (symbol E_h), defined by

$$E_h = \frac{m_e e^4}{16\pi^2 \epsilon_0^2 \hbar^2} = 4.35974394 \times 10^{-18} \text{ J} = 27.21138386 \text{ eV}. \tag{4.116}$$

The energy of the ground state of hydrogen is $E_0 = -\frac{1}{2}\left(\frac{\mu}{m_e}\right) E_h$. (This is almost, but not exactly, minus half a hartree because the reduced mass μ of the proton-electron system is not exactly the mass m_e of the electron.) The hartree is the energy unit of a system of units called "atomic units", in which the equations and formulas of atomic and molecular problems take a particularly simple form. You have already seen one other member of the atomic unit family, the Bohr radius a used in the H atom wavefunctions. That quantity was defined using the *reduced* mass of the electron-nucleus system, while the atomic unit a_0 is defined in terms of the electron mass itself:

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} = 0.52917720859 \times 10^{-10} \text{ m}, \tag{4.117}$$

which is the atomic unit of length, given the name 'bohr'. The atomic unit of mass is that of the electron; the atomic unit of charge is the charge of the proton (which is minus that of the electron.) These choices of units for energy, length, and mass then define a unit of angular momentum, which turns out to be exactly $\hbar$. When the Schrödinger equations for atoms and molecules are written in terms of atomic units, most of the constant factors ($\hbar$, $4\pi\epsilon_0$, and so on) cancel and the equations look much simpler; we will use this trick when we study multielectron problems. For now, just remember that the H atom ground state energy is about minus half a hartree.

4.3.3 Selection rules

Transitions between H atom stationary states caused by light must have nonzero "transition matrix elements" as usual. In hydrogenlike atoms, the electric dipole moment is (by definition!) just equal to $-e\mathbf{r}$, where e is the unit charge and $\mathbf{r}$ is the vector pointing from the nucleus to the electron. The interaction Hamiltonian $\mathbf{E} \cdot \mu$ is therefore just proportional to $\mathbf{r}$, and the transition matrix elements look like (this is the z component)

$$\langle \phi_f | \hat{H}_1 | \phi_i \rangle = \int d\tau \phi_{n'l'm'}(r,\theta,\varphi) r \cos\theta \phi_{nlm}(r,\theta,\varphi) \tag{4.118}$$

$$= \int_0^{2\pi} d\varphi \int_0^\pi \sin\theta \, d\theta \int_0^\infty r^2 \, dr \, R_{n'l'}(r) Y_{l'm'}(\theta,\varphi) r \cos\theta R_{nl}(r) Y_{lm}(\theta,\varphi) \tag{4.119}$$

$$= \left[\int_0^{2\pi} d\varphi \int_0^\pi \sin\theta \, d\theta \, Y_{l'm'}(\theta,\varphi) \cos\theta(r) Y_{lm}(\theta,\varphi) \right]$$
$$\times \left[\int_0^\infty r^2 \, dr \, R_{n'l'}(r) r R_{nl}(r) \right] \tag{4.120}$$

The angular integral we have seen before: it is the same one we encountered in working out the selection rules for the rigid rotor, and it will give us once again the rules $\Delta l = \pm 1$, $\Delta m = 0, \pm 1$. Therefore, for electric dipole transitions within H atoms, the value of l must increase or decrease by one: a single photon can move an H atom from an s to a p state, but not to a d state or to another s state. Similarly, transitions out of p states must go to either s or to d states.

The radial integrals are not easy, but can be done; the classic reference is Condon and Shortley, *Theory of Atomic Spectra* (Cambridge University Press, 1959), section 5.6. They turn out to have no particular tendency to be zero; they are nonzero for most values of n and n'. Therefore, there is no selection rule on n; you can go, in a single jump, from any n to any other n (so long as your photon carries the appropriate energy, of course). So a perfectly legal transition in the H atom is $2s \rightarrow 77p$, but you cannot have $2s \rightarrow 3d$. Laser-excited transitions to high n levels in hydrogen atoms are the basis for a very sophisticated experimental method for studying chemical reactions called *Rydberg tagging*.

Chapter 5

Multiparticle problems and approximation methods

5.1 Helium

In the He atom we add another electron. The new electron has a kinetic energy, a potential energy of interaction with the nucleus, and—this is the new thing—a potential energy of interaction with the first electron. The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Ze^2}{4\pi\epsilon_0|\mathbf{r}_1|} - \frac{Ze^2}{4\pi\epsilon_0|\mathbf{r}_2|} + \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \tag{5.1}$$

where ∇_1^2 is the Laplacian (the sum of second derivatives with respect to the x, y, and z coordinates) of electron 1, and similarly for electron 2. $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the two electrons. $|\mathbf{r}_1 - \mathbf{r}_2|$, often written simply r_{12}, is the distance between the two electrons.

There are no eigenfunctions or eigenvalues of this Hamiltonian that can be written down in terms of elementary functions. (This is not really the fault of quantum mechanics: classical mechanics has no closed form solutions for the three-body problem either.) We must therefore look for approximate eigenfunctions and eigenvalues.

However, it's handy first to rewrite our Schrödinger equation in atomic units. Change variables from r to r/a_0, then divide both sides of the equation by E_h. The eigenvalue equation for $\hat{H}$ becomes (it's a little messy because of the derivatives

on r in the kinetic energy terms; I'll spare you the details)

$$\left[-\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}} \right] \psi(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2)$$
$$= E\psi(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2) \quad (5.2)$$

where $r_1 = |\mathbf{r}_1|$ and similarly for r_2. Now, an energy E that we calculate will come out in hartrees. (One disadvantage of atomic units is that the equations themselves become dimensionless so it's not easy to use units to check for mistakes.)

5.1.1 Perturbation theory approach

Let us regard this problem as one of perturbation theory; that is, as consisting of two parts, one of which we have already solved. The part of the Hamiltonian for which we already know the eigenfunctions and eigenvalues is the Hamiltonian of Eq. (5.2) but without the $\frac{1}{r_{12}}$ piece, which can be written

$$\hat{H}_0 = \left(-\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1}\right) + \left(-\frac{1}{2}\nabla_2^2 - \frac{Z}{r_2}\right). \quad (5.3)$$

This Hamiltonian is the sum of two hydrogen atom Hamiltonians, one in the coordinates of electron 1 and the other in the coordinates of electron 2. A sum of two independent Hamiltonians will have as its eigenfunctions the products of the eigenfunctions of the individual operators. In other words, the eigenfunctions of $\hat{H}_0$ are

$$\psi^{(0)}(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2) = \phi_{nlm}(r_1,\theta_1,\phi_1)\phi_{n'l'm'}(r_2,\theta_2,\phi_2) \quad (5.4)$$
$$= \phi_{nlm}(1)\phi_{n'l'm'}(2) \quad (5.5)$$

where the ϕ functions are the ordinary eigenfunctions of the H atom Hamiltonian, and the second line shows a common shorthand. The corresponding eigenvalues are sums of the H atom eigenvalues. Since for ordinary helium $Z = 2$, and the H atom ground state energy is proportional to Z^2, if we choose the ϕ_{1s} eigenfunction for both electrons, the overall energy of the helium (still neglecting interelectron repulsion) will be $E = \frac{-Z^2}{2} + \frac{-Z^2}{2}$, or -4 hartrees ($-4E_h$). The exact (experimentally measured) energy is $-2.9033E_h$, so neglecting the interelectron repulsion has caused an error in the total energy of more than 30%.

What can we do to account for the part of the Hamiltonian that we left out, $\hat{H}_1 = \frac{1}{r_{12}}$? Recall the simplest result of time dependent perturbation theory that we used before: the probability of making a transition from state $|\phi_i\rangle$ to state $|\phi_f\rangle$ under the influence of a perturbing Hamiltonian $\hat{H}_1$ was proportional to

$\langle\phi_f | \hat{H}_1 | \phi_i \rangle$. The simplest result of time independent perturbation theory is that if an eigenfunction $|\phi_n\rangle$ of $\hat{H}_0$ has nondegenerate eigenvalue E_n, then the *change* in E_n brought on by the presence of a perturbing Hamiltonian $\hat{H}_1$ is approximately $\langle\phi_n | \hat{H}_1 | \phi_n\rangle$. In other words, the first approximation to the change in the energy eigenvalue for a particular eigenfunction is just the expectation value of the perturbing Hamiltonian, computed using the original eigenfunctions from the unperturbed problem.

To get a better estimate of the true ground state energy of helium, then, we must evaluate the expectation value of the perturbing Hamiltonian in the approximate ground state wavefunction:

$$E^{(1)} = \left\langle \phi_{1s}(1)\phi_{1s}(2) \middle| \hat{H}_1 \middle| \phi_{1s}(1)\phi_{1s}(2) \right\rangle \tag{5.6}$$

This integral is a little complicated to do, because r_{12} is messy when written in spherical coordinates. But it can be done,[1] and the result is

$$E^{(1)} = \frac{5}{8}Z \tag{5.7}$$

so the estimated total energy of the ground state of He is (because $Z = 2$) $-4 + \frac{5}{4} = -2.75 E_h$. Now in comparison to the exact energy given above, we have a relative error of about 5%, so we have not done too badly!

I now want to introduce another important method for finding approximate eigenfunctions and eigenvalues of the Hamiltonian.

5.1.2   The variational theorem

We have some Hamiltonian $\hat{H}$ for our problem, and would like to find some   Levine §18.15
approximate eigenfunction and energy.

The variational theorem says this: You can choose *any* "trial" function you like, so long as it depends on the same variables as the true one, satisfies the same boundary conditions, and is normalizable. Then if you calculate the expectation value of the energy,

$$\langle E \rangle = \frac{\langle\psi_0 | \hat{H} | \psi_0\rangle}{\langle\psi_0 | \psi_0\rangle} = \frac{\int \psi_0^* \hat{H} \psi_0 \, d\mathbf{r}}{\int \psi_0^* \psi_0 \, d\mathbf{r}}, \tag{5.8}$$

---

[1] H. Eyring, J. Walter, and G.E. Kimball, *Quantum Chemistry*, Wiley, 1944, p. 103 gives an elegant approach. Both that approach and a more pedestrian one are outlined in McQuarrie and Simon, *Physical Chemistry: A Molecular Approach*, University Science Books, 1997, probs. 7–30 and 7–31.

where ψ_0 is your trial function, that value will always be larger than the true ground state energy. The denominator in this "Rayleigh quotient" is there for normalization; if your trial function ψ_0 is already normalized, it is not needed.

The proof of the theorem is as follows. We know that the trial function ψ_0 can be written as a linear combination of eigenfunctions of $\hat{H}$, which I will call the $|\varphi_n\rangle$. (Notice that we do not actually know what the $|\varphi_n\rangle$ are, or we wouldn't need to be using an approximate method!)

$$\psi_0 = \sum_{n=0}^{\infty} c_n |\varphi_n\rangle \tag{5.9}$$

I'll calculate the denominator first (try a 3-term example if this is unclear):

$$\langle \psi_0 | \psi_0 \rangle = \left[\sum_{m=0}^{\infty} c_m^* \langle \varphi_m | \right] \left[\sum_{n=0}^{\infty} c_n |\varphi_n\rangle \right] \tag{5.10}$$

$$= \sum_m \sum_n \langle \varphi_m | c_m^* c_n |\varphi_n\rangle \tag{5.11}$$

Terms in that double sum for which $m \neq n$ will be zero because of orthogonality, so the double sum reduces to a single sum:

$$= \sum_n \langle \varphi_n | c_n^* c_n |\varphi_n\rangle \tag{5.12}$$

$$= \sum_n |c_n|^2 \tag{5.13}$$

Now, calculate the numerator:

$$\langle \psi_0 | \hat{H} | \psi_0 \rangle = \left[\sum_m c_m^* \langle \varphi_m^* | \right] \hat{H} \left[\sum_n c_n |\varphi_n\rangle \right] \tag{5.14}$$

$$= \sum_m \sum_n c_m^* c_n \langle \varphi_m | \hat{H} | \varphi_n \rangle \tag{5.15}$$

$$= \sum_m \sum_n c_m^* c_n E_n \langle \varphi_m | \varphi_n \rangle \tag{5.16}$$

$$= \sum_n E_n |c_n|^2 \tag{5.17}$$

I'm interested in comparing to the true ground state energy E_0, so introduce it by adding and subtracting:

$$= \sum_n |c_n|^2 (E_0 + (E_n - E_0)) \tag{5.18}$$

$$= \sum_n |c_n|^2 E_0 + \sum_n |c_n|^2 (E_n - E_0) \tag{5.19}$$

$$= E_0 \sum_n |c_n|^2 + \sum_n |c_n|^2 (E_n - E_0) \tag{5.20}$$

Now examine the estimated energy, the quotient:

$$\frac{\langle \psi_0 | H | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \;=\; \frac{E_0 \sum_n |c_n|^2 + \sum_n |c_n|^2 (E_n - E_0)}{\sum_n |c_n|^2} \tag{5.21}$$

$$=\; E_0 + \frac{\overbrace{\sum_n |c_n|^2}^{>0} \overbrace{(E_n - E_0)}^{\ge 0}}{\underbrace{\sum_n |c_n|^2}_{>0}} \tag{5.22}$$

The second term must be zero or positive, so the expectation value of the energy must be greater than or equal to E_0. (It will be 0 only if $\psi_0 = |\varphi_0\rangle$, that is, if you guessed the correct ground state wavefunction.) The variational theorem is therefore proved.

This "variational principle" suggests a strategy for finding approximate wavefunctions: from a large set of normalized trial functions, select the one with lowest $\langle \psi_0 | H | \psi_0 \rangle$. If you can write the trial wavefunction as a function with parameters, and *minimize* $\langle H \rangle$ with respect to the parameters, so much the better.

5.1.3   Variational calculation for He

The ground state of a one-electron atom with nuclear charge Ze has the wavefunction

$$\psi_0 = R_{10}(r) Y_{00}(\theta, \phi) = C e^{-Zr/a_0} \tag{5.23}$$

where the normalization constant $C = \left(\frac{1}{\pi}\right)^{1/2} \left(\frac{Z}{a_0}\right)^{3/2}$. In atomic units we would drop the a_0 and write

$$\psi_0 = \left(\frac{Z^3}{\pi}\right)^{1/2} e^{-Zr}. \tag{5.24}$$

Let us try a product of these wavefunctions (one for each electron) as a trial function for He. However, rather than fixing the nuclear charge Z, let's use the nuclear charge as an adjustable parameter, called Z'. It will represent some kind of effective nuclear charge seen by the electrons; as Z' increases, the trial wavefunction will become more tightly localized around the nucleus, and as Z' decreases, it will "relax" so that
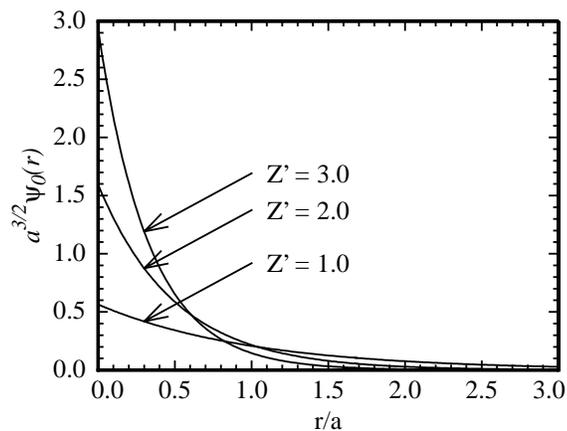
Figure 5.1: He trial wavefunction for the radial coordinate of one electron for three values of Z'. The true Z is 2.

the electrons can travel farther away from the nucleus.

Our trial He function will be

$$\Psi_0 = C^2 e^{-Z' r_1} e^{-Z' r_2} \tag{5.25}$$

where C, the normalization constant, must contain Z' rather than the true nuclear charge so the trial wavefunction is properly normalized for any choice of Z'. The dependence on r for each of the electrons for three values of Z' is shown in Figure 5.1.

To do a variational calculation, we find an expression for the energy $E_{\text{He}} = \left\langle \Psi_0 \middle| \hat{H}_{\text{He}} \middle| \Psi_0 \right\rangle$, and minimize it with respect to Z'. Ψ_0 is the approximate wavefunction that uses Z'. $\hat{H}_{\text{He}}$ is the true helium Hamiltonian, with $Z = 2$, and with the $\frac{1}{r_{12}}$ term included. In atomic units it is

$$\hat{H}_{\text{He}} = -\frac{\nabla_1^2}{2} - \frac{Z}{r_1} - \frac{\nabla_2^2}{2} - \frac{Z}{r_2} + \frac{1}{r_{12}}, \tag{5.26}$$

where the $\frac{Z}{r}$ terms are the Coulomb potential between each electron and the nucleus with true charge $Z = 2$, the ∇^2 terms are the kinetic energies of the two electrons, and the $\frac{1}{r_{12}}$ term is the Coulomb repulsion between the electrons.

It's possible to evaluate $\left\langle \Psi_0 \middle| \hat{H}_{\text{He}} \middle| \Psi_0 \right\rangle$ without having to apply the kinetic energy operators directly. Ψ_0 is a product of two functions, one for electron 1 and one for electron 2, that are eigenfunctions of a hydrogenlike Hamiltonian with effective nuclear charge Z'. If we add and subtract terms $\frac{Z'}{r}$ for each electron, we can write $\hat{H}_{\text{He}}$ as

$$\hat{H}_{\text{He}} = \left[-\frac{\nabla_1^2}{2} - \frac{Z'}{r_1} \right] + \left[-\frac{\nabla_2^2}{2} - \frac{Z'}{r_2} \right] + \frac{Z' - Z}{r_1} + \frac{Z' - Z}{r_2} + \frac{1}{r_{12}}, \tag{5.27}$$

where the operators in square brackets are now hydrogenlike Hamiltonians with effective nuclear charge Z'. Those operators applied to their respective eigenfunctions give eigenvalues $\frac{-(Z')^2}{2}$ times the eigenfunctions back. The remaining electron-nuclear Coulomb terms, $\left\langle \Psi_0 \middle| \frac{Z' - Z}{r_1} \middle| \Psi_0 \right\rangle$ for example, can be evaluated by ordinary integration, and each one gives the result $Z'(Z' - Z)$. The electron-electron repulsion term $\left\langle \Psi_0 \middle| \frac{1}{r_{12}} \middle| \Psi_0 \right\rangle$ was evaluated in Eq. (5.7). Assembling these results we find

$$E_{\text{He}} = 2 \left(\frac{-(Z')^2}{2} \right) + 2 Z'(Z' - Z) + \frac{5}{8} Z' \tag{5.28}$$

$$= (Z')^2 - 2 Z Z' + \frac{5}{8} Z'. \tag{5.29}$$

Then, minimizing with respect to Z', we find

$$Z' = Z - \frac{5}{16}. \tag{5.30}$$

So in a variational calculation using the product of two hydrogenlike wavefunctions with adjustable nuclear charge as the trial wavefunction, the optimum value of Z' for the ground state of He is $Z - \frac{5}{16}$. For He the nuclear charge Z is 2 so the optimized effective nuclear charge Z' is $1\frac{11}{16} = 1.6875$.

This lower "effective nuclear charge" is due to *screening*; the electric field seen by one electron from the nucleus is effectively reduced by the presence of the other electron.

What is the calculated energy?

$$E_{\mathrm{He}} = (Z')^2 - 2ZZ' + \frac{5}{8}Z' = -2.8477E_h. \tag{5.31}$$

The exact energy is $-2.9033E_h$, so the fractional error is

$$\frac{-2.9033 - (-2.8477)}{-2.9033} = 1.9\% \tag{5.32}$$

and the calculated energy is higher than the exact value, as expected.

5.1.4   More flexible trial functions

Now we have shown that if we take as a trial wavefunction for helium the product of two independent exponential decays in r, one for each electron, and we use the variational theorem to find the best value of the decay constant, the resulting decay constant corresponds to an effective nuclear charge a bit less than the true +2 charge of a helium nucleus, and the corresponding energy is about 2% above the true one.

How can we improve on this result? We should enlarge our collection of trial functions, making it more flexible. For example, we could choose

$$\Psi_0 = N[c_1 e^{-\alpha r_1} + c_2 e^{-\beta r_1}][c_1 e^{-\alpha r_2} + c_2 e^{-\beta r_2}] \tag{5.33}$$

and use c_1, c_2, α, and β as variational parameters. Here we will have to solve a set of 4 simultaneous equations (plus normalization) which we obtain by setting $\frac{\partial \langle \psi_0 | \hat{H} | \psi_0 \rangle}{\partial c_1} = 0$, etc.

5.1.5 Linear variational calculations

Since the trial wavefunction just suggested is nonlinear in its parameters, some of the resulting equations will be nonlinear. With many parameters, it is difficult to solve the equations.

Therefore, the standard way to do large variational calculations is to make the trial wavefunction linear in the parameters:

$$\psi_0 = a_1 f_1(\mathbf{r}_1, \mathbf{r}_2, \cdots) + a_2 f_2(\mathbf{r}_1, \mathbf{r}_2, \cdots) + a_3 f_3(\mathbf{r}_1, \mathbf{r}_2, \cdots) + \ldots \tag{5.34}$$

where the a_n are the parameters we adjust to get the best (lowest) energy. In general, the functions f_n are functions of the coordinates of all the electrons; in many cases, each f_n will be a product of functions each depending on coordinates of one electron. The different f_n might have different effective charges, be centered on different nuclei, and so on. Then the optimization problem becomes a set of linear algebraic equations. Very well-established methods are available for solving such sets of equations on computers.

When you evaluate $\langle \psi_0 | H | \psi_0 \rangle$, differentiate with respect to the a_n, and set the derivatives to 0, you obtain a homogeneous set of algebraic equations that look like

$$a_1(H_{11} - ES_{11}) + a_2(H_{12} - ES_{12}) = 0$$
$$a_1(H_{21} - ES_{21}) + a_2(H_{22} - ES_{22}) = 0 \tag{5.35}$$

(with more terms if there are more parameters.) In those equations

$$H_{11} = \left\langle f_1 \left| \hat{H} \right| f_1 \right\rangle \tag{5.36}$$
$$H_{12} = \left\langle f_1 \left| \hat{H} \right| f_2 \right\rangle \tag{5.37}$$
$$S_{11} = \left\langle f_1 | f_1 \right\rangle \tag{5.38}$$
$$S_{12} = \left\langle f_1 | f_2 \right\rangle, \tag{5.39}$$

which can all be computed once the basis set $\{f_1, f_2, \ldots\}$ has been chosen. E is the expectation value of the energy $\left\langle \psi_0 \left| \hat{H} \right| \psi_0 \right\rangle$, which won't be known until the a_n coefficients have been determined.

There is a theorem from linear algebra that says that a nontrivial solution to Eq. (5.35) (one other than $a_1 = a_2 = 0$) exists only if the determinant of the coefficients vanishes:

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - ES_{22} \end{vmatrix} = 0 \tag{5.40}$$

This is the notorious secular determinant of elementary molecular orbital theory, and the lowest value of E which solves the equation is the estimated ground state energy.

Modern quantum chemistry calculations (done with Gaussian and similar programs) use this linear variation technique. In those programs, many basis functions f_n are used; the user of the program selects a particular "basis set", and that specifies a particular collection of basis functions on each atom in the molecule. The program then sets up Eq. (5.35) as a large linear algebra problem; there can easily be hundreds or thousands of equations in the set, but well-developed computational methods exist for finding solutions. In fact often as much work goes into calculating the matrix elements (integrals) H_{12} and so on as into solving the set of equations once the integrals are known.

So far, we have investigated only hydrogen and helium atoms; we can study the ground states of those atoms without worrying about the spin of the electrons. But to make any further progress, we must look at the quantum mechanics of systems containing multiple identical particles. That is the next topic.

5.2 Multiple particles

The wavefunction of a system of many particles is a function of (at most) all the coordinates of all the particles. For example, for 2 particles (labeled 1 and 2) in 3 dimensions we would have

Levine §19.8

$$\psi = \psi\left(x_1, y_1, z_1, x_2, y_2, z_2\right) = \psi\left(\mathbf{r}_1, \mathbf{r}_2\right), \tag{5.41}$$

and similarly for larger number of particles.

The interpretation of ψ is a natural extension of the one-particle case: $|\psi\left(\mathbf{r}_1, \mathbf{r}_2\right)|^2 d\mathbf{r}_1 d\mathbf{r}_2$ gives the probability of finding particle 1 in a box of volume $d\mathbf{r}_1$ at $\mathbf{r}_1$, and simultaneously finding particle 2 in a box of volume $d\mathbf{r}_2$ at position $\mathbf{r}_2$.

Note the units: $|\psi|^2 d\mathbf{r}_1 d\mathbf{r}_2$ is a probability, and must be unitless. $d\mathbf{r}_1$ has units m^3 (volume), as has $d\mathbf{r}_2$. So $\psi\left(\mathbf{r}_1, \mathbf{r}_2\right)$ must have units m^{-3}. In general, for n particles, ψ will have units m$^{-3n/2}$.

If a system includes two particles that are *identical* (also called *indistinguishable*), then there is a special requirement on the wavefunction. If the two particles are the same, then it cannot matter which of them I call "1" and which I call "2": the physical properties must be the same no matter which choice I make.

That means that if I switch the labels on the particles, the most that can happen to the wavefunction is that it gets multiplied by some complex number,

call it γ, with magnitude 1.

$$\psi(\mathbf{r}_2, \mathbf{r}_1) = \gamma \psi(\mathbf{r}_1, \mathbf{r}_2). \tag{5.42}$$

Now, if we do the same trick again, we multiply by γ again:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \gamma \psi(\mathbf{r}_2, \mathbf{r}_1) = \gamma^2 \psi(\mathbf{r}_1, \mathbf{r}_2) \tag{5.43}$$

So $\gamma^2 = 1$. That means $\gamma = 1$ or -1.

This is a general rule of QM: whenever the labels of two identical particles in a multiparticle wavefunction are *exchanged*, the wavefunction must either *remain the same* or *change sign*. Which of those it does depends on the type of particles:

- Particles for which $\gamma = 1$ are called *bosons*. Photons and deuterium nuclei are examples.

- Particles for which $\gamma = -1$ are called *fermions*. Both protons and electrons are fermions.

For chemistry, the most important application is to electrons: any multi-electron wavefunction must change sign when the *labels* of any 2 electrons are exchanged. This is one statement of the *Pauli principle*.

5.2.1 Spin

In quantum mechanics, *spin* means an "intrinsic angular momentum" that is a property of most particles. The spin operators are usually written $\hat{s}$ (or $\hat{I}$ for nuclear spins). They have all the properties we identified in our study of general angular momenta.

A complete wavefunction for a system of electrons must include a "spin part" that describes the spin state for each electron. Electrons always have $s = \frac{1}{2}$, so the two spin eigenfunctions from which we construct the spin part are simply $\left| s = \frac{1}{2}, m_s = \frac{1}{2} \right\rangle$ and $\left| s = \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle$, also known as

$$\left| \frac{1}{2} \, \frac{1}{2} \right\rangle \quad \text{and} \quad \left| \frac{1}{2} \, -\frac{1}{2} \right\rangle, \tag{5.44}$$

$$\alpha \quad \text{and} \quad \beta, \tag{5.45}$$

$$\text{or "spin up" and "spin down".} \tag{5.46}$$

For the H atom, possible stationary state wavefunctions include

$$\psi = R_{nl}(r) Y_{lm}(\theta, \phi) \alpha \tag{5.47}$$

$$\psi = R_{nl}(r) Y_{lm}(\theta, \phi) \beta \tag{5.48}$$

$$\psi = R_{nl}(r) Y_{lm}(\theta, \phi) \frac{1}{\sqrt{2}} (\alpha + \beta) \tag{5.49}$$

Since the spin functions do not appear in the H-atom Hamiltonian anywhere, all three of these functions will have the same energy.

The α and β "functions" (the angular momentum kets $\lvert j\,m\rangle$ with $j = \frac{1}{2}$) are definitely *not* functions of the electron coordinates. It's best to think of them simply as abstract eigenfunctions of $\hat{s}^2$ and $\hat{s}_z$, with properties defined by those operators. If you don't like that, you can think of some fourth coordinate of the electron called its "spin variable" and regard α and β as functions of that.

5.2.2   Spin and the Pauli principle

Consider a generic two-fermion system, and ignore for the moment any interaction between the particles (for instance, Coulomb repulsion). Then the spin part of the two-particle wavefunction is just a product of a spin function for electron 1 and one for electron 2. In the simplest case, when each electron is in some eigenstate of its $\hat{s}_z$ operator, there are four possibilities:

$\alpha(1)\alpha(2)$ both electrons "up"

$\alpha(1)\beta(2)$ first up, second down

$\beta(1)\alpha(2)$ first down, second up

$\beta(1)\beta(2)$ both down

A complete two-electron wavefunction will be the product of one of these with a spatial part. For the overall wavefunction to obey the Pauli principle, each part must change into $\pm$ itself if the electron labels are interchanged. Let's look at these candidates, using the symbol $\hat{P}_{12}$ to indicate an operator that just changes a label 1 to a 2 and vice versa. The first candidate is

$$\hat{P}_{12}[\alpha(1)\alpha(2)] = \alpha(2)\alpha(1), \tag{5.50}$$

which is okay; multiplication is commutative, so this is just what we started with.

$$\hat{P}_{12}[\alpha(1)\beta(2)] = \alpha(2)\beta(1) = \beta(1)\alpha(2) \tag{5.51}$$

which is not plus or minus itself!

$$\hat{P}_{12}[\beta(1)\alpha(2)] = \beta(2)\alpha(1) = \alpha(1)\beta(2) \tag{5.52}$$

which, again, is no good. Finally,

$$\hat{P}_{12}[\beta(1)\beta(2)] = \beta(2)\beta(1) = \beta(1)\beta(2) \tag{5.53}$$

which is okay.

Notice that the middle two candidates turned into each other under $\hat{P}_{12}$. That suggests a trick: Try

$$\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) + \beta(1)\alpha(2)) \tag{5.54}$$

or

$$\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \beta(1)\alpha(2)). \tag{5.55}$$

Let's check the first one:

$$\hat{P}_{12}[\alpha(1)\beta(2) + \beta(1)\alpha(2)] = \alpha(2)\beta(1) + \beta(2)\alpha(1) \tag{5.56}$$

$$= \alpha(1)\beta(2) + \beta(1)\alpha(2) \tag{5.57}$$

which works (with $\gamma = +1$).

How about the second?

$$\hat{P}_{12}[\alpha(1)\beta(2) - \beta(1)\alpha(2)] = \alpha(2)\beta(1) - \beta(2)\alpha(1) \tag{5.58}$$

$$= (-1)\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \tag{5.59}$$

so that is also an acceptable spin function; it has $\gamma = -1$.

Now we have four candidate spin functions with acceptable symmetry properties:

$$\left.\begin{array}{c} \alpha(1)\alpha(2) \\ \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \beta(1)\alpha(2)\right) \\ \beta(1)\beta(2) \end{array}\right\}\gamma = 1$$

$$\frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \qquad \gamma = -1$$

If you form the total spin operator $\mathbf{S}_T = \mathbf{s}_1 + \mathbf{s}_2$, then the first group are all eigenfunctions of S_T^2 with eigenvalue $2\hbar^2$, and are called *triplet* functions; the last is an eigenfunction of S_T^2 with eigenvalue 0, and is called the *singlet* function.

5.2.3 Singlet and triplet helium

Consider a trial wavefunction that is just a product of two H-atom-like functions, $Z = 2$, with one electron in the $1s$ orbital and the other in the $2s$. (This will not

give particularly good numerical results but the qualitative behavior is correct.)
The basic spatial functions therefore look like

$$\varphi_{1s}(r_1)\varphi_{2s}(r_2) \equiv 1s(1)2s(2) \tag{5.60}$$

But because we have two identical particles we must construct functions that
are either symmetric or antisymmetric when the electron labels are interchanged.
We have two choices:

$$\psi_s \propto \varphi_{1s}(r_1)\varphi_{2s}(r_2) + \varphi_{2s}(r_1)\varphi_{1s}(r_2) \tag{5.61}$$

$$\psi_a \propto \varphi_{1s}(r_1)\varphi_{2s}(r_2) - \varphi_{2s}(r_1)\varphi_{1s}(r_2) \tag{5.62}$$

The first one is symmetric if the electron labels are exchanged; it must there-
fore be paired with an antisymmetric spin function. We have only one of those,
the singlet. So

$$\Psi_s = \frac{1}{\sqrt{2}} (1s(1)2s(2) + 2s(1)1s(2)) \frac{1}{\sqrt{2}} \left(\alpha(1)\beta(2) - \beta(1)\alpha(2) \right) \tag{5.63}$$

The second spatial function is antisymmetric under exchange, so it must be
paired with a symmetric spin function. We have 3 of those, the 3 components of
the triplet:

$$\Psi_t = \frac{1}{\sqrt{2}} (1s(1)2s(2) - 2s(1)1s(2)) (\alpha(1)\alpha(2)) \tag{5.64}$$

or

$$\Psi_t = \frac{1}{\sqrt{2}} (1s(1)2s(2) - 2s(1)1s(2)) \frac{1}{\sqrt{2}} \left(\alpha(1)\beta(2) + \beta(1)\alpha(2) \right) \tag{5.65}$$

or

$$\Psi_t = \frac{1}{\sqrt{2}} (1s(1)2s(2) - 2s(1)1s(2)) (\beta(1)\beta(2)) \tag{5.66}$$

You can think of these functions as "linear combinations of freshman chem-
istry electron diagrams". The diagrams corresponding to the different spin com-
binations are shown in Figure 5.2.
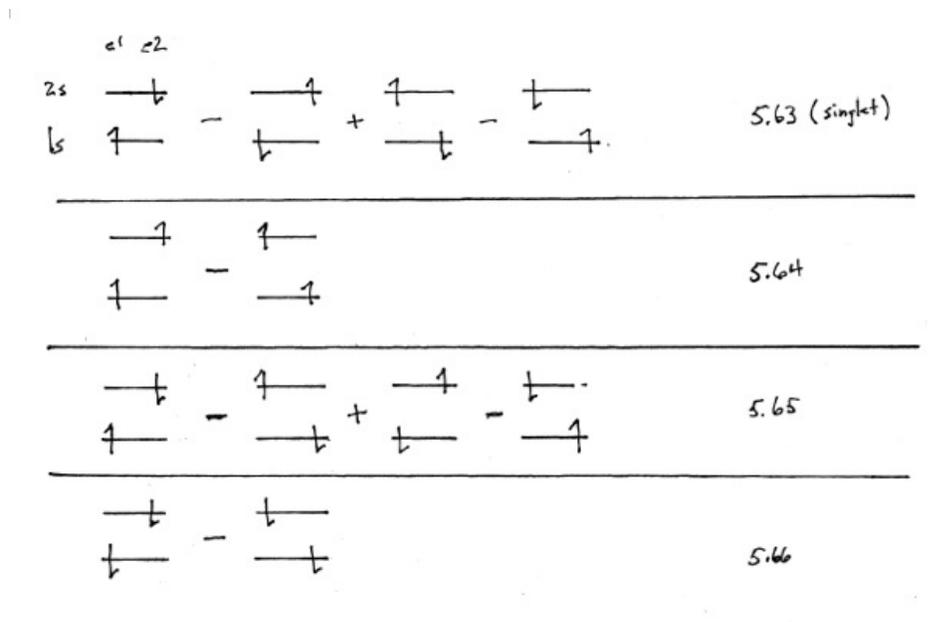
Figure 5.2: Basic orbital diagrams corresponding to equations (5.63) to (5.66). Electron 1 is shown on the left side of each line, and electron 2 on the right.

5.2.4   Hund's first rule

Let's look at the behavior of the two spatial wavefunctions ψ_s and ψ_a, shown in Eq. (5.61) and Eq. (5.62), when the electrons come close together. When $r_1 = r_2$, the antisymmetric spatial wavefunction becomes 0; therefore, two electrons with that spatial wavefunction cannot come together. This is the spatial function that must be paired with the triplet spin function. The symmetric space function ψ_s does not have that property; it is the one that goes with the singlet spin function. The difference in electron correlation provides one component of the difference in energy between the two states, and in fact the triplet is lower.

The lower triplet energy for the two-electron system is an example of "Hund's first rule": for two atomic states with the same electron configuration (here $1s2s$), the one with the larger value of S will be lower in energy. S is the quantum number giving the eigenvalues of the $\hat{S}^2$ operator (total spin). In the singlet $S = 0$; in the triplet $S = 1$.

The experimentally observed values of the He total energies are

$$1s^2 \quad (^1S) \quad -2.9033E_h \tag{5.67}$$

$$1s2s \quad (^3S) \quad -2.1750 \tag{5.68}$$

$$1s2s \quad (^1S) \quad -2.1457 \tag{5.69}$$

so for the two wavefunctions with $1s2s$ electron configurations, the singlet is $.029E_h = 0.8eV$ higher. (The separation between $1s^2$ and $1s2s$ is ~ 20 eV.)

5.3 Multielectron atoms and term symbols

A hydrogen atom (single electron!) stationary state wavefunction is an eigenfunction of at least five operators: $\hat{H}$, $\hat{L}^2$, $\hat{L}_z$, $\hat{s}^2$, and $\hat{s}_z$. For a single electron, there is no $\frac{1}{r_{12}}$ term in the Hamiltonian, because there are not two electrons to have any repulsion.

In larger atoms, the *orbital approximation* consists of assuming that the wavefunction of the atom can still be usefully written as a product of one-electron (often, hydrogenlike) orbitals. Such a wavefunction would be an eigenfunction of the Hamiltonian if it did not contain interelectron repulsion terms. Even with those terms, the orbital approximation is a very useful qualitative tool, though for accurate calculations of molecular properties it must be abandoned.

The interelectron repulsion terms that are proportional to $\frac{1}{r_{ij}}$, where i and j label all the possible pairs of electrons in a multielectron atom, have another important effect. Once they are included in the Hamiltonian, the $\hat{L}^2$ and $\hat{L}_z$ operators for individual electrons no longer commute with $\hat{H}$. That means that the orbital angular momentum for an individual electron is no longer conserved. This should come as no surprise; if two electrons collide, their paths are altered. If one of them had been (for instance) circling the nucleus in the xy plane , so that it had a large z component of angular momentum, then it might be knocked out of that plane by the collision and its z component of angular momentum reduced.

On the other hand, even with the $\frac{1}{r_{ij}}$ terms included, the operator for the total electronic orbital angular momentum, $\hat{\mathbf{L}}_T = \hat{\mathbf{L}}_1 + \hat{\mathbf{L}}_2 + \cdots$, still does commute with the Hamiltonian. That is, while the electrons are continuously trading their individual angular momenta back and forth, the total electronic angular momentum is conserved. It is therefore useful to label different states of atoms with quantum numbers L and M_L giving the eigenvalues with respect to $\hat{L}_T^2$ and $\hat{L}_{Tz}$.

The individual spin operators do still commute with $\hat{H}$ when the interelectron repulsion terms are included. Nevertheless, it is customary to consider a total spin

operator $\hat{\mathbf{S}}_T$ as well, and label atomic states with values of their total *spin* quantum numbers, S and M_S. (This becomes necessary when the next important term in the Hamiltonian, called "spin-orbit coupling", is included.) For a particular atom, we need to learn how to decide which values of L, M_L, S and M_S are possible, and to some extent what their relative energies will be.

5.3.1   Addition of Angular Momenta

In general, when you add two vectors, you add their corresponding components together to get the components of the new "resultant" vector. If the vectors we are interested in are quantum mechanical angular momenta, however, we have the problem that we don't (in fact, *can't*) know all their components at once! We therefore have to settle for just adding the z components together, and then figuring out what overall vector lengths are possible by looking at the list of possible z components in the resultant.

When two angular momenta ℓ_1 and ℓ_2 are added, the resultant quantum number L can have any value from $\ell_1 + \ell_2$ down to $|\ell_1 - \ell_2|$. The resultant quantum number M is always $M_1 + M_2$, and each resulting L must have its full complement of M values.

To see where these rules come from, let's look in detail at an example. If we add two angular momenta, $\ell_1 = 1$ and $\ell_2 = \frac{1}{2}$, the resulting M values for all the possible combinations are given in the following table.

ℓ_1	M_1	ℓ_2	M_2	M
1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
1	1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$
1	-1	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$
1	-1	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$

Seeing an M of $\frac{3}{2}$ means that there must be a resultant L of $\frac{3}{2}$; it requires M's of $\frac{3}{2}$, $\frac{1}{2}$, $-\frac{1}{2}$, and $-\frac{3}{2}$. Those four possibilities make up an $L = \frac{3}{2}$ state. The remaining two rows have $M = \frac{1}{2}$ and $-\frac{1}{2}$; they belong to an $L = \frac{1}{2}$ state. So adding two systems with $\ell_1 = 1$ and $\ell_2 = \frac{1}{2}$ gives possible resulting angular momenta of $\frac{3}{2}$ and $\frac{1}{2}$.

A multielectron atomic *term* (for light atoms, $Z \le 40$, anyway) will have well-defined values of L, M_L, S, M_S. The standard way to describe atomic states is with

a "term symbol", ^{2S+1}L. L is written with the letters $S, P, D, F, \ldots$ for $L = 0, 1, 2, 3 \ldots$, just as in the single-electron case (but for multielectron atoms we use capital letters.) So a ^4F state has $S = \frac{3}{2}$ and $L = 3$; that means its eigenvalue with respect to $\hat{S}_T^2$ is $\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\hbar^2$, and its eigenvalue with respect to $\hat{L}_T^2$ is $(3)(4)\hbar^2$.

5.3.2 Finding the lowest atomic term

Hund's second rule says that for two electronic terms arising from the same electron configuration with the same value of S, the lowest one will be the one with larger L. The combination of L and S corresponding to the lowest energy term for a given electron configuration is therefore easy to find. It is necessary to consider only the electrons in subshells that are not completely filled, where a subshell is defined by particular values of n and l. For example, for nitrogen, we need to consider only the $2p$ electrons. Draw an electron-orbital diagram like you learned to do in freshman chemistry, placing the electrons in the highest m_l orbitals first, and keeping as many parallel-spin electrons as possible. Then, add up all the values of m_l to get the value of L for the ground term, and all the values of m_s to get the value of S.

For example, for N, we need to put three electrons into the $2p$ orbitals. We must put one each into $m_l = 1, 0$, and -1 to keep all their spins parallel. Then the total value of S is $\frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2}$, and the total value of L is $1 + 0 + (-1) = 0$. So the ground term has $S = \frac{3}{2}$ and $L = 0$, and is ^4S.
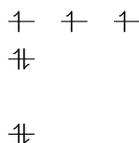
The term symbols for the ground states of all configurations with a single partially filled subshell are shown in Table 5.1. This table is easy to use; for example, the Fe^{2+} ion has an electron configuration of $[Ar]d^6$, and the table indicates that the ground state will have term symbol ^5D.

**Table 5.1** Term symbols for the ground states of atoms or atomic ions with a single partially filled subshell.

s^1						
^2S						

p^1, p^5	p^2, p^4	p^3				
^2P	^3P	^4S				

d^1, d^9	d^2, d^8	d^3, d^7	d^4, d^6	d^5		
^2D	^3F	^4F	^5D	^6S		

f^1, f^{13}	f^2, f^{12}	f^3, f^{11}	f^4, f^{10}	f^5, f^9	f^6, f^8	f^7
^2F	^3H	^4I	^5I	^6H	^7F	^8S

5.3.3 Finding all the possible atomic terms

To find the atomic terms arising from a particular configuration, we construct a *microstate table* as follows:

1. Draw diagrams for all the electron-orbital arrangements corresponding to that configuration.

2. Calculate values of M_L and M_S for each diagram.

3. Look for the highest value of M_L. If it appears more than once, choose the entry with the highest M_S.

4. There must be a term with $L = M_{L_{\max}}$. Go down your list of M_L and M_S, marking off lines with $M_L = M_{L_{\max}}$, $M_{L_{\max}-1}$, etc down to $-M_{L_{\max}}$, all with the same M_S. Then mark off another set with $M_S = M_{S_{\max}-1}$, and so on down to $-M_{L_{\max}}$, $-M_{S_{\max}}$. All the lines you just marked belong to a single atomic state.

5. Repeat from step 3 for remaining lines.

Let's do carbon as an example. The basic orbital diagram is
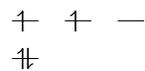
For that particular configuration we would find (counting the leftmost p orbital as $m_l = 1$) $M_L = 0+0+0+0+1+0 = 1$ and $M_S = \frac{1}{2} + -\frac{1}{2} + \frac{1}{2} + -\frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 1$. The ground state (or ground term) is therefore $L = 1, S = 1$ or ^3P.

The plan now is to do a similar analysis for *all* the ways of placing the electrons into the orbitals in a $1s^2 2s^2 2p^2$ configuration. The "filled subshells" will always add to 0 for both M_L and M_S, so don't draw them. Showing only $2p$ orbitals, the possible diagrams are

1	0	-1	M_L	M_S	
↑	↑	—	1	1	^3P
↑	—	↑	0	1	^3P
—	↑	↑	−1	1	^3P
↑↓	—	—	2	0	^1D
↑	↓	—	1	0	^3P
↑	—	↓	0	0	^3P
—	↑↓	—	0	0	^1S
—	↑	↓	−1	0	^3P
↓	↑	—	1	0	^1D
↓	—	↑	0	0	^1D
—	↓	↑	−1	0	^1D
—	—	↑↓	−2	0	^1D
↓	↑	—	1	−1	^3P
↓	—	↓	0	−1	^3P
—	↓	↑	−1	−1	^3P

} not the only possible labeling!

The largest value of M_L is 2; it appears with $M_S = 0$ only. So that state must be $S = 0, L = 2$; it's a ^1D. We mark lines with $M_L = 2, 1, 0, -1, -2$ as belonging to ^1D (all with $M_S = 0$). When we have choices of more than one line, it doesn't matter which we mark.

After marking those off, the largest value of M_L remaining is 1, with $M_S = 1$. So that must be a ^3P state. We must mark nine lines: $M_L = 1, 0, -1$ with $M_S = 1, M_L = 1, 0, -1$ with $M_S = 0$, and $M_L = 1, 0, -1$ with $M_S = -1$.

That leaves only one line: $M_L = 0, M_S = 0$. That line makes up a ^1S state.

So for the electron configuration $1s^2 2s^2 2p^2$, there are three atomic states: $^3P, ^1D, ^1S$.

The energy ordering of these states is given by Hund's rules:

1. Higher multiplicity (higher S) means lower energy.

2. For the same multiplicity, higher L means lower energy.

So ^3P is the ground state, followed by ^1D and ^1S.

5.3.4 Spin-orbit coupling

So long as we consider only the Coulomb terms in the Hamiltonian of an atom, the total electronic angular momentum operator $\hat{\mathbf{L}}$ and the spin angular momenta of the individual electrons $\hat{\mathbf{s}}_i$ all commute with the Hamiltonian; their expectation values are therefore constant. The total spin angular momentum operator $\hat{\mathbf{S}}$ also commutes with $\hat{H}$; the eigenvalues of its associated operators $\hat{S}^2$ and $\hat{S}_z$ are $S(S+1)\hbar^2$ and $M_S\hbar$. However, there is another term in the atomic Hamiltonian that has to do with the magnetic fields produced by the electrons' spins and their orbital motions. This magnetic energy is called "spin-orbit interaction", and to a useful approximation it is given by

$$H_{\mathrm{SO}} \approx A\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} \tag{5.70}$$

where A is a constant called the spin-orbit coupling constant. (A depends on the nuclear charge; in fact, it is roughly proportional to Z^4, so the spin-orbit coupling is much more important for heavy than for light atoms.) Once this term is included in the Hamiltonian, neither $\hat{\mathbf{L}}$ nor $\hat{\mathbf{S}}$ are constants of the motion, and only the total angular momentum of the atom, usually written $\hat{\mathbf{J}}$, is still conserved.

The spin-orbit coupling has two important physical effects:

1. The energies of the different components of a "multiplet" term (such as ^3P) are no longer all the same. Instead, the energies of the states with different values of total J will be different. This difference amounts to a few tens of cm^{-1} in light atoms, but several thousands of cm^{-1} in heavy atoms.

   A ^3P atomic state has $S=1$ and $L=1$. Therefore, J can take values 0, 1, and 2 by the normal rules of addition of angular momenta. Normally the term symbol for the atom is written as before, but with the value of total J added as a right subscript. So the three components of the ^3P atom become $^3P_0$, $^3P_1$, and $^3P_2$. The energy ordering between the states depends on the electron configuration and is given by *Hund's third rule*: for subshells less than half filled, lower J means lower energy, while for subshells more than half filled, higher J means lower energy. So in carbon, which has two $2p$ electrons, the $^3P_0$ state is the ground state, while for oxygen, which has four $2p$ electrons, the $^3P_2$ state is the lowest in energy. Atoms with exactly half-filled subshells have $L=0$ ground states so the spin-orbit coupling is very small.

   The ground state of iodine is ^2P; the two spin-orbit states are therefore $^2P_{\frac{1}{2}}$ and $^2P_{\frac{3}{2}}$. The $^2P_{\frac{3}{2}}$ is the lower energy (ground state) one, and the difference in energy is 7603 cm^{-1}.

2. The selection rules for electric dipole transitions in atoms in the absence of spin-orbit coupling are $\Delta L = 0, \pm 1$ and $\Delta S = 0$. However, with spin-orbit coupling included, neither L nor S is truly a "good quantum number"; that is, the operators they specify eigenvalues for no longer commute with $\hat{H}$. Those rules are therefore relaxed. In particular, $\Delta S \neq 0$ transitions become common and important in heavy atoms. On the other hand, in light atoms, they remain quite weak.

5.3.5 Useful references

The microstate-table procedure for finding all the possible atomic terms becomes tedious for "subshells" with higher l (d and f sets of orbitals), especially for subshells near half-full. There are some shortcuts for that method that reduce the number of microstates that must be listed, and some more powerful group-theory approaches. But essentially this is a problem that can be done once and the answers written down. Complete tables of possible term symbols for partially filled subshells up through f are given in Table 1[7] of E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, 1959), and in Table 1 of W. C. Martin, R. Zalubas, and L. Hagan, *Atomic Energy Levels — The Rare-Earth Elements* (Nat. Bur. Stand., U.S., 1978), available at `http://www.nist.gov/data/nsrds/NSRDS-NBS-60.pdf`.

The National Institute of Standards and Technology maintains an excellent database of atomic energy level properties, available at `http://www.nist.gov/pml/data/asd.cfm`. It has several sections, giving energy levels, spectral lines, and ground-state term symbols and ionization energies. (This database originated in a famous set of published tables entitled *Atomic Energy Levels* compiled by Charlotte Moore of NIST, and many people still call it "Moore's tables.") To use it you must use the annoying atomic-physics nomenclature for atomic ions, where neutral Fe is called Fe I, Fe^+ is called Fe II, and so on.

Chapter 6

Molecules

Consider the simplest molecular ion, H_2^+, with distance variables defined in Figure 6.1. The Hamiltonian is easy to write down:

$$\hat{H} = \frac{P_e^2}{2m_e} + \frac{P_A^2}{2M_A} + \frac{P_B^2}{2M_B} - \frac{Z_A e^2}{r_A} - \frac{Z_B e^2}{r_B} + \frac{Z_A Z_B e^2}{R_{AB}} \tag{6.1}$$

In atomic units,

$$\hat{H} = -\frac{1}{2}\nabla_e^2 - \frac{m_e}{M_A}\frac{1}{2}\nabla_A^2 - \frac{m_e}{M_B}\frac{1}{2}\nabla_B^2 - \frac{Z_A}{r_A} - \frac{Z_B}{r_B} + \frac{Z_A Z_B}{R_{AB}} \tag{6.2}$$

This is a three-body problem, like that for He.
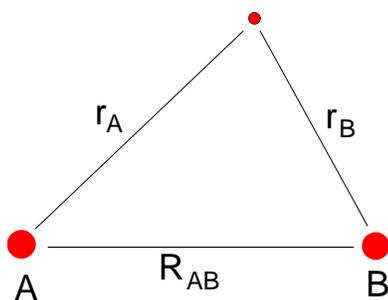
Figure 6.1: Variables for H_2^+.

6.0.1   Born-Oppenheimer Approximation

Nuclei are much heavier than electrons. Let's separate molecular behavior into two timescales: electron motion and nuclear motion.

Assume that the nuclei are fixed at some position, and solve the resulting one-body-in-noncentral-field problem. Then fix the nuclei in a different position, and repeat.

Clamping the nuclei has two effects on $\hat{H}$:

1. The $\frac{P_A^2}{2M_A}$ and $\frac{P_B^2}{2M_B}$ terms that describe the kinetic energies of the nuclei go away.

2. The term $\frac{Z_A Z_B e^2}{R_{AB}}$ that describes the nucleus-nucleus repulsion becomes constant. We can just make it part of the energy eigenvalue.

Now we need to solve the Schrödinger equation

$$\left[\frac{P_e^2}{2m_e} - \frac{Z_A e^2}{r_A} - \frac{Z_B e^2}{r_B} \right] \psi_{\mathrm{el}}(\mathbf{r}_1) = E_{\mathrm{el}} \psi_{\mathrm{el}}(\mathbf{r}_1) \tag{6.3}$$

That gets us an electronic wavefunction $\psi_{\mathrm{el}}(\mathbf{r}_1)$ and energy E_{el} for each internuclear separation R_{AB}.

We will then use the electronic energy $E_{\mathrm{el}}(R_{AB})$, together with the nuclear repulsion, as the *potential* in a Schrödinger equation that describes the motion of the nuclei. The overall wavefunction for the molecule will then look like

$$\Psi(\mathbf{r}_i, \mathbf{R}_j) = \psi_{\mathrm{el}}(\mathbf{r}_i, \mathbf{R}_j) \psi_n(\mathbf{R}_j) \tag{6.4}$$

So, we try to solve the molecular S.E. in two steps:

1. Fix the nuclei in position and calculate the electronic wavefunction and energy. Do this many times, mapping out E_{el} as a function of R_{AB}.

2. Take $E_{\mathrm{el}}(R_{AB}) + \frac{Z_A Z_B e^2}{R_{AB}}$ as the effective potential in a *nuclear* Schrödinger equation; solve that to obtain the nuclear wavefunctions and energies. We have already done that part of the problem, mostly by representing $E_{\mathrm{el}}(R_{AB})$ as a power series in R_{AB}. Let us study the molecular electronic problem now, starting with H_2^+.

6.0.2   Simple variational treatment of H_2^+

The S.E. for H_2^+ can be solved exactly in elliptic coordinates, but the math is a little messy and the results are not easily extensible to larger molecules. So, let's try a variational approximation, using as our trial functions sums of atomic-like orbitals centered on the two nuclei. This is an example of the widely important *linear combination of atomic orbital*, or LCAO, method. We guess

$$\psi_{e1} \approx c_A \varphi_A + c_B \varphi_B \tag{6.5}$$

where φ_A is some one-electron function centered at nucleus A and φ_B is the same function but centered at nucleus B. If the two functions are exponential and normalized then they look like those shown in Figure 6.2 and we have

$$\langle \varphi_A | \varphi_A \rangle = \langle \varphi_B | \varphi_B \rangle = 1 \tag{6.6}$$

$$\langle \varphi_A | \varphi_B \rangle = S_{AB} \neq 0 \tag{6.7}$$
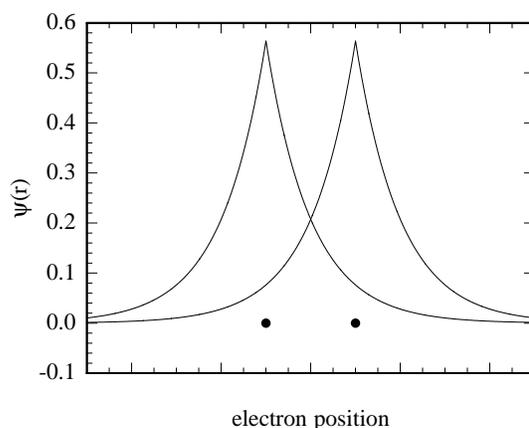

Figure 6.2: Individual $1s$ functions centered on two H atoms with nuclear positions shown as dots. They may be individually normalized but they are *not* orthogonal.

We want to find values of c_A and c_B that minimize the expectation value of the energy. This is exactly the linear variation problem we encountered before; we need solutions of

$$\begin{vmatrix} H_{AA} - E & H_{AB} - ES_{AB} \\ H_{AB} - ES_{AB} & H_{BB} - E \end{vmatrix} = 0 \tag{6.8}$$

where

$$H_{AA} = \langle \varphi_A | H_{\text{el}} | \varphi_A \rangle \tag{6.9}$$

$$H_{BB} = \langle \varphi_B | H_{\text{el}} | \varphi_B \rangle \tag{6.10}$$

$$\langle H_{AB} \rangle = \langle \varphi_A | H_{\text{el}} | \varphi_B \rangle = \langle \varphi_B | H_{\text{el}} | \varphi_A \rangle \tag{6.11}$$

In this problem, $H_{AA} = H_{BB}$ (φ_A and φ_B are the same function and the nuclei

are equivalent) so we have

$$(H_{AA} - E)^2 \quad - \quad (H_{AB} - ES_{AB})^2 = 0 \tag{6.12}$$

$$H_{AA} - E \quad = \quad \pm(H_{AB} - ES_{AB}) \tag{6.13}$$

$$E(1 \mp S_{AB}) \quad = \quad H_{AA} \mp H_{AB} \tag{6.14}$$

$$E \quad = \quad \frac{H_{AA} \mp H_{AB}}{1 \mp S_{AB}} \tag{6.15}$$

So we find two possible values of E,

$$E_- \quad = \quad \frac{H_{AA} - H_{AB}}{1 - S_{AB}}, \text{and} \tag{6.16}$$

$$E_+ \quad = \quad \frac{H_{AA} + H_{AB}}{1 + S_{AB}} \tag{6.17}$$

To find the values of c_A and c_B we substitute each of those values into

$$\left(\begin{array}{cc} H_{AA} - E & H_{AB} - ES_{AB} \\ H_{AB} - ES_{AB} & H_{AA} - E \end{array} \right) \left(\begin{array}{c} c_A \\ c_B \end{array} \right) = 0 \tag{6.18}$$

Substituting E_+ into the first equation gives (after a little algebra) $c_A = c_B$; substituting E_- gives $c_A = -c_B$.

Then, normalizing ψ_{el} gives

$$\psi_+ = (2 + 2S_{AB})^{-\frac{1}{2}}(\varphi_A + \varphi_B) \tag{6.19}$$

$$\psi_- = (2 - 2S_{AB})^{-\frac{1}{2}}(\varphi_A - \varphi_B) \tag{6.20}$$

If I take the φ's as H-atom $1s$ orbitals, with effective nuclear charge ζ, I can draw plots of ψ_+ and ψ_-; these are shown in Figure 6.3 with $\zeta = 1$. (In practice different ζ values are used.)

You should recognize simple bonding and antibonding orbitals. The bonding orbital puts more electron density between the nuclei than the individual H-atom functions; the antibonding MO puts less. With H-Atom $1s$ orbitals as basis functions, the binding energy of H_2^+ (the energy required to separate it into an H atom and a proton) is .065 E_{h}; the experimental value is .102 E_{h}.

6.0.3 Molecular orbitals

Just as you took linear combinations of particle-in-box wavefunctions in your homework to form approximate eigenfunctions of the "particle in sloped box" Hamiltonian, we can take linear combinations of atomic orbitals (LCAOs) to make
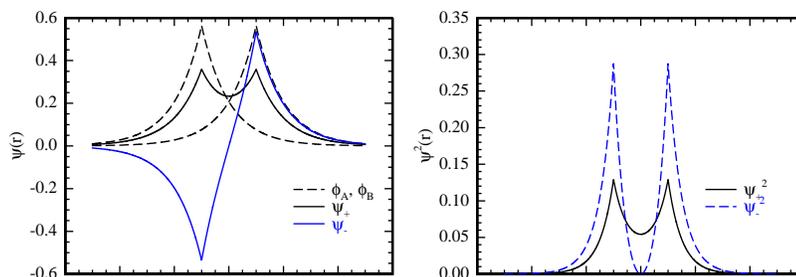
Figure 6.3: The two solutions of the simple variational treatment for H_2^+, using H atom orbitals as the atomic basis functions.

approximate *molecular* orbitals. If we use the two $1s$ hydrogenlike atomic orbitals centered on two atoms, then we form *bonding* and *antibonding* molecular orbitals:

$$\psi_- = N_-(\varphi_{1sA} - \varphi_{1sB}) \tag{6.21}$$

$$\psi_+ = N_+(\varphi_{1sA} + \varphi_{1sB}) \tag{6.22}$$

Figure 6.4 shows a conventional "molecular orbital diagram", indicating what happens to the energy when two atomic orbitals on different atoms are combined to form the two molecular orbitals. The "splitting" depends on the spatial overlap of the two orbitals, as we will see below.
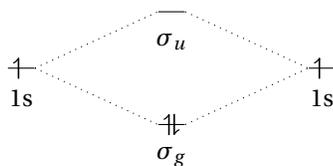

Figure 6.4: Molecular orbital diagram for formation of two molecular orbitals from two $1s$ atomic orbitals. The electrons shown are appropriate for formation of H_2.

In homonuclear diatomic molecules, the two ends of the molecule are just alike. If we place the origin of the coordinate system exactly between the two nuclei, with the z axis along the internuclear axis, then the *inversion* operator $\hat{i}$ that makes $x \to -x$, $y \to -y$, and $z \to -z$ for all electrons must commute with $\hat{H}$. The molecular orbitals can therefore be chosen to be simultaneous eigenfunctions

of $\hat{H}$ and $\hat{i}$. Applying $\hat{i}$ twice must get us back where we started; therefore the eigenvalue of the orbital with respect to $\hat{i}$ must be either 1 or −1. States for which it is +1 are called *g*, for *gerade*, the German word for "even". States for which it is −1 are called *u*, for *ungerade*. Our simple bonding orbital ψ_+ is *g*, and the antibonding ψ_- is *u*.

This should all sound familiar, since the argument is very similar to the one we used in justifying the Pauli principle: the operator $\hat{P}_{12}$, which switches the labels on electrons 1 and 2, is another example of an operator that commutes with the Hamiltonian and must have eigenvalues of 1 or −1. In fact *symmetry operators*, of which $\hat{i}$ and $\hat{P}_{12}$ are both examples, are very widely used in molecular quantum mechanics. A substantial part of most introductory inorganic chemistry courses is devoted to their uses.

The operator $\hat{L}_z$ also commutes with $\hat{H}$ for linear molecules, even though $\hat{L}^2$ no longer does. Molecular orbitals (MOs) can therefore also be labeled with their eigenvalues with respect to $\hat{L}_z$. Just as in atoms, we use lowercase letters to label individual orbitals (one-electron spatial functions), but for molecules we use Greek letters: σ for $m_l = 0$, π for $|m_l| = 1$, δ for $|m_l| = 2$, and ϕ for $|m_l| = 3$.

Our bonding and antibonding orbitals are both made up of *s* atomic orbitals, which have $l = 0$ and therefore necessarily $m_l = 0$; they are therefore σ orbitals. ψ_+ retains its sign when the electron coordinates are inverted through the origin, so it is σ_g. ψ_- changes sign under inversion, so it is σ_u.

A couple of other notation conventions are common. Often a superscript asterisk is used to indicate antibonding orbitals; therefore, in homonuclear diatomic molecules, σ^* means the same as σ_u. (In heteronuclear diatomics and in σ bonds in larger molecules the g/u notation is no longer available but antibonding orbitals are still given asterisks.) Also, it is often useful to indicate which atomic orbitals predominantly make up a particular molecular orbital; we would call our bonding and antibonding orbitals $\sigma_g 1s$ and $\sigma_u 1s$, for example. (There is no agreement on whether the atomic orbital should be given first or last; you'll often see $1s\sigma_g$, as well.) Finally, you will sometimes see orbitals of a particular symmetry (σ, π, etc.) simply given numbers in order of increasing energy; the lowest σ will be called 1σ, the next 2σ or $2\sigma^*$, the next 3σ, and so on.

We can, of course, form molecular orbitals from other atomic orbitals. Consider a MO made from $2p_x$ orbitals on both nuclei. There are two different linear combinations, shown in the top two lines of Figure 6.5. Note that with π orbitals, the bonding ones are *u* and the antibonding ones are *g*.

MOs made from the $2p_z$ are shown in the bottom half of Figure 6.5. Since there is no node that passes through the two nuclei, these are σ orbitals. Alternatively you can think of them as having $M_L = 0$ since they are made from the $2p_0$ atomic orbitals with $m_l = 0$. In fact the best description of diatomic σ MOs arises
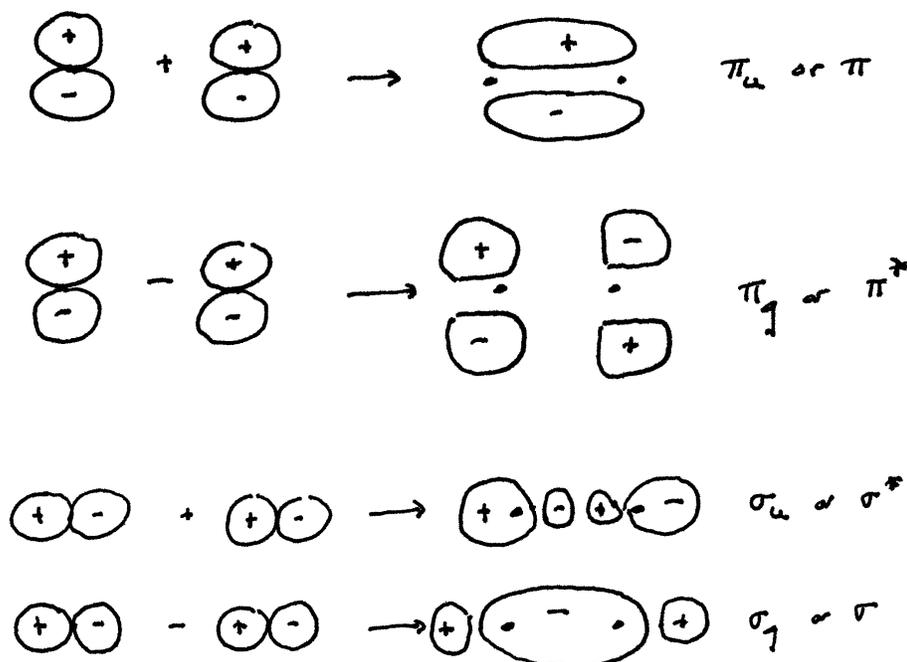
Figure 6.5: Molecular orbitals made from $2p_x$ and $2p_z$ atomic orbitals.

from calculations that allow both the s and p_z (and, for that matter, d_z^2 and so on: all atomic orbitals with $m_l = 0$) to contribute to the function, with weighting coefficients determined variationally. But, you get the same number and types of orbitals that you get in the simple models I have just described.

6.0.4   MOs for diatomic molecules

To describe H_2, we need only place a second electron in the σ_g orbital made from the $1s$ atomic orbitals. Since the two electrons then have the same spatial wavefunction, the spin wavefunction must be the antisymmetric (singlet) one, just as in the ground state of helium.

In fact the simplest version of that MO treatment of H_2, using just the $1s$ atomic orbitals to construct the molecular orbitals and using a simple product wavefunction with both electrons in the bonding orbital, does not give a very good description of the molecule, especially when the bond is stretched. The basic problem is that it gives too much weight to the possibility of having both

electrons near the same nucleus. In modern versions of MO theory this problem is addressed with larger basis sets and *configuration interaction*, a way to write wavefunctions with contributions from more than one electron configuration. However, an early simple treatment gives a good qualitative description, and offers some insights about how bonding works. It modifies the simple MO treatment that uses as its wavefunction

$$\psi_{\text{MO}} = N\psi_+(1)\psi_+(2)[\alpha(1)\beta(2) - \beta(1)\alpha(2)] \tag{6.23}$$

by deleting the terms that look like $\varphi_{1sA}(1)\varphi_{1sA}(2)$ from the spatial function. Let's examine this early treatment before we go on to bigger molecules.

Heitler-London treatment of hydrogen; Coulomb and exchange integrals

Most of the bond strength in H_2 comes from a surprising source. Consider the simplest "valence bond" picture, putting both electrons in the ψ_+ molecular orbitals but ignoring the possibility that both electrons might be in the *same* atomic orbital. This is called the "Heitler-London treatment" of H_2 (*Zeitschrift für Physik* **44**, 455 (1927); an English translation is available in H. Hettema, *Quantum Chemistry: Classic Scientific Papers* (World Scientific, 1998).) In that case the possible spatial wavefunctions are

$$\psi = N[\varphi_{1sA}(1)\varphi_{1sB}(2) \pm \varphi_{1sB}(1)\varphi_{1sA}(2)]. \tag{6.24}$$

Note that the positive linear combination will have to be combined with a singlet (antisymmetric) spin function, and the negative linear combination with a triplet (symmetric) spin function. Different N will be required for the two combinations as well.

When you write out the expectation value of the energy for hydrogen, in addition to ordinary H-atom-like pieces you find two kinds of terms. The first is

$$J = \langle \varphi_{1sA}(1)\varphi_{1sB}(2)| - \frac{1}{r_{B1}} - \frac{1}{r_{A2}} + \frac{1}{r_{12}} + \frac{1}{r_{AB}} |\varphi_{1sA}(1)\varphi_{1sB}(2)\rangle. \tag{6.25}$$

J is called a *Coulomb integral*. It represents the electrostatic interactions of electron 1 around nucleus A with the + charge of nucleus B; the same for electron 2 around B interacting with nucleus A; the $e^- $-$e^-$ repulsion; and the nuclear repulsion.

The second kind of term is

$$K = \langle \varphi_{1sA}(1)\varphi_{1sB}(2)| - \frac{1}{r_{A1}} - \frac{1}{r_{B2}} + \frac{1}{r_{12}} + \frac{1}{r_{AB}} |\varphi_{1sB}(1)\varphi_{1sA}(2)\rangle \tag{6.26}$$

K is called an *exchange integral.* It is a purely quantum-mechanical thing, arising from the requirement of indistinguishability of the two electrons. (It is called a *resonance integral* by some authors). If we hadn't had to make sure that our wavefunction was symmetric (or antisymmetric), we wouldn't get any K terms.

Both J and K, like S, are functions of R_{AB}. J has a simple form in terms of R_{AB}, but K is more complicated. Both J (which is the only contribution to the bonding in H_2^+) and K have minima as functions of R_{AB}, but the minimum in K is much deeper and contributes most of the binding energy of H_2. Both J and K go to zero as $R_{AB} \to \infty$; in order to form a bond, the atomic orbitals involved must overlap in space.

The total energy corresponding to the bonding wavefunction (with the + sign in Eq. (6.24)) of H_2 is

$$E = -1 + \frac{J + K}{1 + S^2}, \tag{6.27}$$

in atomic units, where S is the overlap integral as before. The energy of two separated H atoms is -1 (Hartrees), so the potential curve for the bonding combination with the dissociation asymptote set to zero is

$$E_+ = \frac{J + K}{1 + S^2}. \tag{6.28}$$

The energy for the antibonding combination relative to that of separated H atoms is

$$E_- = \frac{J - K}{1 - S^2}. \tag{6.29}$$

The curves for J and K as functions of R_{AB} are shown in Figure 6.6 along with the resulting potential curves for the two states. The dissociation energy D_e of H_2 is given by the depth of the minimum in the E_+ curve. Using the ordinary H-atom $1s$ orbitals, with nuclear charge $Z = 1$, gives $D_e = 0.116 E_h$ at $R_{AB} = 1.64 a_0$. (This is the curve that's shown.) The experimental value is $D_e = 0.174 E_h$ at $R_{AB} = 1.40 a_0$, so while the qualitative behavior of this basic calculation is reasonable, the quantitative results are not particularly good. A variational calculation that allows Z' to vary (as we did with He) finds the optimum value $Z' = 1.166$ and gives $D_e = 0.139 E_h$ at $R_{AB} = 1.39 a_0$. That result gives a big improvement in the equilibrium bond length, but improves the dissociation energy only modestly.

MO treatment of diatomic He

Returning to the normal MO treatment, consider He_2. To form a helium diatomic molecule, we would put the next two electrons in the antibonding σ_u formed
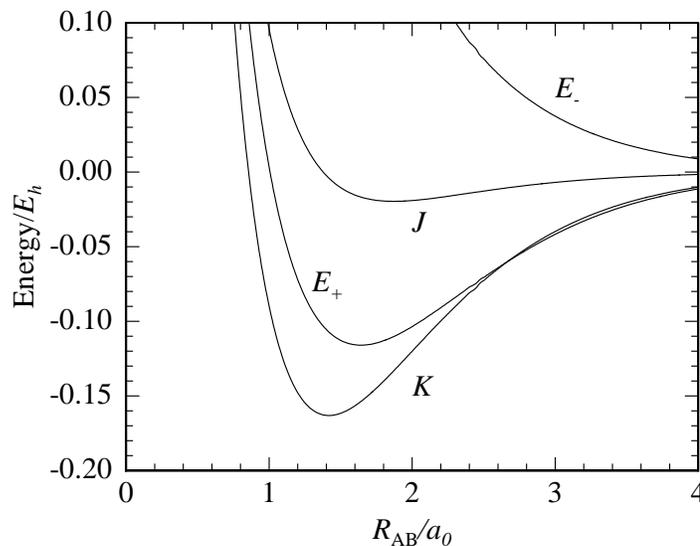
Figure 6.6: Coulomb (J) and exchange (K) integrals for the Heitler-London treatment of H_2, along with the potential curves for the ground and excited states. These curves were computed from expressions in Y. Sugiura, *Zeitschrift für Physik* **45**, 484 (1927). (In that paper, J is called E_1, K is called E_2, and our S^2 is called simply S.)

from the same two $1s$ atomic orbitals. But then the bond order is zero; no ordinary chemical bond forms between a pair of helium atoms.

(If anything in my scientific career can be considered famous, it is this: I was part of the team that found the first convincing experimental evidence that helium forms diatomic molecules at all [*J. Chem. Phys.* **98**, 3564 (1993)]. It generated some controversy, but was eventually confirmed. The bond between two helium atoms is definitely not an ordinary chemical bond, and it is extremely weak: the molecules fall apart if the temperature goes above about 1 mK.)

6.0.5 Second row diatomics

The homonuclear and "nearly-homonuclear" diatomic molecules from the second row of the periodic table all have very similar molecular orbitals. Example MO diagrams for O_2 and N_2 are shown in Figure 6.7.

σ^*

π_1^* π_{-1}^*

$2p_1$ $2p_0$ $2p_{-1}$ $2p_1$ $2p_0$ $2p_{-1}$ $2p_1$ $2p_0$ $2p_{-1}$

π_1 π_{-1}

σ

σ^*

2s σ 2s

σ^*

σ

1s σ^* 1s

σ
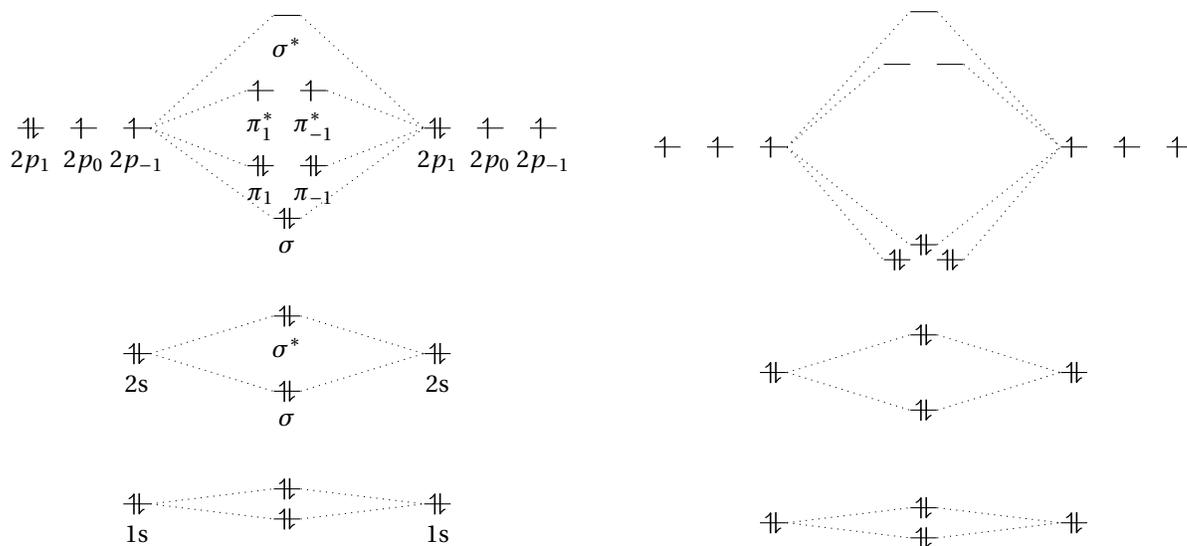
Figure 6.7: Molecular orbital diagram for the ground states of O_2 (left, with labels) and N_2. The energy (vertical) scale is qualitative only, and the two diagrams are not on the same vertical scale.

The lowest two orbitals are the σ_g and σ_u formed from the atomic $1s$ orbitals. The atomic orbitals are localized close to the nuclei and have very little overlap, so these orbitals have similar energies and do not contribute much to bond formation.

The next two MOs are again σ_g and σ_u, formed this time principally from the $2s$ atomic orbitals. In Li_2 the bonding σ_g is occupied by two electrons and the antibonding σ_u is empty. The bond order, defined by half the difference in the numbers of electrons in bonding and antibonding orbitals, is therefore 1.

At this point the next MOs could be either the σ_g formed principally from the $2p_z$ orbitals, or the π_u ones formed from the $2p_x$ and $2p_y$. In fact those two sets of orbitals are very close in energy, and either the σ or the π pair can be lower depending on the nuclear charge and the number of electrons in the molecule. For neutral homonuclear diatomics, the π pair is lower for N_2 and the lighter atoms, while the σ is lower for O_2 and F_2. Both the diagrams are shown in Figure 6.7.

After the bonding σ and π orbitals formed from the $2p$ atomic orbitals, the next MOs are the antibonding π^* and then σ^* formed from the same orbitals (always in that order). At Ne_2 all these orbitals are filled, and we are back to a total bond order of zero.

Table 6.1 gives the bond orders, bond lengths, and bond energies for some first- and second-row homonuclear diatomics. The experimentally determined bond lengths and energies track nicely with the qualitative predictions of this simple molecular orbital theory.

Table 6.1 Properties of light homonuclear diatomic molecules. (From D. A. McQuarrie, *Quantum Chemistry*, University Science Books, 1983).

Species	Bond order	Bond length in pm	Bond energy in kJ/mol	Ground term symbol
H_2^+	$\frac{1}{2}$	106	255	$^2\Sigma$
H_2	1	74	431	$^1\Sigma$
He_2^+	$\frac{1}{2}$	108	251	$^2\Sigma$
He_2	0	~ 6200	$\sim 1 \times 10^{-7}$	$(^1\Sigma)$
Li_2	1	267	105	$^1\Sigma$
Be_2	0			
B_2	1	159	289	$^3\Sigma$
C_2	2	124	599	$^1\Sigma$
N_2	3	110	942	$^1\Sigma$
O_2^+	$2\frac{1}{2}$	112	643	$^2\Pi$
O_2	2	121	494	$^3\Sigma$
O_2^-	$1\frac{1}{2}$	135	395	$^2\Pi$
F_2	1	141	154	$^1\Sigma$
Ne_2	0	~ 400	$\sim 1 \times 10^{-3}$	$(^1\Sigma)$

6.0.6 Term symbols for diatomics

Just as in the atomic case, it is not possible to assign a set of quantum numbers to each individual electron in a diatomic molecule, but it is possible to give quantum numbers that specify eigenvalues of operators that commute with the diatomic molecule Hamiltonian. The two most important such operators are $\hat{S}^2$ (the same as in atoms) and $\hat{L}_z$. The procedure for finding what molecular states can arise from a particular electron configuration is almost the same as that for atoms, but is a little bit easier. You make a table of M_L and M_S values just as before, but for a particular value of $M_L(\text{max})$ you do not need to find all the integers between $M_L(\text{max})$ and $-M_L(\text{max})$, as in the atomic case, but only $M_L(\text{max})$ and $-M_L(\text{max})$ with nothing in between. You still treat the M_S values the same as before. Molecular terms are given the Greek letters corresponding to S, P, D, F; that is Σ, Π, Δ, Φ.

Let's use B_2 as an example. The MO diagram is the one that has the $2p\pi_u$ lower than the $2p\sigma_g$. The diagram therefore looks like that shown in the right side of Figure 6.8. Just as in atoms, any completely filled set of orbitals does not contribute to the total M_L and M_S, so we need consider only the last two electrons, the ones that go into the $2p\pi_u$ orbitals, explicitly. All six of the possible electron diagrams are shown on the right side of the figure, with their values of M_L and M_S.
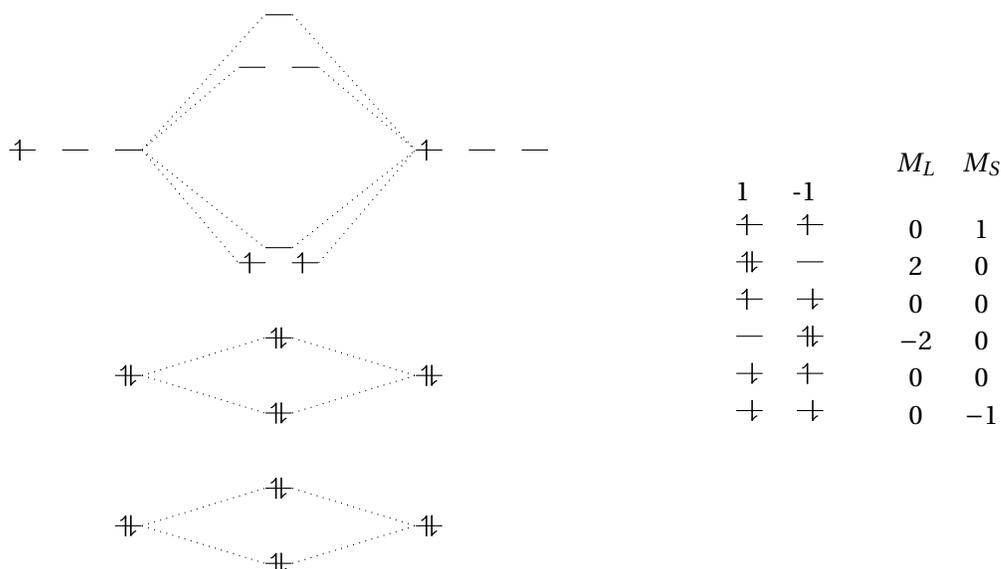

Figure 6.8: MO diagram for B_2 with one possible arrangement of the highest two electrons (left), and the six possible arrangements of those electrons (right).

The largest M_L is 2, and it appears only with $M_S = 0$. The two lines with $M_L = \pm 2$ and $M_S = 0$ make up a $^1\Delta$ state. (Notice that this is where the difference with atomic term symbols comes in. We did not have to look for $M_L = 2, 1, 0, -1$ and -2 with $M_S = 0$, but only for $M_L = \pm 2$.) Mark those two lines off as used.

The next highest M_L is 0, but it occurs with $M_S = 1, 0$ (twice), and -1. We treat M_S just as we did for atoms, so three of those lines $M_S = 1, 0, -1$ make up a $^3\Sigma$ state. We then have only one line remaining; it has $M_L = 0$ and $M_S = 0$, so it makes up a $^1\Sigma$ state.

We have found three possible molecular states: $^1\Delta$, $^3\Sigma$, $^1\Sigma$. Hund's rules apply to molecules just as they do to atoms, so the lowest energy state will be $^3\Sigma$, the next lowest $^1\Delta$, and the highest $^1\Sigma$.

Oxygen has exactly the same set of possible molecular states (its last two elec-

trons go into a pair of π^* orbitals rather than π, but everything works exactly the same way.) Its ground state is therefore also $^3\Sigma$. Triplet molecules have unpaired electrons and are therefore *paramagnetic* (weakly attracted to magnetic field gradients). It is certainly not obvious from looking at a simple Lewis structure of O_2 that it should have unpaired electrons, and the prediction of the paramagnetism of O_2 was one of the early triumphs of molecular orbital theory.

The transition $^1\Delta \leftarrow {}^3\Sigma$ is forbidden in the absence of spin-orbit coupling, but when spin-orbit coupling is included in the Hamiltonian it becomes weakly allowed. The $^1\Delta$ state is about $9000\,\text{cm}^{-1}$ above the $^3\Sigma$, and in fact liquid oxygen has a pale blue color because of the absorption of red photons in this transition. It is also responsible for the blue color of ice that contains dissolved O_2.

Molecules (and for that matter, atoms) in different electronic states can behave quite differently. Ground state oxygen is relatively unreactive (that's why we don't all spontaneously break into flames); $^1\Delta$ oxygen, however, is notoriously reactive, and there is a minor branch of organic chemistry devoted to generating it and getting it to react in desired ways.

Similarly, ground state (^3P) oxygen atoms, while more reactive than ordinary O_2, tend to react slowly. For instance, if ^3P O atoms are mixed with methane at room temperature, it will take many thousands of collisions for them finally to *abstract* H atoms to give OH and CH_3 (which will then react quickly to form a mixture of other products). On the other hand, $O(^1D)$ will react nearly on its first encounter with a methane molecule by *inserting* into a CH bond to give highly excited methanol. This methanol can then dissociate into OH and CH_3, or if it collides with another molecule before falling apart, it can be stabilized and remain intact.