# 1 Introduction

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.

Figure 1: One section of the atomic H emission spectrum, from *Atomic Spectra and Atomic Structure* by G. Herzberg.

The separate lines indicate the electron can only have particular total energies (roughly: be at particular distances from the nucleus). 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. 1870 and 1925, made it clear that $F = ma$ really did not apply to the parts of atoms or to the atoms themselves. Atkins describes four:

1. Black body radiation: if you regard a heated 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 glow like crazy.
2. 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 term "photon". It ex-

tended Planck’s 1900 hypothesis of quantization of oscillator energy levels to quantization of the electromagnetic field itself.

3. 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.
4. Spectra. These cry out for quantization so loudly they are impossible to ignore. One example from my own lab is shown in Figure 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 rotation speed, but only a few select rotation speeds.

Figure 2: $2+1$ resonance enhanced multiphoton ionization spectrum of CO rotationally excited by collisions with He atoms.

Spectra are a reflection of *structure*; if you try to use classical mechanics to describe the behavior of even a 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 75 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. I hope that at the end of our study of QM we can examine some of the philosophically difficult questions, so you can appreciate why eighty 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 do *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!

## 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. I will occasionally indicate the extension to three dimensions and to more particles or extended bodies.

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*:

$$F = ma = m\frac{d^2x}{dt^2}, \tag{1}$$

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.

Finding the solutions to any particular differential equation is not usually easy. Many kinds of equations have been studied, and analytical solutions to many common ones are known and tabulated in books. In addition, it is nearly always possible to find approximate solutions *numerically* for some finite time $t$. There are some theorems about "uniqueness" that are broadly applicable, so you can often be sure that if you have found *some* solution that satisfies both the equation and the initial conditions, that is the only one. Therefore, a profitable way to solve some equations is just to guess-and-try; if you can luck into a solution that works, you know you have the answer.

Eq. (1) is a second-order differential equation, because the highest derivative that appears is a second derivative. It is always possible to convert a second order equation into a pair of "coupled" first order equations. You 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 $p$. Then I have

$$\frac{dx}{dt} = \frac{p}{m} \tag{2}$$

$$\frac{dp}{dt} = F \tag{3}$$

where the second equation 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).

## 2.1 Example: classical harmonic oscillator

Let's use the guess-and-try approach in an example that will show up in several places during the course.

The force for a *simple harmonic oscillator* is given by Hooke's Law: $F = -kx$, where $k$ is a constant with units $\mathrm{kg}/\mathrm{s}^2$. Let us find the trajectory for a particle subject to that force, with the initial conditions $x(0) = 0$ and $v(0) = v_0$. The equation of motion is

$$m\frac{d^2x(t)}{dt^2} = -kx(t). \tag{4}$$

What functions can you think of that "come back" after you differentiate them twice? Two come to mind: exponential fuctions $e^{at}$, and trigonometric functions $\sin(at)$ and $\cos(at)$. In fact this equation can be solved with either, but let me take an educated guess and suggest the function $x(t) = A\sin(\omega t)$, with $A$ and $\omega$ constants we will adjust to fit the differential equation and the initial conditions. Plug that function into the equation of motion:

$$m\frac{d^2x(t)}{dt^2} = -kx(t) \tag{5}$$

$$m\frac{d^2}{dt^2}[A\sin(\omega t)] = -kA\sin(\omega t) \tag{6}$$

$$m\frac{d}{dt}[A\omega\cos(\omega t)] = -kA\sin(\omega t) \tag{7}$$

$$-mA\omega^2\sin(\omega t) = -kA\sin(\omega t) \tag{8}$$

$A\sin(\omega t)$ divides out on both sides of the last equation; it will be satisfied if $k = m\omega^2$ or

$$\omega = \left(\frac{k}{m}\right)^{\frac{1}{2}}. \tag{9}$$

Note that this value of $\omega$ has been determined without using the initial conditions: no matter where the particle started or how it was moving at the beginning, the *frequency* of its motion is determined only by its mass and the force constant $k$.

Now we must use the initial conditions to find $A$. The first initial condition, $x(0) = 0$, is not very useful; since $\sin(0) = 0$, it will be satisfied no matter the value of $A$. (That happened because I was sly and did not include a cosine term in the solution; had the initial position been anything other than $x = 0$, I would have had to do that.) So we turn to the initial velocity. We have $v(0) = v_0$. So

$$v(t) = \frac{d}{dt}[A\sin(\omega t)] \tag{10}$$

$$= A\omega\cos(\omega t) \tag{11}$$

$$v(0) = A\omega\cos(0) \tag{12}$$

$$v_0 = A\omega \tag{13}$$

$$A = \frac{v_0}{\omega} \tag{14}$$

So, we find the solution

$$x(t) = \frac{v_0}{\omega}\sin(\omega t), \tag{15}$$

where

$$\omega = \left(\frac{k}{m}\right)^{\frac{1}{2}}. \tag{16}$$

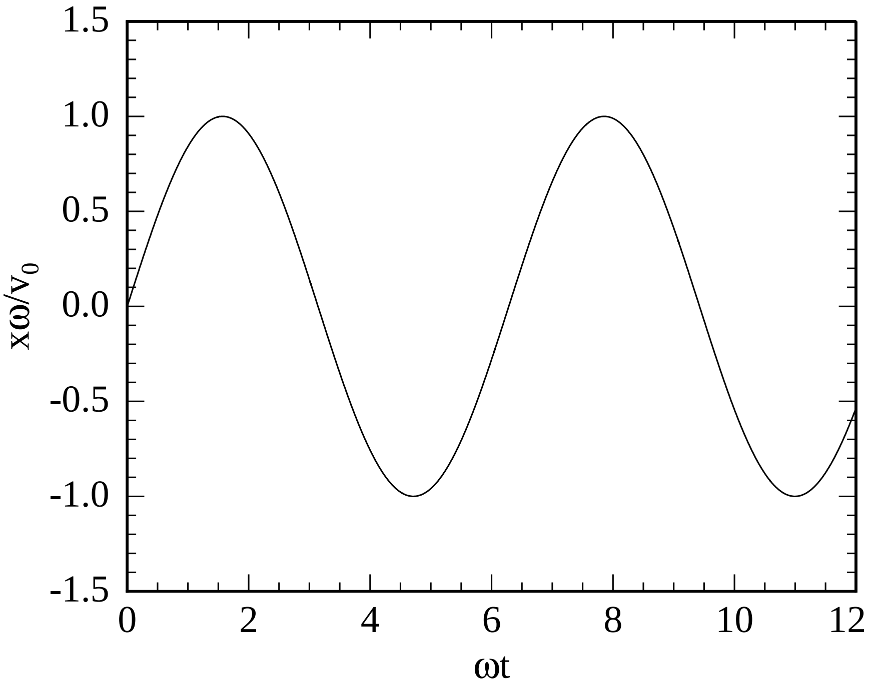


Figure 3: Motion of a particle subject to a Hooke's law force.

Figure 3 shows a graph of the motion of the particle.

We should check that the units are okay: $\omega$ has units $\left(\frac{\mathrm{kg}}{\mathrm{s}^2} \times \frac{1}{\mathrm{kg}}\right)^{\frac{1}{2}} = \mathrm{s}^{-1}$, so $\frac{v_0}{\omega}$ has units $\frac{\mathrm{m}}{\mathrm{s}} \times \frac{1}{\mathrm{s}^{-1}} =$ m, which is appropriate for a position; it checks out. Also note that the argument of the sin function, $\omega t$, is dimensionless; any argument of a pure mathematical function must be.

## 2.2 Potential energy and Hamilton's equations

The Hooke's Law force, $F = -kx$, is one of the simplest examples of a *conservative* force. These are forces for which the concept of potential energy is useful. In a system with potential energy, the force is the negative of the gradient of the potential: $\mathbf{F} = -\nabla V$, where $\nabla$ is the *gradient* operator. This means

$$\mathbf{F} = \left(-\frac{\partial V}{\partial x}\right)\hat{x} + \left(-\frac{\partial V}{\partial y}\right)\hat{y} + \left(-\frac{\partial V}{\partial z}\right)\hat{z}. \tag{17}$$

In one dimension,

$$F = -\frac{dV}{dx}. \tag{18}$$

For our harmonic oscillator example, $V = \frac{1}{2}kx^2$ so $F = -kx$. The shape of the potential is a parabola: you can think of the particle sliding back and forth without friction in a "well" of parabolic shape. The initial conditions I gave before corresponded to the particle beginning exactly at the bottom of the well but moving sideways with speed $v_0$. For conservative systems there is a very useful way of writing down the equations of motion without considering the forces. First you write down the *kinetic energy* $T$

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}, \tag{19}$$

and then you write down the *potential energy* $V$ (which is $\frac{1}{2}kx^2$ in our example). The *Hamiltonian* is just $T + V$:

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2. \tag{20}$$

Now you can write down the equations of motion by applying Hamilton's equations, which are (for a single particle in one dimension)

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} \tag{21}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} \tag{22}$$

In our example, these yield immediately

$$\frac{dx}{dt} = \frac{p}{m} \tag{23}$$

$$\frac{dp}{dt} = -kx \tag{24}$$

just as we had before.

For the single particle in a parabolic well, it is easy to find the equations of motion whether you use Newton's second law directly, as we did before, or use Hamilton's equations. But for more complicated systems it is often much easier to use Hamilton's equations. For example, think of two particles connected by a rigid rod. It is not always easy to figure out how much force is being exerted on the two particles by the rod, but it is usually easy to figure out the potential energy (which is just the sum of the potential energies of the two particles.)

Also, Hamilton's equations adapt very easily to coordinate systems other than Cartesian. In a pendulum, the bob swings in a arc, so both its $x$ and $y$ coordinates change as the pendulum moves; to find the equations of motion, you have to resolve the forces on the bob from gravity and string into $x$ and $y$ components at each bob position, then solve for the $x$ and $y$ motions separately. But the problem is a natural for polar coordinates, since the pendulum is just swinging back and forth without changing its length. With Hamilton's equations, you can use the single "position" coordinate $\theta$ and the single "momentum" coordinate $I\frac{d\theta}{dt}$ (with the moment of inertia $I$ acting as the "mass") and get your equation of motion directly in the polar coordinate. The problem is much easier that way.

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.

# 3 Wave nature of matter

## 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 through a screen that had two parallel slits close together. Beyond the slits was a screen. An alternating pattern of light and dark bands appears on the screen if both slits are open.

Figure 4: Schematic of Young's double slit experiment.

The pattern arises from interference. If the light waves from the two slits arrive "in phase" (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 $\theta$, the difference in distances traveled by waves from the two slits is $d \sin\theta$, where $d$ is the distance between the two slits. If that distance is exactly an integral number of wavelengths—that is, if $d \sin\theta = m\lambda$, where $m$ is an integer—the two waves will arrive at the screen in phase. If it is “$m + \frac{1}{2}$” wavelengths, they will arrive exactly out of phase and cancel.

Figure 5: Variable definitions for Young’s experiment.

It is easy to estimate how far apart the fringes should be. If the distance to the screen is $D$, and we call the distance from the center line to the fringe in question $y$, then $\theta \approx \frac{y}{D}$. (See Figure 5.) For small $\theta$, $\sin\theta \approx \theta$ (where $\theta$ must be measured in radians.) Then we have

$$\frac{y}{D} \approx \frac{m\lambda}{d} \tag{25}$$

$$y \approx \frac{m\lambda D}{d} \tag{26}$$

so the distance between fringes must be

$$\tag{27}$$

$$y_{m+1} - y_m \approx \frac{(m+1)\lambda D}{d} - \frac{m\lambda D}{d} \tag{28}$$

$$\approx \frac{\lambda D}{d} \tag{29}$$

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

## 3.2 Matter waves

In 1924 de Broglie, trying to understand the origin 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. Thompson in England showed that electrons bounced off a crystal or a foil exhibited a diffraction pattern (which looks like many double-slit patterns added together.) These experiments confirmed that material particles could have wavelike properties.

For large objects like cells and cars, the momentum $p$ is quite large; therefore the de Broglie wavelength is small, and quantum effects are not apparent. For small objects, though, interference and diffraction will begin to play a role when the wavelength becomes comparable in size to some feature of the apparatus. 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.

# 4 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.

There are two notations that can be used to state the rules of QM and do calculations. One uses differential equations, functions of coordinates, and integrals; the other uses abstract operators, vectors, and inner products ("dot products"). The vector notation (called Dirac notation) is more compact and a little easier to manipulate, and several of the most important problems in QM can be solved very easily with it. The function notation is often 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.

## 4.1 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 the system. $\psi(\mathbf{r})$ must be defined, continuous, and finite 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.

### 4.1.1 Comments and definitions

1. Any function that meets these conditions is a physically possible wavefunction. There is no requirement of relationship to any particular operators, or that the wavefunction solve any particular differential equation. In particular, Atkins in sections 11–3 through 11-6 implies that the wavefunctions must be solutions of the time-independent Schrödinger equation, that is, eigenfunctions of the Hamiltonian; on p. 301 he even says "the wavefunctions are the eigenfunctions of the hamiltonian operator." This statement is false, and importantly so. Wide ranges of wavefunctions are physically meaningful and possible.

   For specific cases, there are additional requirements on the wavefunction: for example, if the potential is finite everywhere, the first derivative of $\psi(\mathbf{r})$ must be continuous; if the potential is smooth (infinitely differentiable), the wavefunction must also be smooth. We will not worry too much about these individual cases.

2. Any multiple of an acceptable wavefunction, and 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 $\lambda_1$ and $\lambda_2$ are complex numbers. In Dirac notation, we say that the vectors $|\psi\rangle$ and $|\phi\rangle$ are members of a *vector space*.

3. To make the Dirac notation more concrete, recall that vectors are usually written in terms of their components. The state vectors of quantum mechanics are not ordinary vectors in 3-dimensional space with three components, but more abstract vectors in a "function space" that can have many (even infinitely many!) components. But you still handle them the same way you do ordinary vectors. To add two state vectors you add the corresponding components; to multiply a vector by a number you multiply each component by that number.

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix} \tag{30}$$

$$\lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \tag{31}$$

4. The *inner product* (also called *scalar product* or *dot product*) of two real vectors is the number given by adding up the products of correspond-

ing components:

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix} = ac + bd \tag{32}$$

In a complex vector space (where the vector components can be complex numbers), the definition is slightly modified: you use the complex conjugates of the first vector's components rather than the components themselves. So Eq. (32) is replaced by

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix} = a^* c + b^* d \tag{33}$$

In general, for two vectors $|\psi\rangle$ and $|\phi\rangle$ whose components are $c_n$ and $d_n$, the inner product is given by

$$\langle \psi | \phi \rangle = \sum_n c_n^* d_n . \tag{34}$$

Since taking the complex conjugate of a real number does nothing, this definition reduces to Eq. (32) for real vectors. For general state vectors $|\psi\rangle$ and $|\phi\rangle$, the inner product is a complex number.

In the function notation, the inner product is defined by an integral:

$$\langle \psi(\mathbf{r}) | \phi(\mathbf{r}) \rangle = \int d\mathbf{r}\, \psi^*(\mathbf{r}) \phi(\mathbf{r}), \tag{35}$$

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.

The inner product of a vector with itself is just

$$\langle \psi | \psi \rangle = \sum_n c_n^* c_n \tag{36}$$

$$= \sum_n |c_n|^2 . \tag{37}$$

Since $|c_n|^2$ is always a real nonnegative number, the inner product of a vector with itself is always real and positive, since it is the sum of a bunch of real, nonnegative numbers, and cannot be zero unless all the components of the vector are zero (which is not permissible for state vectors.) For ordinary three-dimensional vectors in Cartesian space, the inner product of a vector with itself gives the square of the length of the vector: $r^2 = x^2 + y^2 + z^2$. In the more general case of wavefunctions or state vectors, the inner product of a vector or function with itself continues to behave like the square of a length.

5. Often it is convenient to require that the wavefunction or state vector be *normalized*; that is,

$$\langle \psi | \psi \rangle = \int d\mathbf{r}\, \psi^*(\mathbf{r})\psi(\mathbf{r}) = 1. \tag{38}$$

   In other words, we can require that the length of all our state vectors be 1. If you already have a wavefunction $\psi(\mathbf{r})$, it is easy to make a new one $\psi_{\mathrm{N}}(\mathbf{r})$ that is normalized:

$$\psi_{\mathrm{N}}(\mathbf{r}) = \frac{1}{\sqrt{\int d\mathbf{r}\, \psi^*(\mathbf{r})\psi(\mathbf{r})}}\, \psi(\mathbf{r}) \tag{39}$$

   or

$$|\psi_{\mathrm{N}}\rangle = \frac{1}{\sqrt{\langle \psi | \psi \rangle}}\, |\psi\rangle. \tag{40}$$

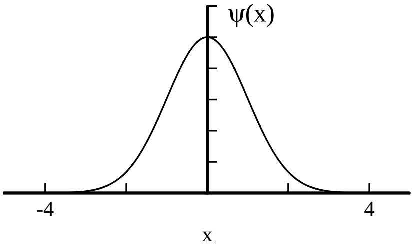
   Notice that a wavefunction for which this is not possible is not an acceptable wavefunction for a physical system anyway. The fraction multiplying $|\psi\rangle$ is called a *normalization constant*.

6. We have not yet said anything about how you use this wavefunction or state vector to predict anything. But the wavefunction contains all the information that *can 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.

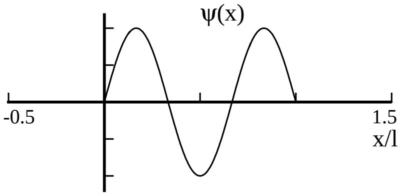
### 4.1.2 Examples

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

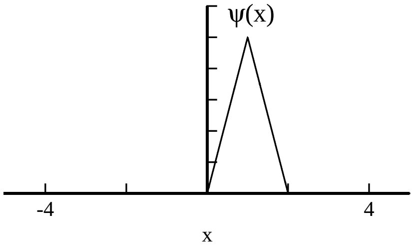
$\psi(x) = e^{-ax^2}$



$$\psi(x) = \begin{cases} \sin\left(\frac{3\pi x}{l}\right) & \text{for } 0 \le x \le l, \\ 0 & \text{otherwise.} \end{cases}$$



$$\psi(x) = \begin{cases} x & \text{for } 0 \le x \le 1, \\ 2 - x & \text{for } 1 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases}$$



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

Notice that $\psi =$ constant is acceptable in the last example because of the finite limits of integration.

## 4.2 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.

### 4.2.1 Comments and definitions

**Definition of operators** An *operator* is a mathematical object that does something to a function; usually, it 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:

$$(\hat{A}+\hat{B})f(x) = \hat{A}f(x) + \hat{B}f(x) \tag{41}$$

$$(\lambda\hat{A})f(x) = \lambda(\hat{A}f(x)) \tag{42}$$

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

$$\hat{A} = x, \tag{43}$$

$$\hat{B} = \frac{d}{dx}. \tag{44}$$

Then

$$\hat{A}\hat{B}f(x) = x\frac{d}{dx}f(x). \tag{45}$$

But

$$\hat{B}\hat{A}f(x) = \frac{d}{dx}\left[xf(x)\right] \tag{46}$$

$$= x\frac{d}{dx}f(x) + f(x)\frac{d}{dx}x \tag{47}$$

$$= x\frac{d}{dx}f(x) + f(x), \tag{48}$$

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,

$$(\hat{A}+\hat{B})^2 = (\hat{A}+\hat{B})(\hat{A}+\hat{B}) \tag{49}$$

$$= \hat{A}^2 + \hat{A}\hat{B} + \hat{B}\hat{A} + \hat{B}^2 \tag{50}$$

$$\neq \hat{A}^2 + 2\hat{A}\hat{B} + \hat{B}^2 \tag{51}$$

**Commutators** The lack of commutativity of operator multiplication is so important 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

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{52}$$

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:

$$[\hat{A}, \hat{B}]f(x) = \hat{A}\hat{B}f(x) - \hat{B}\hat{A}f(x) \tag{53}$$

$$= x\frac{d}{dx}f(x) - \left(x\frac{d}{dx}f(x) + f(x)\right) \tag{54}$$

$$= -f(x) \tag{55}$$

So we would say

$$[\hat{A}, \hat{B}] = -1. \tag{56}$$

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, $[\hat{x}, \hat{\frac{d}{dy}}] = 0$.
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. $\sqrt{af(x)} = \sqrt{a}\sqrt{f(x)} \neq a\sqrt{f(x)}$.

**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 $x$:

$$\hat{x}f(x) = xf(x) \tag{57}$$

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

$$\hat{p}f(x) = -i\hbar\frac{d}{dx}f(x). \tag{58}$$

The most important commutator in QM is

$$[\hat{x}, \hat{p}] = i\hbar \tag{59}$$

Let me show this explicitly:

$$\begin{aligned}
[\hat{x}, \hat{p}]f(x) &= (\hat{x}\hat{p} - \hat{p}\hat{x})f(x) && (60)\\
&= x\left(-i\hbar\frac{d}{dx}f(x)\right) - \left(-i\hbar\frac{d}{dx}\left(xf(x)\right)\right) && (61)\\
&= -i\hbar x\frac{d}{dx}f(x) + i\hbar\left(x\frac{d}{dx}f(x) + f(x)\right) && (62)\\
&= \left(-i\hbar x\frac{d}{dx} + i\hbar x\frac{d}{dx} + i\hbar\right)f(x) && (63)\\
&= i\hbar f(x). && (64)
\end{aligned}$$

So, we write

$$[\hat{x}, \hat{p}] = i\hbar. \tag{65}$$

**Calculating commutators** It is always possible to calculate a commutator in the way I have just demonstrated, by applying the commutator to a "dummy function", working it out as far as you can, then factoring the dummy function back out of your expression. But, in QM it is nearly always possible to write any operator in terms of $\hat{x}$ and $\hat{p}$, and then you can use the identity

$$\hat{p}\hat{x} = \hat{x}\hat{p} - i\hbar \tag{66}$$

Everywhere in your expression for a commutator, if you see $\hat{p}\hat{x}$, replace it with $\hat{x}\hat{p} - i\hbar$. If you see $\hat{x}\hat{p}$, leave it alone. Keep going until $\hat{p}\hat{x}$ does not appear any more.

As an example, let me calculate $[\hat{p}, \hat{x}^2]$. I have

$$\begin{align}
[\hat{p}, \hat{x}^2] &= \hat{p}\hat{x}\hat{x} - \hat{x}\hat{x}\hat{p} \tag{67}\\
&= (\hat{x}\hat{p} - i\hbar)\hat{x} - \hat{x}\hat{x}\hat{p} \tag{68}\\
&= \hat{x}\hat{p}\hat{x} - i\hbar\hat{x} - \hat{x}\hat{x}\hat{p} \tag{69}\\
&= \hat{x}(\hat{x}\hat{p} - i\hbar) - i\hbar\hat{x} - \hat{x}\hat{x}\hat{p} \tag{70}\\
&= \hat{x}\hat{x}\hat{p} - i\hbar\hat{x} - i\hbar\hat{x} - \hat{x}\hat{x}\hat{p} \tag{71}\\
&= -2i\hbar\hat{x} \tag{72}
\end{align}$$

Let me make the same calculation in the other style:

$$\begin{align}
[\hat{p}, \hat{x}^2]f(x) &= (\hat{p}\hat{x}^2 - \hat{x}^2\hat{p})f(x) \tag{73}\\
&= -i\hbar\frac{d}{dx}\left(x^2 f(x)\right) - \left(-i\hbar x^2 \frac{d}{dx} f(x)\right) \tag{74}\\
&= -i\hbar\left(x^2\frac{d}{dx}f(x) + 2xf(x)\right) - \left(-i\hbar x^2\frac{d}{dx}f(x)\right) \tag{75}\\
&= \left(-i\hbar x^2\frac{d}{dx} - 2i\hbar x + i\hbar x^2\frac{d}{dx}\right) f(x) \tag{76}\\
&= -2i\hbar x f(x) \tag{77}
\end{align}$$

So,

$$[\hat{p}, \hat{x}^2] = -2i\hbar x. \tag{78}$$

I find the first calculation easier than the second, but they give the same answer.

**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

$$\hat{A}\left(3x^4\right) = \frac{d^2}{dx^2}\left(3x^4\right) \tag{79}$$

$$= \frac{d}{dx}\left(12x^3\right) \tag{80}$$

$$= 36x^2, \tag{81}$$

and

$$\hat{A}(e^{4x}) = \frac{d^2}{dx^2}\left(e^{4x}\right) \tag{82}$$

$$= \frac{d}{dx}\left(4e^{4x}\right) \tag{83}$$

$$= 16e^{4x}. \tag{84}$$

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

$$\hat{A}f(x) = \lambda f(x) \tag{85}$$

where $\lambda$ (the eigenvalue) is a constant.

If $\hat{A}$ is a linear operator, and $f(x)$ is an eigenfunction with eigenvalue $\lambda$, 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 $\alpha$:

$$\hat{A}(\alpha f(x)) = \alpha\hat{A}f(x) \tag{86}$$

$$= \alpha\lambda f(x) \tag{87}$$

$$= \lambda\alpha f(x) \tag{88}$$

$$= \lambda(\alpha f(x)) \qquad \text{QED.} \tag{89}$$

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*.

In the vector notation, an operator can be represented by a matrix:

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \tag{90}$$

Then a vector $|\psi\rangle$ is called an eigenvector if

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle, \tag{91}$$

where, as usual, $\lambda$ is a constant called the eigenvalue. (You operate with $\hat{A}$ on $|\psi\rangle$ with ordinary matrix-vector multiplication; if the elements of $\hat{A}$ are $a_{ij}$, and the components of $|\psi\rangle$ are $c_j$, then

$$\left(\hat{A}|\psi\rangle\right)_n = \sum_j a_{nj} c_j. \tag{92}$$

In other words, to get the $n$th element of the product vector, you take the dot product of the $n$th row of the matrix $\hat{A}$ with the vector $|\psi\rangle$.)

For example, examine

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{93}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{94}$$

$$= -1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{95}$$

so $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 1, and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector with eigenvalue $-1$.

Notice that the matrix above is *symmetric*: if we *transpose* the matrix (flip it about the diagonal, or interchange rows and columns), we get the same matrix back. In other words, $a_{ij} = a_{ji}$. Such matrices always have eigenvectors.

In QM, the matrices often have complex elements. The analog of a symmetric matrix for complex matrices is called a Hermitian matrix; for a Hermitian matrix, $a_{ij} = a_{ji}^*$. For real matrices Hermitian means the same as symmetric, since taking the complex conjugate of a real number does nothing. An example of a Hermitian matrix is

$$\begin{pmatrix} 0 & i \\ -i & 1 \end{pmatrix} \tag{96}$$

Hermitian matrices always have real elements along the diagonal, and corresponding elements on opposite sides of the diagonal are complex conjugates of each other.

In vector notation, if $\hat{A}$ is Hermitian, then

$$\langle\psi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\psi\rangle^* \tag{97}$$

(notice that the two symbols $\phi$ and $\psi$ have been switched; this corresponds to transposing the matrix.) Notice that if you let the two functions $\phi$ and $\psi$ be the same in Eq. (97), you get

$$\langle\phi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\phi\rangle^*. \tag{98}$$

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

In the function notation, the definition of a Hermitian operator is the same as Eq. (97); that is, $\hat{A}$ is a Hermitian operator if for any functions $\psi(x)$ and $\phi(x)$,

$$\int dx\, \phi(x)^* \hat{A}\psi(x) = \left(\int dx\, \psi(x)^* \hat{A}\phi(x)\right)^*. \tag{99}$$

Hermitian operators have some important properties.

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

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle \tag{100}$$

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

$$\langle\psi|\hat{A}|\psi\rangle = \langle\psi|\lambda|\psi\rangle \tag{101}$$

$$\langle\psi|\hat{A}|\psi\rangle = \lambda\langle\psi|\psi\rangle \tag{102}$$

Since $\langle\psi|\hat{A}|\psi\rangle$ and $\langle\psi|\psi\rangle$ are both real, $\lambda$ 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.

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

$$\hat{A}|\phi\rangle = \lambda|\phi\rangle, \tag{103}$$

then

$$\begin{align} \langle\phi|\hat{A}|\psi\rangle &= \langle\psi|\hat{A}|\phi\rangle^* \tag{104}\\ &= \langle\psi|\lambda|\phi\rangle^* \tag{105}\\ &= \lambda^*\langle\psi|\phi\rangle^* \tag{106}\\ &= \lambda\langle\psi|\phi\rangle^* \tag{107}\\ &= \lambda\langle\phi|\psi\rangle \tag{108}\end{align}$$

2. The eigenvectors of a Hermitian operator corresponding to different eigenvalues are orthogonal. Assume

$$\hat{A}|\phi_1\rangle = \lambda_1|\phi_1\rangle \tag{109}$$

and

$$\hat{A}|\phi_2\rangle = \lambda_2|\phi_2\rangle. \tag{110}$$

Then if we act on the left of Eq. (109) with $\langle\phi_2|$, we have

$$\langle\phi_2|\hat{A}|\phi_1\rangle = \langle\phi_2|\lambda_1|\phi_1\rangle = \lambda_1\langle\phi_2|\phi_1\rangle. \tag{111}$$

Since $\hat{A}$ is Hermitian, we can do

$$\langle\phi_1|\hat{A}|\phi_2\rangle^* = \lambda_1\langle\phi_2|\phi_1\rangle \tag{112}$$

Now from Eq. (110) we have

$$\langle\phi_1|\lambda_2|\phi_2\rangle^* = \lambda_1\langle\phi_2|\phi_1\rangle \tag{113}$$

and since $\lambda_2$ is real, we have

$$\lambda_2\langle\phi_1|\phi_2\rangle^* = \lambda_1\langle\phi_2|\phi_1\rangle \tag{114}$$

Now we reverse the inner product to find

$$\lambda_2\langle\phi_2|\phi_1\rangle = \lambda_1\langle\phi_2|\phi_1\rangle \tag{115}$$

$$(\lambda_2 - \lambda_1)\langle\phi_2|\phi_1\rangle = 0 \tag{116}$$

so that if $\lambda_2 \neq \lambda_1$, $\langle\phi_2|\phi_1\rangle$ must be zero; therefore $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthogonal. QED.

**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,

$$\begin{pmatrix} 18 \\ -4 \\ 6 \end{pmatrix} = 18\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + -4\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 6\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{117}$$

Note that these basis vectors are orthogonal,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0, \tag{118}$$

and normalized,

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} = 1. \tag{119}$$

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:

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + c_3|\phi_3\rangle + \cdots, \tag{120}$$

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 $\langle\phi_n|$:

$$\langle\phi_2|\psi\rangle = c_1\langle\phi_2|\phi_1\rangle + c_2\langle\phi_2|\phi_2\rangle + c_3\langle\phi_2|\phi_3\rangle + \cdots \tag{121}$$

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

$$c_2 = \langle\phi_2|\psi\rangle; \tag{122}$$

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!

## 4.3 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.

**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.

### 4.3.1 Continuous and discrete eigenvalue spectra

You have already seen what an eigenvalue is. An operator can have either a *continuous* eigenvalue spectrum or a *discrete* one (or some combination of the two); we need to examine these two cases because their treatment in the next postulate is somewhat different.

**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

$$-i\hbar\frac{d}{dx}\psi(x) = p\psi(x), \tag{123}$$

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

$$\phi_p(x) = e^{\frac{ipx}{\hbar}} \tag{124}$$

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

$$\begin{aligned} -i\hbar\frac{d}{dx}e^{\frac{ipx}{\hbar}} &= -i\hbar(\frac{ip}{\hbar})e^{\frac{ipx}{\hbar}} && (125) \\ &= pe^{\frac{ipx}{\hbar}} && (126) \end{aligned}$$

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.)

**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 eigenvalues.

**Discrete spectra** You have already seen one example of an operator with discrete eigenvalues, the symmetric matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, whose eigenvalues are 1 and -1. (That operator does correspond to some measurable properties, relevant for example to nmr experiments.) Another example is the total energy for the harmonic oscillator, whose operator is

$$\hat{T} + \hat{V} = \frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2. \tag{127}$$

This operator has eigenvalues $\frac{1}{2}\hbar\omega$, $\frac{3}{2}\hbar\omega$, $\frac{5}{2}\hbar\omega$, and so on. (We have not yet derived this result, but it's coming.)

## 4.4 Predicting measurement results

### 4.4.1 Nondegenerate eigenfunctions

I want to concentrate on predicting the results of measurements whose operators have nondegenerate eigenfunctions, first. 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 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.

**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:

$$|\psi\rangle = c_0|\phi_0\rangle + c_1|\phi_1\rangle + c_2|\phi_2\rangle + \cdots, \tag{128}$$

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

$$c_n = \langle\phi_n|\psi\rangle. \tag{129}$$

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.

**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 $|\langle\phi_a|\psi\rangle|^2$, 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

$$P(2 \leq a \leq 4) = \int_2^4 P(a)\,da. \tag{130}$$

**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,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = \langle\psi|\psi\rangle = 1; \tag{131}$$

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 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

$$\langle a\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{132}$$

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.

### 4.4.2 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.

### 4.4.3 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 just $|\phi_n\rangle$, 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.

**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”.

### 4.4.4 Time evolution

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

$$i\hbar\frac{d|\psi\rangle}{dt} = \hat{H}|\psi\rangle \tag{133}$$

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 $\hat{H}$:

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle. \tag{134}$$

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 $\hat{H}$:

$$|\psi(t)\rangle = \sum_n c_n(t)|\phi_n\rangle, \tag{135}$$

where

$$c_n(t) = \langle\phi_n|\psi(t)\rangle. \tag{136}$$

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:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{137}$$

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

$$i\hbar\frac{d}{dt}\langle\phi_n|\psi(t)\rangle = \langle\phi_n|\hat{H}|\psi(t)\rangle \tag{138}$$

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:

$$i\hbar\frac{d}{dt}\langle\phi_n|\psi(t)\rangle = \langle\phi_n|E_n|\psi(t)\rangle \tag{139}$$

$$i\hbar\frac{d}{dt}\langle\phi_n|\psi(t)\rangle = E_n\langle\phi_n|\psi(t)\rangle \tag{140}$$

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

$$i\hbar \frac{d}{dt} c_n(t) = E_n c_n(t) \tag{141}$$

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

$$c_n(t) = c_n(0) e^{-\frac{iE_n t}{\hbar}}, \tag{142}$$

as we can verify by substitution:

$$i\hbar \frac{d}{dt} \left( c_n(0) e^{-\frac{iE_n t}{\hbar}} \right) = E_n c_n(0) e^{-\frac{iE_n t}{\hbar}} \tag{143}$$

$$i\hbar \left( \frac{-iE_n}{\hbar} \right) c_n(0) e^{-\frac{iE_n t}{\hbar}} = E_n c_n(0) e^{-\frac{iE_n t}{\hbar}} \tag{144}$$

$$E_n c_n(0) e^{-\frac{iE_n t}{\hbar}} = E_n c_n(0) e^{-\frac{iE_n t}{\hbar}} \tag{145}$$

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

$$|\psi(0)\rangle = \sum_n c_n(0) |\phi_n\rangle, \tag{146}$$

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

$$|\psi(t)\rangle = \sum_n c_n(0) e^{\frac{-iE_n t}{\hbar}} |\phi_n\rangle \tag{147}$$

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.

### 4.4.5 Stationary states

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

$$|\psi(0)\rangle = |\phi_n\rangle \tag{148}$$

$$|\psi(t)\rangle = e^{\frac{-iE_n t}{\hbar}} |\phi_n\rangle \tag{149}$$

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

$$\hat{A}|v_m\rangle = b_m|v_m\rangle \tag{150}$$

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

$$\begin{aligned} P(b_m) &= |\langle v_m|\psi(t)\rangle|^2 && (151)\\ &= |\langle v_m|e^{\frac{-iE_n t}{\hbar}}|\phi_n\rangle && (152)\\ &= |e^{\frac{-iE_n t}{\hbar}}\langle v_m|\phi_n\rangle|^2 && (153)\\ &= |e^{\frac{-iE_n t}{\hbar}}|^2|\langle v_m|\phi_n\rangle|^2 && (154)\\ &= |\langle v_m|\phi_n\rangle|^2 && (155) \end{aligned}$$

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 $\theta$ 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.

### 4.4.6 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.

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; 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 $\hat{H}$,

$$\hat{H}|\psi\rangle = E|\psi\rangle, \tag{156}$$

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.

### 4.4.7 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

$$\frac{d}{dt}\langle\psi|\hat{A}|\psi\rangle = \frac{1}{i\hbar}\langle\psi|[\hat{A},\hat{H}]|\psi\rangle \tag{157}$$

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

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{158}$$

yields the results

$$\frac{d}{dt}\langle\hat{x}\rangle = \frac{1}{m}\langle\hat{p}\rangle \tag{159}$$

$$\frac{d}{dt}\langle\hat{p}\rangle = \langle-\nabla V(\hat{x})\rangle \tag{160}$$

This result is known as Ehrenfest's theorem. 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.

# 5 Harmonic oscillator

We now want to begin studying a simple but important and realistic quantum mechanical system, the one-dimensional harmonic oscillator. This system makes a good model for the vibrations of most molecules, the motion of atoms in a crystal, and (with more subtlety) the behavior of light.

The classical Hamiltonian for a particle in a harmonic (parabolic) potential is

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2, \tag{161}$$

where $k$ is the force constant of the restoring force and $m$ is the particle's mass. A classical particle in such a potential has a sinusoidal motion with angular frequency $\omega = \sqrt{k/m}$, so we can also write

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{162}$$

We obtain the quantum mechanical Hamiltonian by replacing $p$ and $x$ with $\hat{p}$ and $\hat{x}$, respectively:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{163}$$

The first task in understanding the behavior of a quantum harmonic oscillator is to find the eigenvalues and corresponding eigenvectors or eigenfunctions of $\hat{H}$; if we find those, we can then describe the behavior of any possible wavefunction in a simple way.

The "eigen-equation" for $\hat{H}$ is

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle \tag{164}$$

## 5.1 Raising and lowering operators

In order to simplify the math that follows, let me introduce two new operators that are linear combinations of $\hat{x}$ and $\hat{p}$:

$$\hat{a} = \left(\frac{1}{2\hbar}\right)^{\frac{1}{2}}\left(\sqrt{m\omega}\hat{x} + \frac{i}{\sqrt{m\omega}}\hat{p}\right) \tag{165}$$

$$\hat{a}^\dagger = \left(\frac{1}{2\hbar}\right)^{\frac{1}{2}}\left(\sqrt{m\omega}\hat{x} - \frac{i}{\sqrt{m\omega}}\hat{p}\right) \tag{166}$$

Then it is straightforward to show the following:

$$\hat{x} = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} (\hat{a}^\dagger + \hat{a}) \tag{167}$$

$$\hat{p} = i\left(\frac{\hbar m\omega}{2}\right)^{\frac{1}{2}} (\hat{a}^\dagger - \hat{a}) \tag{168}$$

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{169}$$

and substituting those expressions for $\hat{x}$ and $\hat{p}$ into $\hat{H}$ and simplifying gives

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \tag{170}$$

The operators $\hat{a}$ and $\hat{a}^\dagger$ are *not* Hermitian; they do not correspond to any measureable property. They are, however, extremely useful in calculations.

We know that $\hat{H}$ must have a complete set of eigenfunctions. Call them $|\phi_n\rangle$. I want to give (without proof, for now) the behavior of $\hat{a}$ and $\hat{a}^\dagger$ on the $|\phi_n\rangle$:

$$\hat{a}|\phi_n\rangle = \sqrt{n}|\phi_{n-1}\rangle \tag{171}$$

$$\hat{a}^\dagger|\phi_n\rangle = \sqrt{n+1}|\phi_{n+1}\rangle \tag{172}$$

$\hat{a}$ is called a *lowering operator* (or, in more colorful terminology, a destruction or annihilation operator). It changes an eigenfunction of $\hat{H}$ with $n$ quanta of excitation into one with $n-1$ quanta. $\hat{a}^\dagger$ is a *raising operator* (or creation operator). It changes an eigenfunction with $n$ quanta into one with $n+1$ quanta.

## 5.2 Energy eigenvalues

I have not said what the restrictions on $n$ are yet; $n$ must be a nonnegative integer. So we can immediately work out the eigenvalues of the harmonic oscillator Hamiltonian:

$$\hat{H}|\phi_n\rangle = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|\phi_n\rangle \tag{173}$$

$$= \hbar\omega\left(\hat{a}^\dagger\hat{a}|\phi_n\rangle + \frac{1}{2}|\phi_n\rangle\right) \tag{174}$$

$$= \hbar\omega\left(\hat{a}^\dagger\sqrt{n}|\phi_{n-1}\rangle + \frac{1}{2}|\phi_n\rangle\right) \tag{175}$$

$$= \hbar\omega\left(\sqrt{n}\hat{a}^\dagger|\phi_{n-1}\rangle + \frac{1}{2}|\phi_n\rangle\right) \tag{176}$$

$$= \hbar\omega\left(\sqrt{n}\sqrt{(n-1)+1}|\phi_n\rangle + \frac{1}{2}|\phi_n\rangle\right) \tag{177}$$

$$= \hbar\omega\left(n|\phi_n\rangle + \frac{1}{2}|\phi_n\rangle\right) \tag{178}$$

$$= \hbar\omega\left(n + \frac{1}{2}\right)|\phi_n\rangle \tag{179}$$

Therefore, the eigenvalues of the harmonic oscillator Hamiltonian are $(n + \frac{1}{2})\hbar\omega$, where $n$ is an integer greater than or equal to zero.

## 5.3 Zero-point energy

Notice that the lowest state of the harmonic oscillator does not have zero energy, but instead has energy $\frac{1}{2}\hbar\omega$. This unavoidable energy is called *zero-point energy* and is characteristic of all quantum mechanical particles that are confined in some way. At the core, it arises from the uncertainty principle; since the particle's position is restricted by the potential, its momentum must have some uncertainty, and that means it must have some kinetic energy on average. A classical harmonic oscillator, by contrast, has $E = 0$ as its lowest energy state; that corresponds to the particle sitting motionless at the bottom of the potential well.

## 5.4 Energy eigenfunctions

If I put the explicit expressions for $\hat{x}$ and $\hat{p}$ into Eq. (163), I get

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2}{2}x^2\right)\phi_n(x) = E\phi_n(x), \tag{180}$$

or, after dividing by $-\frac{\hbar^2}{2m}$,

$$\left(\frac{d^2}{dx^2} + \frac{2mE_n}{\hbar^2} - \beta^2 x^2\right)\phi_n(x) = 0 \tag{181}$$

where $\beta = \frac{m\omega}{\hbar}$.

There are several ways to find the solutions to Eq. (181). The classical way in the theory of differential equations is to use a power series expansion of $\phi_n$, and to work out the expansion coefficients. I will do something different: show you the lowest-energy eigenfunction, prove that it solves the equation, then show that the others can be constructed by repeated applications of $\hat{a}^\dagger$ followed by normalization.

It is easy to verify that

$$\phi_0(x) = \left(\frac{\beta}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\beta}{2}x^2} \tag{182}$$

solves Eq. (181) if $E_0 = \frac{1}{2}\hbar\omega$. To prove it, work out the first two derivatives of $\phi_0(x)$:

$$\frac{d}{dx}\phi_0(x) = \left(\frac{\beta}{\pi}\right)^{\frac{1}{4}} (-\beta x)e^{-\frac{\beta}{2}x^2} \tag{183}$$

$$\frac{d^2}{dx^2}\phi_0(x) = \left(\frac{\beta}{\pi}\right)^{\frac{1}{4}} (\beta^2 x^2 - \beta)e^{-\frac{\beta}{2}x^2} \tag{184}$$

Substitute into Eq. (181):

$$\left(\beta^2 x^2 - \beta + \frac{2mE_0}{\hbar^2} - \beta^2 x^2\right)\left(\frac{\beta}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\beta}{2}x^2} = 0 \tag{185}$$

This equation is solved if

$$\left(\beta^2 x^2 - \beta + \frac{2mE_0}{\hbar^2} - \beta^2 x^2\right) = 0 \tag{186}$$

which will be true if $E_0 = \frac{1}{2}\hbar\omega$. So we know now that $\phi_0(x)$ is an eigenfunction of $\hat{H}$ with eigenvalue $\frac{1}{2}\hbar\omega$.

### 5.4.1 Behavior of $\hat{a}^\dagger$ and $\hat{a}$

Now I want to show that $\hat{a}^\dagger$ does what I said it does: it raises the quantum number of an eigenfunction of $\hat{H}$.

First, let's make it clear that the eigenfunctions of $\hat{H}$ are the same as the eigenfunctions of the "number operator" $\hat{a}^\dagger\hat{a}$. If

$$\hat{a}^\dagger\hat{a}|\phi_n\rangle = \lambda_n|\phi_n\rangle, \tag{187}$$

then we can add $\frac{1}{2}|\phi_n\rangle$ to each side and multiply both sides by $\hbar\omega$ to get

$$\hbar\omega\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)|\phi_n\rangle = \hbar\omega\left(\lambda_n+\frac{1}{2}\right)|\phi_n\rangle \tag{188}$$

so an eigenfunction of $\hat{a}^\dagger\ \hat{a}$ with eigenvalue $\lambda_n$ is an eigenvalue of $\hat{H}$ with eigenvalue $\hbar\omega\left(\lambda_n+\frac{1}{2}\right)$.

Now let me prove the following theorem:

**Theorem 1** If $|\phi_n\rangle$ is an eigenfunction of $\hat{a}^\dagger\ \hat{a}$ with eigenvalue $\lambda_n$, that is, if

$$\hat{a}^\dagger\hat{a}|\phi_n\rangle = \lambda_n|\phi_n\rangle, \tag{189}$$

then $\hat{a}^\dagger\ |\phi_n\rangle$ is also an eigenfunction of $\hat{a}^\dagger\ \hat{a}$ but with the eigenvalue $\lambda_n+1$.

*Proof*: We want to show that

$$\hat{a}^\dagger\hat{a}\left(\hat{a}^\dagger|\phi_n\rangle\right) = (\lambda_n+1)(\hat{a}^\dagger|\phi_n\rangle) \tag{190}$$

Begin by rearranging the operators, and using the commutation relation $[\hat{a},\hat{a}^\dagger]=1$:

$$\begin{aligned}\hat{a}^\dagger\hat{a}\left(\hat{a}^\dagger|\phi_n\rangle\right) &= \hat{a}^\dagger(\hat{a}\hat{a}^\dagger)|\phi_n\rangle && (191)\\ &= \hat{a}^\dagger(1+\hat{a}^\dagger\hat{a})|\phi_n\rangle && (192)\end{aligned}$$

But applying Eq. (189) gives

$$\begin{aligned}&= \hat{a}^\dagger(1+\lambda_n)|\phi_n\rangle && (193)\\ &= (\lambda_n+1)\hat{a}^\dagger|\phi_n\rangle \text{ QED.} && (194)\end{aligned}$$

There is a similar theorem about the lowering operator $\hat{a}$:

**Theorem 2** If $|\phi_n\rangle$ is an eigenfunction of $\hat{a}^\dagger\ \hat{a}$ with eigenvalue $\lambda_n \neq 0$, that is, if

$$\hat{a}^\dagger \hat{a} |\phi_n\rangle = \lambda_n |\phi_n\rangle, \tag{195}$$

then $\hat{a}\ |\phi_n\rangle$ is also an eigenfunction of $\hat{a}^\dagger\ \hat{a}$ but with the eigenvalue $\lambda_n - 1$.

I will not prove this one, but you probably can by following the same sort of reasoning I just outlined.

### 5.4.2 Generating the eigenfunctions of $\hat{H}$

You can see how we can figure out more eigenfunctions of $\hat{H}$, given that we know one of them: apply $\hat{a}^\dagger$ (in its explicit differential form: it's tedious but not hard at all) to the wavefunction $\phi_0(x)$ to get a new function with an eigenvalue one higher; normalize that new function; repeat until tired. We might also be able to work our way *down* from $\phi_0(x)$, but $\hat{a}\phi_0(x) = 0$ and we have reached the bottom of the ladder.

Applying this prescription generates a series of functions that look like

$$\phi_n(x) = N_n H_n(x) e^{-\frac{\beta}{2}x^2}, \tag{196}$$

where $H_n(x)$ is a polynomial in $x$ called a *Hermite polynomial*. The Hermite polynomials for even $n$ contain only even powers of $x$, and those for odd $n$ contain only odd powers, so the overall wavefunctions all have either even or odd symmetry. The highest power of $x$ in the Hermite polynomial $H_n(x)$ is $n$. Atkins gives the first few Hermite polynomials in Table 12.1 on p. 325. He encourages you to use the properties of Hermite polynomials in solving problems, but it is much cleaner and quicker to rewrite the $\hat{x}$ and $\hat{p}$ operators in terms of $\hat{a}$ and $\hat{a}^\dagger$ and work directly with those.

## 5.5 Nodes

The most interesting thing about the wavefunctions is that the number of *nodes*, or passages through zero, goes up as the energy goes up. The ground state, $\phi_0(x)$, has no nodes; it is positive everywhere. The first excited state, $\phi_1(x)$, has a node at $x = 0$. The second excited state has two nodes, and so on. This is a general principle in quantum mechanics: for the same system, if two wavefunctions have different numbers of nodes, the one with more nodes will have higher energy. In the harmonic oscillator case, the number of nodes is just equal to $n$. This makes for a very easy way to figure out which value of $n$ a particular eigenfunction corresponds to: just count the nodes in the graph!

## 5.6 Harmonic oscillator as model of diatomic molecule vibrations

### 5.6.1 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 electron's 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. 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 6. 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.

In Figure 6 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.

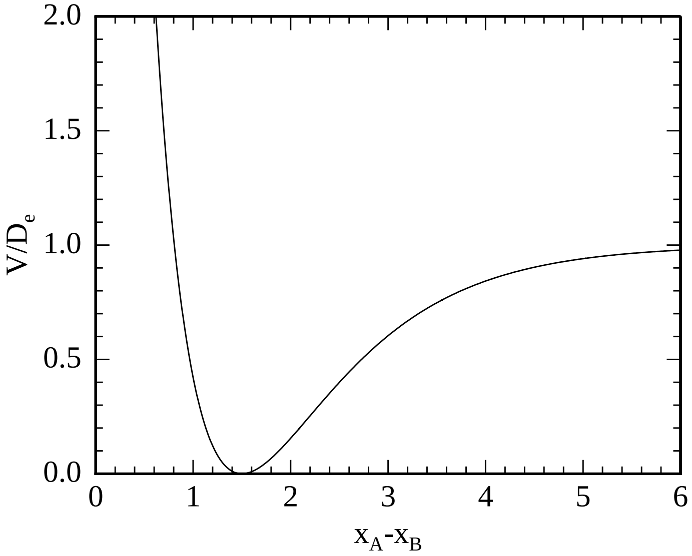


Figure 6: Typical diatomic molecule effective potential curve.

### 5.6.2 Reduction to one-particle problem

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

$$H = \frac{p_A^2}{2m_A} + \frac{p_B^2}{2m_B} + V(x_A - x_B) \tag{197}$$

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,

$$X = \frac{m_A x_A + m_B x_B}{m_A + m_B} \tag{198}$$

$$x = x_A - x_B \tag{199}$$

in terms of which the old variables are

$$x_A = X + \frac{m_B}{m_A + m_B} x \tag{200}$$

$$x_B = X - \frac{m_A}{m_A + m_B} x \tag{201}$$

Now since

$$p_A = m_A \frac{dx_A}{dt}, \tag{202}$$

and similarly for $p_B$, by differentiating Eq. (200) and Eq. (201) you can find a new form of the classical Hamiltonian written in terms of $X$ and $x$. When you do this (homework!) you find that all the terms involving products of time derivatives of $X$ and $x$ cancel, leaving you with the following simple expression:

$$H = \frac{P^2}{2M} + \frac{p^2}{2\mu} + V(x), \tag{203}$$

where

$$M = m_A + m_B, \tag{204}$$

$$\mu = \frac{m_A m_B}{m_A + m_B}, \tag{205}$$

$$P = M \frac{dX}{dt}, \tag{206}$$

$$p = \mu \frac{dx}{dt}. \tag{207}$$

$$\tag{208}$$

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 $\mu$ 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 $\Phi(X)\phi(x)$:

$$\left( \frac{\hat{P}^2}{2M} + \frac{\hat{p}^2}{2\mu} + V(\hat{x}) \right) \Phi(X)\phi(x) = E\Phi(X)\phi(x) \tag{209}$$

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

$$\phi(x)\frac{\hat{P}^2}{2M}\Phi(X) + \Phi(X)\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E\Phi(X)\phi(x) \tag{210}$$

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

$$\frac{1}{\Phi(X)}\frac{\hat{P}^2}{2M}\Phi(X) + \frac{1}{\phi(x)}\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E \tag{211}$$

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, 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:

$$\frac{\hat{P}^2}{2M}\Phi(X) = E_M\Phi(X) \tag{212}$$

$$\left(\frac{\hat{p}^2}{2\mu} + V(\hat{x})\right)\phi(x) = E_\mu\phi(x) \tag{213}$$

where we have

$$E_M + E_\mu = E \tag{214}$$

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 $\mu$ 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.

### 5.6.3 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 $\mu$ 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 $x_e$):

$$V(x) = V(x_e) + \left(\frac{dV(x)}{dx}\right)_{x_e}(x - x_e) + \frac{1}{2}\left(\frac{d^2V(x)}{dx^2}\right)_{x_e}(x - x_e)^2 + \cdots \tag{215}$$

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

$$\left(-\frac{\hbar^2}{2\mu}\frac{d^2}{dy^2} + \frac{1}{2}ky^2\right)\phi_n(y) = E_n\phi_n(y) \tag{216}$$

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 6. 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.

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, and you will see that the predictions made by the harmonic oscillator model are accurate again to a few percent for low vibrational levels, and become worse (as you would expect) for higher levels.

# 6 Transitions

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

$$|\psi(t=0)\rangle = \sum_n c_n|\phi_n\rangle, \tag{217}$$

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

$$|\psi(t)\rangle = \sum_n c_n e^{-\frac{-iE_n t}{\hbar}}|\phi_n\rangle \tag{218}$$

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

$$|\psi(t)\rangle = \sum_n c_n e^{-i\omega_n t}|\phi_n\rangle. \tag{219}$$

If the sum has only one term—that is, if $|\psi(t=0)\rangle$ 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.

## 6.1 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 $|\phi_k\rangle$ 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. (219), where now the $|\phi_n\rangle$ 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, it 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 overall process looks like this: we have two different Hamiltonians, $\hat{H}_a$ and $\hat{H}_b$, that have different sets of eigenfunctions and eigenvalues:

$$\hat{H}_a|\phi_n^a\rangle = E_n^a|\phi_n^a\rangle \tag{220}$$

$$\hat{H}_b|\phi_n^b\rangle = E_n^b|\phi_n^b\rangle, \tag{221}$$

where the functions labeled with $a$ and with $b$ are different functions. The system starts out at time $t = 0$ in an eigenfunction of $\hat{H}_a$, and under the influence of $\hat{H}_a$ it evolves simply:

$$|\psi(t = 0)\rangle = |\phi_k^a\rangle \tag{222}$$

$$|\psi(t = t_1)\rangle = e^{i\omega_k^a t_1}|\phi_k^a\rangle \tag{223}$$

Then, at time $t_1$, the Hamiltonian changes to $\hat{H}_b$. The wavefunction does not change at that instant, but we need to write it in terms of the eigenfunctions of $\hat{H}_b$:

$$|\psi(t = t_1)\rangle = \sum_n c_n|\phi_n^b\rangle, \tag{224}$$

where

$$c_n = e^{i\omega_k^a t_1}\langle\phi_n^b|\phi_k^a\rangle. \tag{225}$$

Several of the $c_n$ are likely to be nonzero. The wavefunction will then evolve under the influence of $\hat{H}_b$:

$$|\psi(t = t_2)\rangle = \sum_n c_n e^{-i\omega_n^b(t_2-t_1)}|\phi_n^b\rangle. \tag{226}$$

Then, at time $t_2$, we switch the Hamiltonian back to $\hat{H}_a$. Again, the wavefunction does not change, but now we need to write it in terms of the original eigenfunctions, the $|\phi_m^a\rangle$:

$$|\psi(t = t_2)\rangle = \sum_m d_m|\phi_m^a\rangle, \tag{227}$$

where

$$d_m = \sum_n c_n e^{-i\omega_n^b(t_2-t_1)} \langle \phi_m^a | \phi_n^b \rangle \tag{228}$$

Now, if we measure the total energy, we have probability $|d_m|^2$ of finding energy $E_m^a$. We say that the system has probability $|d_m|^2$ of making a *transition* from state $|\phi_k^a\rangle$ to state $|\phi_m^a\rangle$.

The transition probabilities are often 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.

### 6.1.1 Time dependent perturbation theory

There is a theory of great usefulness 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 study helium, you can choose $\hat{H}_0$ to be the kinetic energies of both electrons plus the Coulomb attractions of both electrons for the nucleus. That problem looks just like two hydrogen-atom problems, and the H-atom problem has already been solved. Then, you use for $\hat{H}_1$ the Coulomb repulsion of the two electrons for each other. Perturbation theory lets you calculate the changes in the energy eigenvalues and eigenfunctions that result from the interelectron repulsion (which is the piece that makes the He problem hard), and with very simple formulas gets you within 5% of the correct answer. (With slightly more complicated formulas it gets you within 0.2% of the right answer!)

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 $E = E_0 \cos \omega t$, where $\omega$ is the 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 $\langle\phi_m|\hat{H}_1|\phi_k\rangle$.

## 6.2 Selection rules

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:

$$\mu(x) = \mu(0) + \mu'(0)x + \frac{\mu''(0)}{2}x^2 + \cdots \tag{229}$$

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 $\mu$ is proportional to $\mathbf{E} \cdot \mu$. Therefore,

$$\hat{H}_1 \propto E_x(\mu(0) + \mu'(0)\hat{x}) \tag{230}$$

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

$$\hat{H}_1 \propto a + b\hat{x} = a + b'(\hat{a}^\dagger + \hat{a}). \tag{231}$$

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

$$\langle\phi_m|a + b'(\hat{a}^\dagger + \hat{a})|\phi_k\rangle \neq 0. \tag{232}$$

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 studied. 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 $\mu$m.

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.

# 7 Particle in a Box

One of the most famous, and most often taught, quantum mechanical systems is the “particle in a box”: 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

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) \tag{233}$$

with

$$V(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq L, \\ \infty & \text{otherwise.} \end{cases} \tag{234}$$

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

## 7.1 Eigenfunctions of $\hat{H}$

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

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\phi(x) = E\phi(x) \tag{235}$$

Outside the box the potential is infinite. Any finite $\phi(x)$ 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

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right]\phi(x) = E\phi(x) \tag{236}$$

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 requirements.

We could write the solution of Eq. (236) in terms of either exponentials or trig functions; the latter course is a little more direct and I will adopt it here. A solution of Eq. (236) is

$$\phi(x) = A\cos ax + B\sin ax, \tag{237}$$

because the second derivative of $\phi(x)$ is

$$\frac{d^2}{dx^2}\phi(x) = -a^2(A\cos ax + B\sin ax). \tag{238}$$

$A$, $B$, 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. (237) and requiring $\phi(0) = 0$ gives immediately $A = 0$ since $\cos(0) \neq 0$. We therefore have

$$\phi(x) = B\sin ax. \tag{239}$$

$B$ 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 $\pi$. So,

$$aL = n\pi \tag{240}$$

which gives

$$a = \frac{n\pi}{L}, \tag{241}$$

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

$$\phi_n(x) = B\sin\frac{n\pi x}{L}. \tag{242}$$

How do we find $B$? We normalize!

$$1 = \int_0^L |\phi(x)|^2\,dx \tag{243}$$

$$= \int_0^L B^2\sin^2\frac{n\pi x}{L}\,dx \tag{244}$$

$$= B^2\left[\frac{x}{2} - \frac{1}{4}\frac{L}{n\pi}\sin\frac{2n\pi x}{L}\right]_0^L \tag{245}$$

$$= B^2\frac{L}{2} \tag{246}$$

so

$$B = \left(\frac{2}{L}\right)^{\frac{1}{2}}. \tag{247}$$

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

$$\phi_n(x) = \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\frac{n\pi x}{L} \tag{248}$$

with $n$ a positive integer. The first few of these functions are shown in Figure 7 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.

## 7.2 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. (248) into the eigenvalue equation Eq. (236) and applying the second derivative formula Eq. (238) gives us

$$-\frac{\hbar^2}{2m}\left[-\left(\frac{n\pi}{L}\right)^2\right]\sin\frac{n\pi x}{L} = E_n \sin\frac{n\pi x}{L} \tag{249}$$

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2} \tag{250}$$

$$= \frac{h^2 n^2 \pi^2}{(2\pi)^2 2mL^2} \tag{251}$$

$$= \frac{n^2 h^2}{8mL^2} \tag{252}$$

So, the energies go up *quadratically* with $n$. 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.

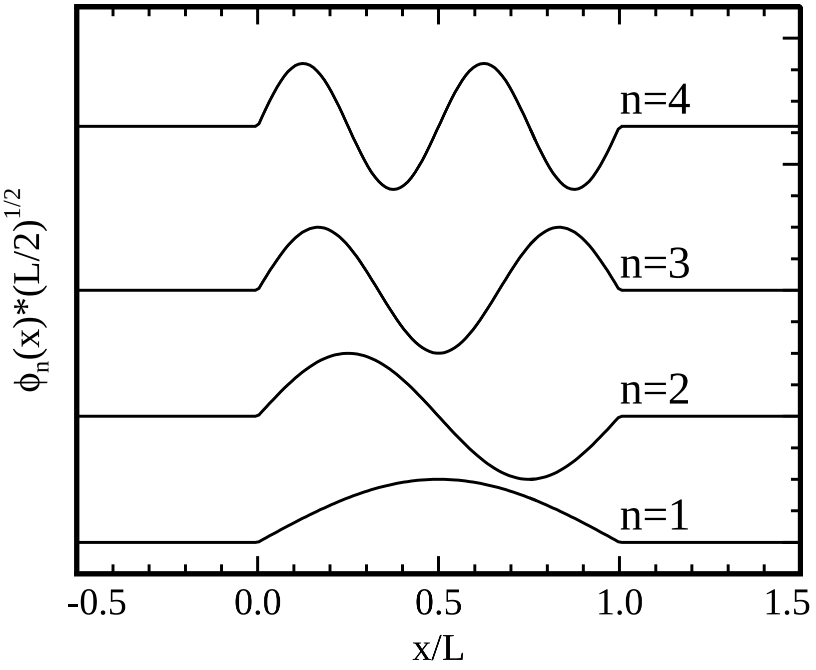


Figure 7: Stationary state wavefunctions for the particle in a one-dimensional box. The functions are offset vertically for clarity; all of them have value zero outside the range $0 \leq \frac{x}{L} \leq 1$. The vertical offsets do *not* scale with the energy differences between the functions. (In this potential, the energies go up quadratically with the quantum number $n$.)

# 8 Angular Momentum

(Much of the substance of this section is due to Prof. Sherwin Singer of The Ohio State University.) In classical mechanics, as soon as we consider more than one-dimensional problems (particles constrained to move along lines), the concept of *angular momentum* appears. Angular momentum is the "moment of momentum"; that is,

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{253}$$

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 (in three-dimensional, ordinary space) components of the angular momentum vector, and moving to the quantum problem just requires the placements of hats on all the variables:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \tag{254}$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \tag{255}$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{256}$$

Note how each of the above lines follows from the previous by the replacements, $x \to y$, $y \to z$, $z \to x$. This is referred to as 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.

## 8.1 Approach to the problem

In the harmonic oscillator problem, I showed you the most important parts of an algebraic derivation of the eigenvalues and eigenfunctions, using kets, raising and lowering operators, commutators, and so on. I also hinted at a way to find the eigenvalues and eigenfunctions by solving differential equations, finding directly the eigenfunctions $\phi(x)$ as functions of $x$. For the one-dimensional harmonic oscillator, those two approaches are exactly equivalent; all the same physical conclusions are reached with one as the other.

In the particle in a box problem, I only showed you the differential equation approach, which is particularly easy in that case. In fact I have not seen a pure-algebraic solution of that problem. (I have made a start on one myself, but it's not clear that I'll be able to finish it.)

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 does find 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.

## 8.2 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 $\hat{L}_y$.

$$\left[\hat{L}_x, \hat{L}_y\right] = \left[\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right] \tag{257}$$

Writing the commutator out and collecting terms gives

$$= \left[\hat{y}\hat{p}_z, \hat{z}\hat{p}_x\right] + \left[\hat{z}\hat{p}_y, \hat{x}\hat{p}_z\right] - \left[\hat{y}\hat{p}_z, \hat{x}\hat{p}_z\right] - \left[\hat{z}\hat{p}_y, \hat{z}\hat{p}_x\right] \tag{258}$$

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

$$= \left[\hat{y}\hat{p}_z, \hat{z}\hat{p}_x\right] + \left[\hat{z}\hat{p}_y, \hat{x}\hat{p}_z\right] \tag{259}$$

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

$$= \hat{y}\hat{p}_x\left[\hat{p}_z, \hat{z}\right] + \hat{x}\hat{p}_y\left[\hat{z}, \hat{p}_z\right] \tag{260}$$

$$= \hat{y}\hat{p}_x\left(-i\hbar\right) + \hat{x}\hat{p}_y\left(i\hbar\right) \tag{261}$$

$$= i\hbar\hat{L}_z \tag{262}$$

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

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar \hat{L}_z \tag{263}$$

$$\left[\hat{L}_y, \hat{L}_z\right] = i\hbar \hat{L}_x \tag{264}$$

$$\left[\hat{L}_z, \hat{L}_x\right] = i\hbar \hat{L}_y \tag{265}$$

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, eigenstates that are simultaneously eigenstates of two different components of $\hat{\mathbf{L}}$, say, $\hat{L}_x$ and $\hat{L}_y$, do not exist.

## 8.3 Commutators with $\hat{\mathbf{L}}^2$

$\hat{\mathbf{L}}$ is an 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.

$$\begin{aligned}
\left[\hat{\mathbf{L}}^2, \hat{L}_x\right] &= \left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right)\hat{L}_x - \hat{L}_x\left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right) && (266)\\
&= \left(\hat{L}_y^2 + \hat{L}_z^2\right)\hat{L}_x - \hat{L}_x\left(\hat{L}_y^2 + \hat{L}_z^2\right) && (267)\\
&= \hat{L}_y\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\hat{L}_z && (268)\\
&= \hat{L}_y(\hat{L}_y\hat{L}_x) - (\hat{L}_x\hat{L}_y)\hat{L}_y + \hat{L}_z(\hat{L}_z\hat{L}_x) - (\hat{L}_x\hat{L}_z)\hat{L}_z && (269)
\end{aligned}$$

Eq. (269) differs from Eq. (268) 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.

$$\begin{aligned}
\left[\hat{\mathbf{L}}^2, \hat{L}_x\right] &= \hat{L}_y(\hat{L}_x\hat{L}_y - i\hbar\hat{L}_z) - (\hat{L}_y\hat{L}_x + i\hbar\hat{L}_z)\hat{L}_y\\
&\quad + \hat{L}_z(\hat{L}_x\hat{L}_z + i\hbar\hat{L}_y) - (\hat{L}_z\hat{L}_x - i\hbar\hat{L}_y)\hat{L}_z && (270)\\
&= 0
\end{aligned}$$

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

$$\left[\hat{\mathbf{L}}^2, \hat{L}_y\right] = \left[\hat{\mathbf{L}}^2, \hat{L}_z\right] = 0 \tag{271}$$

From the fact that $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do not commute with each other [Eqs.(263-265)] but they all separately commute with $\hat{\mathbf{L}}$, 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 by $\hat{\mathbf{L}}^2$, but this is only convention.

## 8.4 Eigenvalues of angular momentum

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

$$\hat{\mathbf{L}}^2|c\,b\rangle = c|c\,b\rangle \tag{272}$$

$$\hat{L}_z|c\,b\rangle = b|c\,b\rangle. \tag{273}$$

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,

$$\left\langle \hat{\mathbf{L}}^2 \right\rangle \geq \left\langle \hat{L}_z \right\rangle^2 , \tag{274}$$

and therefore

$$c \geq 0 \text{ and } c \geq b^2 . \tag{275}$$

### 8.4.1 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,

$$\hat{L}_+ \equiv \hat{L}_x + i\hat{L}_y. \tag{276}$$

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

$$\hat{L}_z \left(\hat{L}_+|c\,b\rangle\right) = \hat{L}_z \left(\hat{L}_x + i\hat{L}_y\right)|c\,b\rangle \tag{277}$$

Now apply commutation relation (265) in the form $\hat{L}_z\hat{L}_x = i\hbar\hat{L}_y + \hat{L}_x\hat{L}_z$, and commutation relation (264) in the form $\hat{L}_z\hat{L}_y = -i\hbar\hat{L}_x + \hat{L}_y\hat{L}_z$:

$$= \left(i\hbar\hat{L}_y + \hat{L}_x\hat{L}_z\right)|c\,b\rangle + i\left(-i\hbar\hat{L}_x + \hat{L}_y\hat{L}_z\right)|c\,b\rangle \tag{278}$$

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

$$= \hbar\left(\hat{L}_x + i\hat{L}_y\right)|c\,b\rangle + \left(\hat{L}_x + i\hat{L}_y\right)b|c\,b\rangle \tag{279}$$

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

$$= (b + \hbar)\left(\hat{L}_+|c\,b\rangle\right) \tag{280}$$

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

$$\hat{L}_- \equiv \hat{L}_x - i\hat{L}_y\,. \tag{281}$$

By very similar manipulations to Eqs.(277-280), you can show that, if $\hat{L}_z|c\,b\rangle = b|c\,b\rangle$, then

$$\hat{L}_z\left(\hat{L}_-|c\,b\rangle\right) = (b - \hbar)\left(\hat{L}_-|c\,b\rangle\right)\ , \tag{282}$$

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

Since

$$\left[\hat{\mathbf{L}}^2, \hat{L}_\pm\right] = \left[\hat{\mathbf{L}}^2, \hat{L}_x\right] \pm i\left[\hat{\mathbf{L}}^2, \hat{L}_y\right] = 0\,, \tag{283}$$

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.)

### 8.4.2 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 $b_{\mathrm{min}}$ and a $b_{\mathrm{max}}$. In the case of the maximum,

$$\hat{L}_+|c\,b_{\mathrm{max}}\rangle = 0. \tag{284}$$

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

$$\hat{L}_-\hat{L}_+|c\,b_{\mathrm{max}}\rangle = 0. \tag{285}$$

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

$$\begin{aligned} 0 = \hat{L}_-\hat{L}_+|c\, b_{\max}\rangle &= \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right)|c\, b_{\max}\rangle && (286)\\ &= \left\{\hat{L}_x^2 + \hat{L}_y^2 + i\left(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x\right)\right\}|c\, b_{\max}\rangle && (287)\\ &= \left\{\hat{L}_x^2 + \hat{L}_y^2 - \hbar\hat{L}_z\right\}|c\, b_{\max}\rangle && (288)\\ &= \left\{\hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z\right\}|c\, b_{\max}\rangle && (289)\\ &= \left\{c - b_{\max}^2 - \hbar b_{\max}\right\}|c\, b_{\max}\rangle && (290)\end{aligned}$$

Therefore

$$c = b_{\max}^2 + \hbar b_{\max}\,. \qquad (291)$$

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

$$c = b_{\min}^2 - \hbar b_{\min}\,. \qquad (292)$$

Subtract Eq.(292) from Eq.(291).

$$b_{\max}^2 + \hbar b_{\max} - b_{\min}^2 + \hbar b_{\min} = 0 \qquad (293)$$

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:

$$b_{\max} = -b_{\min}\,. \qquad (294)$$

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 $\hat{L}_+$. Lets say, for a given value of $c$, it takes a total of $n$ applications of $\hat{L}_+$ to turn $|c\, b_{\min}\rangle$ into $|c\, b_{\max}\rangle$.

$$(\hat{L}_+)^n|c\, b_{\min}\rangle \propto |c\, b_{\max}\rangle \qquad (295)$$

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

$$b_{\max} - b_{\min} = n\hbar, \quad n = 0, 1, 2, \ldots, \qquad (296)$$

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

$$\begin{aligned} b_{\max} &= \frac{1}{2}n\hbar, \quad n = 0, 1, 2, \ldots && (297)\\ &\equiv j\hbar, \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots, && (298)\end{aligned}$$

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

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

$$b_{\min} = -j\hbar,\ j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{299}$$

Finally, upon substitution of $b_{\max} = j\hbar$ into Eq.(291), we obtain

$$c = b_{\max}^2 + \hbar b_{\max} = j^2\hbar^2 + j\hbar^2 = j(j+1)\hbar^2 \tag{300}$$

### 8.4.3 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 $|j\,m\rangle$ with the properties

$$\hat{\mathbf{L}}^2|j\,m\rangle = \hbar^2 j(j+1)|j\,m\rangle \tag{301}$$

$$\hat{L}_z|j\,m\rangle = \hbar m|j\,m\rangle \tag{302}$$

where

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{303}$$

$$m = -j, -j+1, \ldots j. \tag{304}$$

The raising and lowering operators are

$$\hat{L}_+ \equiv \hat{L}_x + i\hat{L}_y \tag{305}$$

$$\hat{L}_- \equiv \hat{L}_x - i\hat{L}_y \tag{306}$$

and, inverting those,

$$\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-) \tag{307}$$

$$\hat{L}_y = \frac{1}{2i}(\hat{L}_+ - \hat{L}_-) \tag{308}$$

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

$$\hat{L}_+|j\,m\rangle = \hbar\sqrt{j(j+1) - m(m+1)}|j\,m+1\rangle \tag{309}$$

$$\hat{L}_-|j\,m\rangle = \hbar\sqrt{j(j+1) - m(m-1)}|j\,m-1\rangle \tag{310}$$

**Table 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 | $\hat{\mathbf{I}}$ | $I$, $m_I$ |

## 8.5 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 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.

## 8.6 Properties of Rotational Angular Momentum States

When we applied the harmonic oscillator model to molecular vibrations, you 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.

Converting differential operators into spherical coordinates is a straightforward but tedious exercise; I'll just give the results here. The $\hat{\mathbf{L}}^2$ operator becomes

$$\hat{\mathbf{L}}^2 = -\hbar^2 \left( \frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) \tag{311}$$

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 $\theta$ and $\phi$, the normal spherical polar angles. The polar radius $r$ does not appear.

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

$$\hat{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}. \tag{312}$$

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

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

$$\begin{aligned}\hat{\mathbf{L}}^2 Y_{lm}(\theta,\phi) &= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) Y_{lm}(\theta,\phi) \\ &= (\text{eigenvalue of } \hat{\mathbf{L}}^2)\ Y_{lm}(\theta,\phi) \qquad (313) \\ \hat{L}_z Y_{lm}(\theta,\phi) &= \frac{\hbar}{i}\frac{\partial}{\partial\phi} Y_{lm}(\theta,\phi) \\ &= (\text{eigenvalue of } \hat{L}_z)\ Y_{lm}(\theta,\phi) \qquad (314)\end{aligned}$$

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

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta\ |Y_{lm}(\theta,\phi)|^2 = 1\ , \tag{315}$$

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

$$Y_{lm}(\theta,\phi+2\pi) = Y_{lm}(\theta,\phi)\ . \tag{316}$$

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.

### 8.6.1 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 $\phi$ with a real function of $\theta$. 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. 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 $-l \leq m \leq l$, there is only one $s$-state (with $m = 0$).

$$Y_{00}(\theta, \phi) = \sqrt{\frac{1}{4\pi}} \tag{317}$$

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.

Since $m = -1, 0, 1$ for $l = 1$, there are three $p$-states.

$$Y_{11}(\theta, \phi) = -\sqrt{\frac{3}{4\pi}}\sqrt{\frac{1}{2}} \sin\theta \, e^{i\phi} \tag{318}$$

$$Y_{10}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos\theta \tag{319}$$

$$Y_{1,-1}(\theta, \phi) = \sqrt{\frac{3}{4\pi}}\sqrt{\frac{1}{2}} \sin\theta \, e^{-i\phi} \tag{320}$$

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 oscillating and complex elsewhere. You are used to seeing linear combinations of the $m = \pm 1$ states which are real:

$$p_x = \sqrt{\frac{1}{2}}\left[-Y_{11} + Y_{1,-1}\right] = \sqrt{\frac{3}{4\pi}} \sin\theta \ \cos\phi = \sqrt{\frac{3}{4\pi}}\frac{x}{r} \tag{321}$$

$$p_y = \sqrt{\frac{1}{2}}\left[-Y_{11} - Y_{1,-1}\right] = \sqrt{\frac{3}{4\pi}} \sin\theta \ \sin\phi = \sqrt{\frac{3}{4\pi}}\frac{y}{r} \tag{322}$$

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

$$p_z = Y_{10} = \sqrt{\frac{3}{4\pi}} \cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r}. \tag{323}$$

You can verify that the $p_x$ and $p_y$ states are still eigenfunctions of $\hat{\mathbf{L}}^2$, but are not eigenfunctions of $\hat{L}_z$. 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 $z = 0$ and you are looking in the $xy$ plane.

The five $d$-states are

$$Y_{20}(\theta,\phi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta - 1) \tag{324}$$

$$Y_{2\pm1}(\theta,\phi) = \frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta\ \cos\theta\ e^{\pm i\phi} \tag{325}$$

$$Y_{2\pm2}(\theta,\phi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\ e^{\pm 2i\phi} \tag{326}$$

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 $d_{xy}$.

### 8.6.2 Orthonormality of the $Y_{lm}$

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

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta\, Y^*_{l'm'}(\theta,\phi) Y_{lm}(\theta,\phi) = \begin{cases} 1 & \text{if } l = l' \text{ and } m = m', \\ 0 & \text{otherwise.} \end{cases} \tag{327}$$

# 9 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:

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(r) \tag{328}$$

where

$$\hat{\mathbf{p}}^2 = -\hbar^2 \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right). \tag{329}$$

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

$$\hat{H} = \left[ -\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} \right) + \frac{1}{2\mu r^2}\hat{\mathbf{L}}^2 + V(r) \right], \tag{330}$$

where

$$\hat{\mathbf{L}}^2 = -\hbar^2 \left( \frac{\partial^2}{\partial \theta^2} + \cot\theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial \phi^2} \right) \tag{331}$$

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

$$\psi(r,\theta,\phi) = R(r)Y_{lm}(\theta,\phi). \tag{332}$$

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

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{1}{2\mu r^2}\hat{\mathbf{L}}^2+V(r)\right]R(r)Y_{lm}(\theta,\phi) = ER(r)Y_{lm}(\theta,\phi) \tag{333}$$

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

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{\hbar^2 l(l+1)}{2\mu r^2}+V(r)\right]R(r)Y_{lm}(\theta,\phi) = ER(r)Y_{lm}(\theta,\phi) \tag{334}$$

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

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{\hbar^2 l(l+1)}{2\mu r^2}+V(r)\right]R_{nl}(r) = ER_{nl}(r) \tag{335}$$

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. (335) so that we must find a different $R(r)$ for every $l$.

This "radial equation" appears in every two-body potential problem. With $V(r) = \frac{1}{2}kr^2$, 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.

## 9.1 The rigid rotor

What about a simple limiting case: we have a system with $r$ *fixed.* (This amounts to a V(r) that is very large except right at a particular value, where it is zero; call that value $r_e$.) If $r$ is no longer a variable but simply a constant parameter, then the first term in Eq. (330) disappears; that term is the kinetic energy corresponding to the radial motion (increases and decreases in the interparticle distance) and since $r$ is fixed that kinetic energy is zero. The third, $V(r)$, term simply becomes a constant, which we can absorb into the eigenvalue $E$. The Schrödinger equation becomes

$$\frac{1}{2\mu r_e^2}\hat{\mathbf{L}}^2\psi(\theta,\phi) = E\psi(\theta,\phi) \tag{336}$$

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 eigenfunctions of two particles connected 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 $\theta$ and $\phi$ describe the orientation of the molecule in space.

What are the eigenvalues? We have

$$\hat{H}Y_{lm}(\theta,\phi) = \frac{1}{2\mu r_e^2}\hat{\mathbf{L}}^2 Y_{lm}(\theta,\phi) \tag{337}$$

$$= \frac{\hbar^2}{2\mu r_e^2} l(l+1) Y_{lm}(\theta,\phi) \tag{338}$$

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

$$B = \frac{\hbar^2}{2\mu r_e^2}, \tag{339}$$

the *rotational constant*. The energies of the rotating molecule are then

$$E_J = BJ(J+1) \tag{340}$$

Typically $B$ is 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.

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.

### 9.1.1 Selection rules for the rigid rotor

A heteronuclear diatomic molecule (modeled for the moment as a rigid rotor) can have a permanent dipole moment $\mu_0$, 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:

$$\hat{H}_1 = \mathbf{E} \cdot \mu \tag{341}$$

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 $\theta$ and $\phi$, we need the transformation between spherical polar and Cartesian coordinates:

$$x = r \sin\theta \cos\phi \tag{342}$$

$$y = r \sin\theta \sin\phi \tag{343}$$

$$z = r \cos\theta \tag{344}$$

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

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

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta \, d\theta Y_{j'm'}(\theta, \phi) E_z \cos\theta Y_{jm}(\theta, \phi) \neq 0 \tag{345}$$

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

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

$$\cos\theta Y_{jm}(\theta,\phi) = f(j,m)Y_{j+1\,m}(\theta,\phi) + g(j,m)Y_{j-1\,m}(\theta,\phi) \tag{346}$$

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

$$\Delta j = \pm 1 \tag{347}$$

$$\Delta m = 0, \pm 1 \tag{348}$$

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 $\text{cm}^{-1}$ for large molecules. The energy difference between two rotational states of a molecule with quantum numbers $J$ and $J+1$ is (in $\text{cm}^{-1}$)

$$\Delta E = B(J+1)(J+2) - B(J)(J+1) = 2B(J+1) \tag{349}$$

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 $1.2\times 10^{12}$ Hz. This light is in the microwave frequency region. (Microwave ovens work at $2.4\times 10^{12}$ Hz; they excite rotational transitions in water molecules, which have a large $B$ because of the light hydrogen atoms.)

## 9.2 Vibrational motion in 3D problems

I claimed that the radial equation, Eq. (335), 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 us consider only the rotationless molecule: $l = 0$. That removes the middle term, the "centrifugal potential" (which really corresponds to rotational kinetic energy). Next, search for radial wavefunctions $R(r)$ of the form

$$R_{nl}(r) = \frac{1}{r} u_{nl}(r) \tag{350}$$

In your homework, you will show that substituting that form of $R(r)$ into the radial equation gives a simple, one-dimensional equation for $u(r)$ that looks exactly like the harmonic oscillator problem if $V(r) = \frac{1}{2}kr^2$. 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.

Actually, this trick of looking for $u_{nl}(r)$ instead of $R_{nl}(r)$ works for all kinds of radial potentials. Let us turn next to the hydrogen atom and use it there.

# 10 Hydrogen atom

Eq. (335) describes hydrogenlike atoms if we use

$$V(r) = \frac{1}{4\pi\epsilon_0} \frac{q_e q_N}{r} \tag{351}$$

$$= \frac{-Ze^2}{4\pi\epsilon_0} \frac{1}{r}, \tag{352}$$

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 ($1.6 \times 10^{-19}$ Coulombs), and $\epsilon_0$ is the permittivity of free space; $4\pi\epsilon_0 = 1.11265 \times 10^{-10}\ \mathrm{C}^2\,\mathrm{J}^{-1}\,\mathrm{m}^{-1}$. (A hydrogenlike atom is any atom with only a single electron: H, $\mathrm{He}^+$, $\mathrm{Li}^{2+}$, etc.)

Making the substitution $R_{nl}(r) = \frac{1}{r} u_{nl}(r)$, and plugging that into Eq. (335), we find after a little work (which you do in your homework)

$$\left[ -\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \right] u_{nl}(r) = E_{nl} u_{nl}(r) \tag{353}$$

This looks like a one-dimensional Schrödinger equation for a particle of mass $\mu$ in a potential

$$V_{\text{eff}}(r) = \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \tag{354}$$

If $l \neq 0$, the first 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 second 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.

Eq. (353) can be solved by the standard method of power series expansion, taught in every course on differential equations. You may recall that in the harmonic oscillator case, we found solutions that looked like a polynomial in $x$ times $e^{-ax^2}$, and the polynomials were called Hermite polynomials. The H-atom case is similar; the solutions are polynomials in $r$ times $e^{-ar}$. (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".

(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 $\theta$ part of the $Y_{lm}$ are Legendre polynomials.)

The H-atom stationary states are simpler to write down if we define the quantity

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} \tag{355}$$

$a_0$ is called the Bohr radius; it is the average distance from the nucleus of an electron in the ground state of a hydrogen atom, and has the value 0.05291772083 nm (that is, just over half an Å).

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

Figure 13.5 in Atkins gives plots of the radial functions $R_{nl}(r)$ for $1 \leq n \leq 3$ and all the permissible values of $l$. 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 1$s$ radial function has no nodes, the 2$s$ has one node, the 2$p$ has no nodes, and so on.

## 10.1 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

$$E_n = -\frac{1}{n^2}E_I = -\frac{1}{n^2}\frac{\mu Z^2 e^4}{32\pi\epsilon_0^2\hbar^2} \tag{356}$$

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 hydrogen-like 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).

## 10.2 Introduction to atomic units

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

$$E_h = \frac{m_e e^4}{16\pi^2\epsilon_0^2\hbar^2} = 4.35974381 \times 10^{-18}\,\mathrm{J} \approx 27.20\,\mathrm{eV} \tag{357}$$

The energy of the ground state of hydrogen is $E_0 \approx -\frac{1}{2}E_h$. (It is not exactly half a hartree because the reduced mass of the proton-electron system is not exactly the mass 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, $a_0 = 0.5291772083 \times 10^{-10}$ m, which is the atomic unit of length. 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.

## 10.3 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)

$$\langle\phi_f|\hat{H}_1|\phi_i\rangle = \int d\text{\textbf{ø}}\phi_{n'l'm'}(r,\theta,\varphi)r\cos\theta\phi_{nlm}(r,\theta,\varphi) \tag{358}$$

$$= \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\infty} r^2\, dr R_{n'l'}(r) Y_{l'm'}(\theta,\varphi) r\cos\theta R_{nl}(r) Y_{lm}(\theta,\varphi) \tag{359}$$

$$= \left[\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta Y_{l'm'}(\theta,\varphi)\cos\theta(r) Y_{lm}(\theta,\varphi)\right] \times \left[\int_0^{\infty} r^2\, dr R_{n'l'}(r) r R_{nl}(r)\right] \tag{360}$$

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 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 $2s \rightarrow 3d$.

# 11 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

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Ze^2}{4\pi\epsilon_0|\mathbf{r}_1|} - \frac{Ze^2}{4\pi\epsilon_0|\mathbf{r}_2|} + \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_2 - \mathbf{r}_2|} \tag{361}$$

where $\nabla_1^2$ 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. $|\mathbf{r}_2 - \mathbf{r}_2|$, 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)

$$\begin{aligned}\left[-\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}\right]\psi(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2)&\\ = E\psi(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2)&\end{aligned} \tag{362}$$

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.)

## 11.1 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

$$\hat{H}_0 = \left(-\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1}\right) + \left(-\frac{1}{2}\nabla_2^2 - \frac{Z}{r_2}\right) \tag{363}$$

This Hamiltonian looks like 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

$$\begin{aligned} \psi^{(0)}(r_1,\theta_1,\phi_1,r_2,\theta_2,\phi_2) &= \phi_{nlm}(r_1,\theta_1,\phi_1)\phi_{n'l'm'}(r_2,\theta_2,\phi_2) \qquad (364)\\ &= \phi_{nlm}(1)\phi_{n'l'm'}(2) \qquad (365) \end{aligned}$$

where the $\phi$ functions are the ordinary eigenfunctions of the H atom Hamiltonian, and the second line shows a commone 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 $\phi_{1s}$ 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.

Now what do we do about the part of the Hamiltonian that we left out? 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.

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:

$$E^{(1)} = \langle\phi_{1s}(1)\phi_{1s}(2)|\hat{H}_1|\phi_{1s}(1)\phi_{1s}(2)\rangle \qquad (366)$$

This integral is a little complicated to do, because $r_{12}$ is messy when written in spherical coordinates. But you can do it, and the result is

$$E^{(1)} = \frac{5}{8}Z \qquad (367)$$

so the estimated total energy of the ground state of He is (because $Z = 2$) $-4 + \frac{5}{4} = -2.75$ hartree. The exact (experimentally measured) energy is $-2.9033$ hartree; that's 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.

## 11.2 The variational theorem

We have some Hamiltonian $H$ for our problem, and would like to find some *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,

$$\langle E\rangle = \frac{\langle\psi_0|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} = \frac{\int \psi_0^* \, H \, \psi_0 \, d\mathbf{r}}{\int \psi_0^* \, \psi_0 \, d\mathbf{r}}, \tag{368}$$

where $\psi_0$ 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 $\psi_0$ is already normalized, it it not needed.

The proof of the theorem is as follows. We know that the trial function $\psi_0$ 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!)

$$\psi_0 = \sum_{n=0}^{\infty} c_n|\varphi_n\rangle \tag{369}$$

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

$$\begin{aligned}\langle\psi_0|\psi_0\rangle &= \left[\sum_{m=0}^{\infty} c_m^*\langle\varphi_m|\right]\left[\sum_{n=0}^{\infty} c_n|\varphi_n\rangle\right] && (370)\\ &= \sum_m\sum_n\langle\varphi_m|c_m^* c_n|\varphi_n\rangle && (371)\end{aligned}$$

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:

$$\begin{aligned}&= \sum_n\langle\varphi_n|c_n^* c_n|\varphi_n\rangle && (372)\\ &= \sum_n |c_n|^2 && (373)\end{aligned}$$

Now, calculate the numerator:

$$\begin{aligned}
\langle\psi_0|\hat{H}|\psi_0\rangle &= \left[\sum_m c_m^*\langle\varphi_m^*|\right]\ \hat{H}\ \left[\sum_n c_n|\varphi_n\rangle\right] && (374)\\
&= \sum_m\sum_n c_m^* c_n\langle\varphi_m|\hat{H}|\varphi_n\rangle && (375)\\
&= \sum_m\sum_n c_m^* c_n E_n\langle\varphi_m|\varphi_n\rangle && (376)\\
&= \sum_n E_n|c_n|^2 && (377)\\
&= \sum_n |c_n|^2\left(E_0+(E_n-E_0)\right) && (378)\\
&= \sum_n |c_n|^2 E_0+\sum_n |c_n|^2(E_n-E_0) && (379)\\
&= E_0\sum_n |c_n|^2+\sum_n |c_n|^2(E_n-E_0) && (380)
\end{aligned}$$

Now look at quotient

$$\begin{aligned}
\frac{\langle\psi_0|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} &= \frac{E_0\ \sum_n\ |c_n|^2+\sum_n|c_n|^2(E_n-E_0)}{\sum_n\ |c_n|^2} && (381)\\
&= E_0+\frac{\sum_n \overbrace{|c_n|^2}^{>0}\overbrace{(E_n-E_0)}^{\geq 0}}{\underbrace{\sum_n |c_n|^2}_{>0}} && (382)
\end{aligned}$$

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.

## 11.3 Variational calculation for He

The ground state of H (hydrogen) has the wavefunction

$$\psi_0 = Ce^{-Zr/a_0} \tag{383}$$

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 $Z$ in the wavefunction as an adjustable parameter, called $Z'$. It will represent some kind of effective nuclear charge seen by the electron. Our trial He function will be

$$\Psi_0 = c^2 e^{-Z' r_1 / a_0} e^{-Z' r_2 / a_0} \tag{384}$$

where $c$, the normalization constant, must contain $Z'$ rather than the true nuclear charge.

To do a variational calculation, we must evaluate the energy $E_{\mathrm{He}} = \langle \Psi_0 | \hat{H}_{\mathrm{He}} | \Psi_0 \rangle$, and minimize it with respect to $Z'$. In that expression $\Psi_0$ is the approximate wavefunction, with $Z = Z'$. $\hat{H}_{\mathrm{He}}$ is the true helium Hamiltonian, with $Z = 2$, and with the $\frac{1}{r_{12}}$ term included.

In homework, you'll show that

$$E_{\mathrm{He}} = (Z')^2 - 2ZZ' + \frac{5}{8} Z' \tag{385}$$

Then, minimizing with respect to $Z'$, we find

$$Z' = Z - \frac{5}{16}. \tag{386}$$

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?

$$E_{\mathrm{He}} = (Z')^2 - 2ZZ' + \frac{5}{8} Z' = -2.8477 E_h. \tag{387}$$

The exact energy is $-2.9033 E_h$, so the fractional error is

$$\frac{-2.9033 - (-2.8477)}{-2.9033} = 1.9\% \tag{388}$$

and the calculated energy is higher than the exact value, as expected.

### 11.3.1 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

$$\Psi_0 = N[c_1 e^{-\alpha r_1} + c_2 e^{-\beta r_1}][c_1 e^{-\alpha r_2} + c_2 e^{-\beta r_2}] \tag{389}$$

and use $c_1$, $c_2$, $\alpha$, and $\beta$ 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.

### 11.3.2 Linear variational calculations

Since our trial wavefunction 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:

$$\psi_0 = a_1 f_1(\mathbf{r}_1, \mathbf{r}_2, \cdots) + a_2 f_2(\mathbf{r}_1, \mathbf{r}_2, \cdots) + a_3 f_3(\mathbf{r}_1, \mathbf{r}_2, \cdots) + \ldots \tag{390}$$

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 (usually on computers) are available for solving such sets of equations.

The linear variation technique is described in section 14.7 of Atkins, in the context of molecular orbital theory. 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 homogenous set of algebraic equations that look like

$$a_1(H_{11} - ES_{11}) + a_2(H_{12} - ES_{12}) = 0 \tag{391}$$

$$a_1(H_{21} - ES_{21}) + a_2(H_{22} - ES_{22}) = 0 \tag{392}$$

(with more terms if there are more parameters.) In those equations

$$H_{11} = \langle f_1|\hat{H}|f_1\rangle \tag{393}$$

$$H_{12} = \langle f_1|\hat{H}|f_2\rangle \tag{394}$$

$$S_{11} = \langle f_1|f_1\rangle \tag{395}$$

$$S_{12} = \langle f_1|f_2\rangle \tag{396}$$

and $E$ is the expectation value of the energy.

There is a theorem from linear algebra that says that a nontrivial solution to Eq. (392) (one other than $a_1 = a_2 = 0$) exists only if the determinant of the coefficients vanishes:

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - ES_{22} \end{vmatrix} = 0 \tag{397}$$

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. (392) 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.

# 12 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

$$\psi = \psi\left(x_1, y_1, z_1, x_2, y_2, z_2\right) = \psi\left(\mathbf{r}_1, \mathbf{r}_2\right), \tag{398}$$

and similarly for larger number of particles.

The interpretation of $\psi$ is a natural extension of the one-particle case: $|\psi(\mathbf{r}_1,\mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2$ 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. $\mathrm{d}\mathbf{r}_1$ has units $\mathrm{m}^3$ (volume), as has $d\mathbf{r}_2$. So $\psi(\mathbf{r}_1,\mathbf{r}_2)$ must have units $\mathrm{m}^{-3}$. In general, for n particles, $\psi$ will have units $\mathrm{m}^{-3n/2}$.

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 is gets multiplied by some complex number, call it $\gamma$, with magnitude 1.

$$\psi(\mathbf{r}_2,\mathbf{r}_1) = \gamma\psi(\mathbf{r}_1,\mathbf{r}_2)\,. \tag{399}$$

Now, if we do the same trick again, we multiply by $\gamma$ again:

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \gamma\psi(\mathbf{r}_2,\mathbf{r}_1) = \gamma^2\psi(\mathbf{r}_1,\mathbf{r}_2) \tag{400}$$

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 multielectron wavefunction must change sign when the *labels* of any 2 electrons are exchanged. This is one statement of the *Pauli principle.*

## 12.1 Spin

In quantum mechanics, *spin* means an “intrinsic angular momentum” that is a property of most particles. The spin operators are usually written $s$ (or

$I$ 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 $|s = \frac{1}{2}, m_s = \frac{1}{2}\rangle$ and $|s = \frac{1}{2}, m_s = -\frac{1}{2}\rangle$, also known as

$$|\frac{1}{2}\,\frac{1}{2}\rangle \quad \text{and} \quad |\frac{1}{2}\,-\frac{1}{2}\rangle \tag{401}$$

$$\alpha \quad \text{and} \quad \beta, \tag{402}$$

$$\text{or ”spin up” and “spin down”.} \tag{403}$$

For the H atom, possible stationary state wavefunctions include

$$\psi = R_{nl}(r)Y_{lm}(\theta,\phi)\alpha \tag{404}$$

$$\psi = R_{nl}(r)Y_{lm}(\theta,\phi)\beta \tag{405}$$

$$\psi = R_{nl}(r)Y_{lm}(\theta,\phi)(\alpha+\beta) \tag{406}$$

Since the spin functions do not appear in the H-atom Hamiltonian anywhere, all three of these functions will have the same energy.

The $\alpha$ and $\beta$ “functions” (the angular momentum kets $|j\,m\rangle$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 $\alpha$ and $\beta$ as functions of that.

## 12.2 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 system 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 ± 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

$$\hat{P}_{12}[\alpha(1)\alpha(2)] = \alpha(2)\alpha(1), \tag{407}$$

which is okay; multiplication is commutative, so this is just what we started with.

$$\hat{P}_{12}[\alpha(1)\beta(2)] = \alpha(2)\beta(1) = \beta(1)\alpha(2) \tag{408}$$

which is not plus or minus itself!

$$\hat{P}_{12}[\beta(1)\alpha(2)] = \beta(2)\alpha(1) = \alpha(1)\beta(2) \tag{409}$$

which, again, is no good. Finally,

$$\hat{P}_{12}[\beta(1)\beta(2)] = \beta(2)\beta(1) = \beta(1)\beta(2) \tag{410}$$

which is okay.

Notice that the middle two candidates turned into each other under $\hat{P}_{12}$. That suggests a trick: Try

$$\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) + \beta(1)\alpha(2)) \tag{411}$$

or

$$\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \beta(1)\alpha(2)). \tag{412}$$

Let's check the first one:

$$\begin{aligned} \hat{P}_12[\alpha(1)\beta(2) + \beta(1)\alpha(2)] &= \alpha(2)\beta(1) + \beta(2)\alpha(1) && (413) \\ &= \alpha(1)\beta(2) + \beta(1)\alpha(2) && (414) \end{aligned}$$

which works (with $\gamma = +1$).

How about the second?

$$\begin{aligned} \hat{P}_12[\alpha(1)\beta(2) - \beta(1)\alpha(2)] &= \alpha(2)\beta(1) - \beta(2)\alpha(1) && (415) \\ &= (-1)(\alpha(1)\beta(2) - \beta(1)\alpha(2)) && (416) \end{aligned}$$

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.

Consider a trial wavefunction that is just a product of two H-atom-like functions, $Z = 2$, with one electron in the 1$s$ orbital and the other in the 2$s$. (This will not give particularly good numerical results but the qualitative behavior is correct.) The basic spatial functions therefore look like

$$\varphi_{1s}(r_1)\varphi_{2s}(r_2) \equiv 1s(1)2s(2) \tag{417}$$

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:

$$\psi_s \propto \varphi_{1s}(r_1)\varphi_{2s}(r_2) + \varphi_{2s}(r_1)\varphi_{1s}(r_2) \tag{418}$$
$$\psi_a \propto \varphi_{1s}(r_1)\varphi_{2s}(r_2) - \varphi_{2s}(r_1)\varphi_{1s}(r_2) \tag{419}$$

The first one is symmetric if the electron labels are exchanged; it must therefore be paired with an antisymmetric spin function. We have only one of those, the singlet. So

$$\Psi_s = \frac{1}{\sqrt{2}}\left(1s(1)2s(2) + 2s(1)1s(2)\right)\frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \tag{420}$$

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:

$$\Psi_t = \frac{1}{\sqrt{2}}(1s(1)2s(2) - 2s(1)1s(2))(\alpha(1)\alpha(2)) \tag{421}$$

or

$$\Psi_t = \frac{1}{\sqrt{2}}(1s(1)2s(2) - 2s(1)1s(2))\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) + \beta(1)\alpha(2)) \tag{422}$$

or

$$\Psi_t = \frac{1}{\sqrt{2}}(1s(1)2s(2) - 2s(1)1s(2))(\beta(1)\beta(2)) \tag{423}$$

You can think of these functions as “linear combinations of freshman chemistry electron diagrams”. The diagrams corresponding to the different spin combinations are shown on the next page.

### 12.3 Hund's first rule

Let's look at the behavior of the two spatial wavefunctions $\psi_s$ and $\psi_a$, shown in Eq. (418) and Eq. (419), 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 $\psi_s$ does not have that property; it is the one that goes with the singlet spin function. It is therefore possible to get large contributions to repulsive energy in the singlet by bringing the two electrons very close together. The triplet spatial function does not allow them to come close together (there is a "Fermi hole" in the probability density of electron 2 at the position of electron 1) so these large repulsive energies never occur, and the triplet total energy is lower.

This is an example of "Hund's first rule": for two atomic states with the same electron configuration, the one with the larger value of $S$ will be lower in energy. $S$ is the quantum number giving the eigenvalues of the $S^2$ operator (total spin). In the singlet $S = 0$; in the triplet $S = 1$, as you show in your homework.

The experimentally observed values of the He total energies are

$$1s^2 \quad (^1S) \quad -2.9033E_h \tag{424}$$

$$1s2s \quad (^3S) \quad -2.1750 \tag{425}$$

$$1s2s \quad (^1S) \quad -2.1457 \tag{426}$$

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 $\sim eV$.)

## 13 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.

## 13.1 Addition of Angular Momenta

When two angular momenta $\ell_\mathbf{1}$ and $\ell_\mathbf{2}$ 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.

| $\ell_1$ | $M_1$ | $\ell_2$ | $M_2$ | $M$ |
|---|---|---|---|---|
| 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 state (for light atoms, $Z \leq 40$, 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 $(3)(4)\hbar^2$.

## 13.2 Finding the possible atomic states

To find the atomic states arising from a particular configuration:

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.

Example (for carbon):

$$M_L = 0+0+0+0+-1+0 = -1$$
$$M_S = \tfrac{1}{2} + -\tfrac{1}{2} + \tfrac{1}{2} + -\tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} = 1$$

The "filled subshells" will always add to 0, 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 determinant, 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 $^1S$.

## 13.3 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

$$H_{\mathrm{SO}} \approx A\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \tag{427}$$

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 $\mathrm{cm}^{-1}$ in light atoms, but several thousands of $\mathrm{cm}^{-1}$ 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 $7603\,\mathrm{cm}^{-1}$.

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.

# 14 Molecules

Consider the simplest molecular ion, $H_2^+$.

The Hamiltonian is easy to write down:

$$\hat{H} = \frac{P_e^2}{2m_e} + \frac{P_A^2}{2M_A} + \frac{P_B^2}{2M_B} - \frac{Z_A e^2}{r_A} - \frac{Z_B e^2}{r_B} + \frac{Z_A Z_B e^2}{R_{AB}} \tag{428}$$

In atomic units,

$$\hat{H} = -\frac{1}{2}\nabla_e^2 - \frac{m_e}{M_A}\frac{1}{2}\nabla_A^2 - \frac{m_e}{M_B}\frac{1}{2}\nabla_B^2 - \frac{Z_A}{r_A} - \frac{Z_B}{r_B} + \frac{Z_A Z_B}{R_{AB}} \tag{429}$$

This is a three-body problem, like that for He.

## 14.1 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 go away.
2. The term $\frac{Z_A Z_B e^2}{R_{AB}}$ becomes constant. We can just make it part of the energy eigenvalue.

Now we need to solve the Schrödinger equation

$$\left[\frac{P_e^2}{2m_e} - \frac{Z_A e^2}{r_A} - \frac{Z_B e^2}{r_B}\right]\psi_{\mathrm{el}}(\mathbf{r}_1) = E_{\mathrm{el}}\psi_{\mathrm{el}}(\mathbf{r}_1) \tag{430}$$

That gets us an electronic wavefunction $\psi_{\mathrm{el}}(\mathbf{r}_1)$ and energy $E_{\mathrm{el}}$ 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

$$\Psi(\mathbf{r}_i, \mathbf{R}_j) = \psi_{\text{el}}(\mathbf{r}_i, \mathbf{R}_j)\psi_n(\mathbf{R}_j) \tag{431}$$

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 $E_{\text{el}}$ as a function of $R_{AB}$.

2. Take $E_{\text{el}}(R_{AB}) + \frac{Z_A Z_B e^2}{R_{AB}}$ 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 $E_{\text{el}}(R_{AB})$ as a power series in $R_{AB}$. Let us study the molecular electronic problem now, starting with $\text{H}_2^+$.

## 14.2 Simple variational treatment of $\text{H}_2^+$

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

$$\psi_{e1} \approx c_A \varphi_A + c_B \varphi_B \tag{432}$$

where $\varphi_A$ is some one-electron function centered at nucleus $A$ and $\varphi_B$ is the same function but centered at nucleus $B$. If the two functions are exponential and normalized then they look like

and we have

$$\langle \varphi_A | \varphi_A \rangle = \langle \varphi_B | \varphi_B \rangle = 1 \tag{433}$$

$$\langle \varphi_A | \varphi_B \rangle = S_{AB} \neq 0 \tag{434}$$

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

$$\begin{vmatrix} H_{AA} - E & H_{AB} - ES_{AB} \\ H_{AB} - ES_{AB} & H_{BB} - E \end{vmatrix} = 0 \tag{435}$$

where

$$H_{AA} = \langle \varphi_A | H_{\text{el}} | \varphi_A \rangle \tag{436}$$
$$H_{BB} = \langle \varphi_B | H_{\text{el}} | \varphi_B \rangle \tag{437}$$
$$\langle H_{AB} \rangle = \langle \varphi_A | H_{\text{el}} | \varphi_B \rangle = \langle \varphi_B | H_{\text{el}} | \varphi_A \rangle \tag{438}$$

In this problem, $H_{AA} = H_{BB}(\varphi_A$ and $\varphi_B$ are the same function and the nuclei are equivalent) so we have

$$\begin{aligned} (H_{AA} - E)^2 &- (H_{AB} - ES_{AB})^2 = 0 && (439) \\ H_{AA} - E &= \pm(H_{AB} - ES_{AB}) && (440) \\ E(1 \mp S_{AB}) &= H_{AA} \mp H_{AB} && (441) \\ E &= \frac{H_{AA} \mp H_{AB}}{1 \mp S_{AB}} && (442) \end{aligned}$$

So we find two possible values of E,

$$\begin{aligned} E_- &= \frac{H_{AA} - H_{AB}}{1 - S_{AB}}, \text{and} && (443) \\ E_+ &= \frac{H_{AA} + H_{AB}}{1 + S_{AB}} && (444) \end{aligned}$$

To find the values of $c_A$ and $c_B$ we substitute each of those values into

$$\begin{pmatrix} H_{AA} - E & H_{AB} - ES_{AB} \\ H_{AB} - ES_{AB} & H_{AA} - E \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix} = 0 \tag{445}$$

Substituting $E_+$ into the first equation gives (after a little algebra) $c_A = c_B$; substituting $E_-$ gives $c_A = -c_B$.

Then, normalizing $\psi_{\text{el}}$ gives

$$\psi_+ = (2 + 2S_{AB})^{-\frac{1}{2}}(\varphi_A + \varphi_B) \tag{446}$$
$$\psi_- = (2 - 2S_{AB})^{-\frac{1}{2}}(\varphi_A - \varphi_B) \tag{447}$$

If I take the $\varphi$'s as H-atom 1$s$ orbitals, I can draw plots of $\psi_+$ and $\psi_-$; these are shown in Figure 8. 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 1$s$ 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 $E_h$.

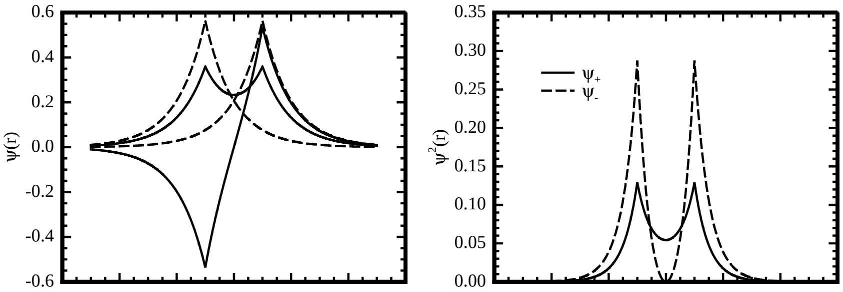


Figure 8: The two solutions of the simple variational treatment for $H_2^+$, using H atom orbitals as the atomic basis functions.

## 14.3 Molecular orbitals

We have seen how to make a bonding and an antibonding orbital from 1$s$ hydrogenlike atomic orbitals centered on two atoms:

$$\psi_- = N_-(\varphi_{1sA} - \varphi_{1sB}) \tag{448}$$

$$\psi_+ = N_+(\varphi_{1sA} + \varphi_{1sB}) \tag{449}$$

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 $\psi_+$ is $g$, and the antibonding $\psi_-$ 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: $\sigma$ for $m_l = 0$, $\pi$ for $|m_l| = 1$, $\delta$ for $|m_l| = 2$, and so on.

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 $\sigma$ orbitals. $\psi_+$ retains its sign when the electron coordinates are inverted through the origin, so it is $\sigma_g$. $\psi_-$ changes sign under inversion, so it is $\sigma_u$.

A couple of other notation conventions are common. Often a superscript asterisk is used to indicate antibonding orbitals; therefore, in homonuclear diatomic molecules, $\sigma^*$ means the same as $\sigma_u$. (In heteronuclear diatomics and in $\sigma$ 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 ($\sigma$, $\pi$, etc.) simply given numbers in order of increasing energy; the lowest $\sigma$ will be called $1\sigma$, the next $2\sigma$ or $2\sigma^*$, the next $3\sigma$, 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:

With $\pi$ orbitals, the bonding ones are $u$ and the antibonding ones are $g$. How about MOs made from the $2p_z$?

In fact the best description of diatomic $\sigma$ MOs arises 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.

## 14.4 MOs for diatomic molecules

To describe $H_2$, we need only place a second electron in the $\sigma_g$ orbital made from the $1s$ atomic orbitals. Since the two electrons then have the same spatial wavefunction, the spin wavefunction must be the antisymmetric a (singlet) one, just as in the ground state of helium.

### 14.4.1 Coulomb and exchange integrals

The $H_2$ bond is much more than twice as strong as the bond in $H_2^+$. The extra strength comes from a surprising source. Consider the simplest molecular orbital picture, putting both electrons in the $\psi_+$ MO. 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

$$J = \langle \varphi_A(1)\varphi_B(2)| - \frac{1}{r_{B1}} - \frac{1}{r_{A2}} + \frac{1}{r_{12}} + \frac{1}{r_{AB}} |\varphi_A(1)\varphi_B(2)\rangle. \tag{450}$$

$J$ 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 term is

$$K = \langle \varphi_A(1)\varphi_B(2)| - \frac{1}{r_{A1}} - \frac{1}{r_{B2}} + \frac{1}{r_{12}} + \frac{1}{r_{AB}} |\varphi_B(1)\varphi_A(2)\rangle \tag{451}$$

$K$ 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).

The total energy of $H_2$ is (in atomic units)

$$E = -1 + \frac{J+K}{1+S^2}, \tag{452}$$

where $S$ is the overlap integral as before. $-1$ is the energy of two separated H atoms, so the "binding energy" is

$$D_e = \frac{J+K}{1+S^2} \tag{453}$$

$J$ has a simple form in terms of $R_{AB}$, but $K$ does not; it must be integrated numerically. Both $J$ (which is the only contribution to the bonding in $H_2^+$) and $K$ have minima as functions of $R_{AB}$, but the minumum in $K$ is much deeper and contributes most of the binding energy of $H_2$.

To form a helium diatomic molecule, we would put the next two electrons in the antibonding $\sigma_u$ formed from the same two 1$s$ 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. 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. This work is mentioned in a footnote at the bottom of page 399 of Atkins. It generated some controversy, but was eventually confirmed.)

## 14.5 Second row diatomics

The homonuclear and "nearly-homonuclear" diatomic molecules from the second row of the periodic table all have very similar molecular orbitals. The lowest two orbitals are the $\sigma_g$ and $\sigma_u$ formed from the atomic 1$s$ 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 $\sigma_g$ and $\sigma_u$, formed this time principally from the 2$s$ atomic orbitals. In $Li_2$ the bonding $\sigma_g$ is occupied by two electrons and the antibonding $\sigma_u$ 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 $\sigma_g$ formed principally from the $2p_z$ orbitals, or the $\pi_u$ ones formed from the $2p_x$ and $2p_y$. In fact those two sets of orbitals are very close in energy, and either the $\sigma$ or the $\pi$ pair can be lower depending on the nuclear charge and the number of electrons in the molecule. For neutral homonuclear diatomics, the $\pi$ pair is lower for $N_2$ and the lighter atoms, while the $\sigma$ is lower for $O_2$ and $F_2$. Both the diagrams are shown on the next page.

After the bonding $\sigma$ and $\pi$ orbitals formed from the 2$p$ atomic orbitals, the next MOs are the antibonding $\pi^*$ and then $\sigma^*$ 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 2 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.

## 14.6 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 $M_L$(max) you do not need to find all the integers between $M_L$(max) and $-M_L$(max),

**Table 2** 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 | $\sim 6200$ | $\sim 1 \times 10^{-7}$ | $(^1\Sigma)$ |
| $Li_2$ | 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 | $\sim 400$ | $\sim 1 \times 10^{-3}$ | $(^1\Sigma)$ |

as in the atomic case, but only $M_L$(max) and $-M_L$(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 $\Sigma$, $\Pi$, $\Delta$, $\Phi$.

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 left side of Figure 9. 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 $M_S$.

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 $M_L = 0$

Figure 9: MO diagram for $B_2$ with one possible arrangement of the highest two electrons (left), and the six possible arrangements of those electrons (right).

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 electrons go into a pair of $\pi^*$ orbitals rather than $\pi$, 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.

## 14.7 Hybrid orbitals

One useful way to describe bonding in more complicated molecules is to use *hybrid* orbitals. A molecular orbital, as we have seen, is formed by taking linear combinations of atomic orbitals on two different nuclei. To form a hybrid orbital, we take a linear combination of atomic orbitals on the same nucleus. This gives us a new set of “atomic” orbitals, but by taking such linear combinations, we can get new atomic orbitals whose main electron density lobes point in “useful” directions, that is, toward other atoms in the molecule. Then, we get bonding orbitals by taking linear combinations of appropriate atomic or hybrid orbitals on the two adjacent nuclei.

For example, consider $BH_3$. Look at the following linear combinations of atomic orbitals:

$$\phi_1 = \sqrt{\frac{1}{3}}2s + \sqrt{\frac{2}{3}}2p_z \tag{454}$$

$$\phi_2 = \sqrt{\frac{1}{3}}2s - \sqrt{\frac{1}{6}}2p_z + \sqrt{\frac{1}{2}}2p_x \tag{455}$$

$$\phi_3 = \sqrt{\frac{1}{3}}2s - \sqrt{\frac{1}{6}}2p_z - \sqrt{\frac{1}{2}}2p_x \tag{456}$$

These three functions form three $sp^2$ hybrid orbitals; in each one, there is a main positive lobe and a smaller negative lobe, and the three main lobes point away from the nucleus in directions that are 120° apart. With these three hybrid orbitals, we can now describe each of the single bonds in $BH_3$ by taking a simple linear combination of one hybrid orbital and one $1s$ orbital located on an H atom. Such a combination makes a $\sigma$-type bonding orbital. There is, of course, another linear combination of hybrid orbital and $1s$ with opposite signs, that puts a node between the two nuclei; that is an antibonding orbital.

Note that the $2p_y$ atomic orbital was not used at all in the hybridization scheme I just described. It is still there, and can be "used" in describing other characteristics of $BH_3$. In isolated, neutral $BH_3$ that third atomic orbital is usually treated as empty, but in $CH_3$, for example, it holds a single unpaired electron. (The methyl radical is a well-known species; it is paramegnetic because of the single unpaired electron, and is trigonal planar with three equivalent single bonds.)