

Chapter 1

The basics of quantum mechanics

1.1 Why quantum mechanics is necessary for describing molecular properties

We know that all molecules are made of atoms which, in turn, contain nuclei and electrons. As I discuss in this introductory section, the equations that govern the motions of electrons and of nuclei are not the familiar Newton equations,

$$\mathbf{F} = m\mathbf{a}. \quad (1.1)$$

but a new set of equations called Schrödinger equations. When scientists first studied the behavior of electrons and nuclei, they tried to interpret their experimental findings in terms of classical Newtonian motions, but such attempts eventually failed. They found that such small light particles behaved in a way that simply is not consistent with the Newton equations. Let me now illustrate some of the experimental data that gave rise to these paradoxes and show you how the scientists of those early times then used these data to suggest new equations that these particles might obey. I want to stress that the Schrödinger equation was not derived but postulated by these scientists. In fact, to date, no one has been able to derive the Schrödinger equation.

From the pioneering work of Bragg on diffraction of x-rays from planes of atoms or ions in crystals, it was known that peaks in the intensity of diffracted x-rays having wavelength λ would occur at scattering angles θ determined by the famous Bragg equation:

$$n\lambda = 2d \sin \theta, \quad (1.2)$$

where d is the spacing between neighboring planes of atoms or ions. These quantities are illustrated in Fig. 1.1. There are many such diffraction peaks, each labeled by a different value of the integer n ($n = 1, 2, 3, \dots$). The Bragg formula can be derived by considering when two photons, one scattering from the second plane in the figure and the second scattering from the third plane, will undergo constructive interference. This condition is met when the “extra path length”

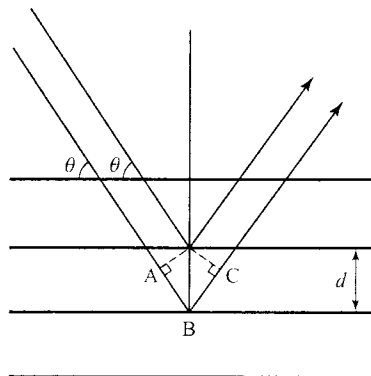


Figure 1.1 Scattering of two beams at angle θ from two planes in a crystal spaced by d .

covered by the second photon (i.e., the length from points A to B to C) is an integer multiple of the wavelength of the photons.

The importance of these x-ray scattering experiments to the study of electrons and nuclei appears in the experiments of Davisson and Germer, in 1927, who scattered electrons of (reasonably) fixed kinetic energy E from metallic crystals. These workers found that plots of the number of scattered electrons as a function of scattering angle θ displayed “peaks” at angles θ that obeyed a Bragg-like equation. The startling thing about this observation is that electrons are particles, yet the Bragg equation is based on the properties of waves. An important observation derived from the Davisson–Germer experiments was that the scattering angles θ observed for electrons of kinetic energy E could be fit to the Bragg $n\lambda = 2d \sin \theta$ equation if a wavelength were ascribed to these electrons that was defined by

$$\lambda = h/(2m_e E)^{1/2}, \quad (1.3)$$

where m_e is the mass of the electron and h is the constant introduced by Max Planck and Albert Einstein in the early 1900s to relate a photon’s energy E to its frequency ν via $E = h\nu$. These amazing findings were among the earliest to suggest that electrons, which had always been viewed as particles, might have some properties usually ascribed to waves. That is, as de Broglie suggested in 1925, an electron seems to have a wavelength inversely related to its momentum, and to display wave-type diffraction. I should mention that analogous diffraction was also observed when other small light particles (e.g., protons, neutrons, nuclei, and small atomic ions) were scattered from crystal planes. In all such cases, Bragg-like diffraction is observed and the Bragg equation is found to govern the scattering angles if one assigns a wavelength to the scattering particle according to

$$\lambda = h/(2mE)^{1/2}, \quad (1.4)$$

where m is the mass of the scattered particle and h is Planck’s constant (6.62×10^{-27} erg s).

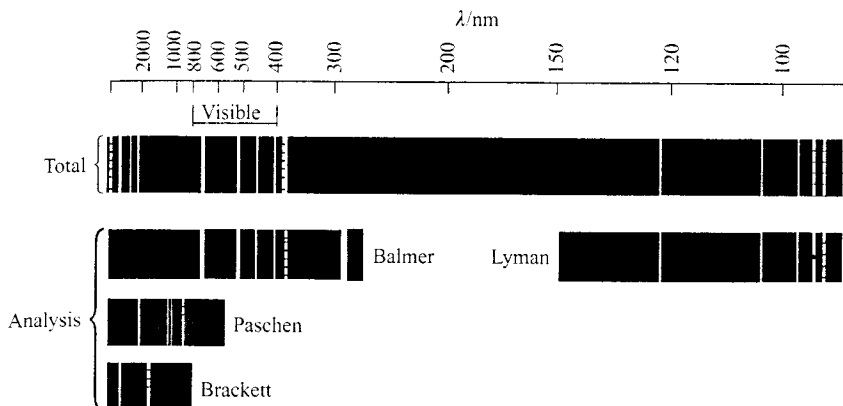


Figure 1.2 Emission spectrum of atomic hydrogen with some lines repeated below to illustrate the series to which they belong.

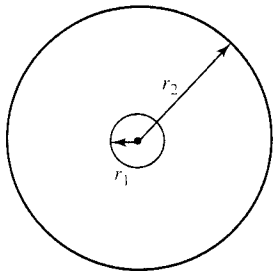
The observation that electrons and other small light particles display wave-like behavior was important because these particles are what all atoms and molecules are made of. So, if we want to fully understand the motions and behavior of molecules, we must be sure that we can adequately describe such properties for their constituents. Because the classical Newton equations do not contain factors that suggest wave properties for electrons or nuclei moving freely in space, the above behaviors presented significant challenges.

Another problem that arose in early studies of atoms and molecules resulted from the study of the photons emitted from atoms and ions that had been heated or otherwise excited (e.g., by electric discharge). It was found that each kind of atom (i.e., H or C or O) emitted photons whose frequencies ν were of very characteristic values. An example of such emission spectra is shown in Fig. 1.2 for hydrogen atoms. In the top panel, we see all of the lines emitted with their wavelengths indicated in nanometers. The other panels show how these lines have been analyzed (by scientists whose names are associated) into patterns that relate to the specific energy levels between which transitions occur to emit the corresponding photons.

In the early attempts to rationalize such spectra in terms of electronic motions, one described an electron as moving about the atomic nuclei in circular orbits such as shown in Fig. 1.3. A circular orbit was thought to be stable when the outward centrifugal force characterized by radius r and speed v ($m_e v^2/r$) on the electron perfectly counterbalanced the inward attractive Coulomb force (Ze^2/r^2) exerted by the nucleus of charge Z :

$$m_e v^2/r = Ze^2/r^2. \quad (1.5)$$

This equation, in turn, allows one to relate the kinetic energy $\frac{1}{2}m_e v^2$ to the Coulombic energy Ze^2/r , and thus to express the total energy E of an orbit in

**Figure 1.3**

Characterization of small and large stable orbits of radii r_1 and r_2 for an electron moving around a nucleus.

terms of the radius of the orbit:

$$E = \frac{1}{2}m_e v^2 - Ze^2/r = \frac{-1}{2}Ze^2/r. \quad (1.6)$$

The energy characterizing an orbit of radius r , relative to the $E = 0$ reference of energy at $r \rightarrow \infty$, becomes more and more negative (i.e., lower and lower) as r becomes smaller. This relationship between outward and inward forces allows one to conclude that the electron should move faster as it moves closer to the nucleus since $v^2 = Ze^2/(rm_e)$. However, nowhere in this model is a concept that relates to the experimental fact that each atom emits only certain kinds of photons. It was believed that photon emission occurred when an electron moving in a larger circular orbit lost energy and moved to a smaller circular orbit. However, the Newtonian dynamics that produced the above equation would allow orbits of any radius, and hence any energy, to be followed. Thus, it would appear that the electron should be able to emit photons of any energy as it moved from orbit to orbit.

The breakthrough that allowed scientists such as Niels Bohr to apply the circular-orbit model to the observed spectral data involved first introducing the idea that the electron has a wavelength and that this wavelength λ is related to its momentum by the de Broglie equation $\lambda = h/p$. The key step in the Bohr model was to also specify that the radius of the circular orbit be such that the circumference of the circle $2\pi r$ equal an integer (n) multiple of the wavelength λ . Only in this way will the electron's wave experience constructive interference as the electron orbits the nucleus. Thus, the Bohr relationship that is analogous to the Bragg equation that determines at what angles constructive interference can occur is

$$2\pi r = n\lambda. \quad (1.7)$$

Both this equation and the analogous Bragg equation are illustrations of what we call boundary conditions; they are extra conditions placed on the wavelength to produce some desired character in the resultant wave (in these cases, constructive interference). Of course, there remains the question of why one must impose these extra conditions when the Newtonian dynamics do not require them. The resolution of this paradox is one of the things that quantum mechanics does.

Returning to the above analysis and using $\lambda = h/p = h/(mv)$, $2\pi r = n\lambda$, as well as the force-balance equation $m_e v^2/r = Ze^2/r^2$, one can then solve for the radii that stable Bohr orbits obey:

$$r = (nh/2\pi)^2/(m_e Ze^2) \quad (1.8)$$

and, in turn, for the velocities of electrons in these orbits,

$$v = Ze^2/(nh/2\pi). \quad (1.9)$$

These two results then allow one to express the sum of the kinetic ($\frac{1}{2}m_e v^2$) and Coulomb potential ($-Ze^2/r$) energies as

$$E = -\frac{1}{2}m_e Z^2 e^4 / (nh/2\pi)^2. \quad (1.10)$$

Just as in the Bragg diffraction result, which specified at what angles special high intensities occurred in the scattering, there are many stable Bohr orbits, each labeled by a value of the integer n . Those with small n have small radii, high velocities and more negative total energies (n.b., the reference zero of energy corresponds to the electron at $r = \infty$, and with $v = 0$). So, it is the result that only certain orbits are “allowed” that causes only certain energies to occur and thus only certain energies to be observed in the emitted photons.

It turned out that the Bohr formula for the energy levels (labeled by n) of an electron moving about a nucleus could be used to explain the discrete line emission spectra of all one-electron atoms and ions (i.e., H, He⁺, Li²⁺, etc.) to very high precision. In such an interpretation of the experimental data, one claims that a photon of energy

$$h\nu = R (1/n_f^2 - 1/n_i^2) \quad (1.11)$$

is emitted when the atom or ion undergoes a transition from an orbit having quantum number n_i to a lower-energy orbit having n_f . Here the symbol R is used to denote the following collection of factors:

$$R = \frac{1}{2}m_e Z^2 e^4 / (h/2\pi)^2. \quad (1.12)$$

The Bohr formula for energy levels did not agree as well with the observed pattern of emission spectra for species containing more than a single electron. However, it does give a reasonable fit, for example, to the Na atom spectra if one examines only transitions involving the single valence electron. The primary reason for the breakdown of the Bohr formula is the neglect of electron–electron Coulomb repulsions in its derivation. Nevertheless, the success of this model made it clear that discrete emission spectra could only be explained by introducing the concept that not all orbits were “allowed”. Only special orbits that obeyed a constructive-interference condition were really accessible to the electron’s motions. This idea that not all energies were allowed, but only certain “quantized” energies could occur was essential to achieving even a qualitative sense of agreement with the experimental fact that emission spectra were discrete.

In summary, two experimental observations on the behavior of electrons that were crucial to the abandonment of Newtonian dynamics were the observations of electron diffraction and of discrete emission spectra. Both of these findings seem to suggest that electrons have some wave characteristics and that these waves have only certain allowed (i.e., quantized) wavelengths.

So, now we have some idea why the Newton equations fail to account for the dynamical motions of light and small particles such as electrons and nuclei. We see that extra conditions (e.g., the Bragg condition or constraints on the de Broglie wavelength) could be imposed to achieve some degree of agreement with experimental observation. However, we still are left wondering what the equations are that can be applied to properly describe such motions and why the extra conditions are needed. It turns out that a new kind of equation based on combining wave and particle properties needed to be developed to address such issues. These are the so-called Schrödinger equations to which we now turn our attention.

As I said earlier, no one has yet shown that the Schrödinger equation follows deductively from some more fundamental theory. That is, scientists did not derive this equation: they postulated it. Some idea of how the scientists of that era “dreamed up” the Schrödinger equation can be had by examining the time and spatial dependence that characterizes so-called traveling waves. It should be noted that the people who worked on these problems knew a great deal about waves (e.g., sound waves and water waves) and the equations they obeyed. Moreover, they knew that waves could sometimes display the characteristic of quantized wavelengths or frequencies (e.g., fundamentals and overtones in sound waves). They knew, for example, that waves in one dimension that are constrained at two points (e.g., a violin string held fixed at two ends) undergo oscillatory motion in space and time with characteristic frequencies and wavelengths. For example, the motion of the violin string just mentioned can be described as having an amplitude $A(x, t)$ at a position x along its length at time t given by

$$A(x, t) = A(x, 0) \cos(2\pi \nu t), \quad (1.13)$$

where ν is its oscillation frequency. The amplitude’s spatial dependence also has a sinusoidal dependence given by

$$A(x, 0) = A \sin(2\pi x/\lambda), \quad (1.14)$$

where λ is the crest-to-crest length of the wave. Two examples of such waves in one dimension are shown in Fig. 1.4. In these cases, the string is fixed at $x = 0$ and at $x = L$, so the wavelengths belonging to the two waves shown are $\lambda = 2L$ and $\lambda = L$. If the violin string were not clamped at $x = L$, the waves could have any value of λ . However, because the string is attached at $x = L$, the allowed wavelengths are quantized to obey

$$\lambda = L/n, \quad (1.15)$$

where $n = 1, 2, 3, 4, \dots$. The equation that such waves obey, called the wave equation, reads

$$\frac{d^2 A(x, t)}{dt^2} = c^2 \frac{d^2 A}{dx^2}, \quad (1.16)$$

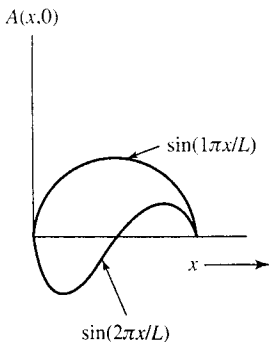


Figure 1.4
Fundamental and first overtone notes of a violin string.

where c is the speed at which the wave travels. This speed depends on the composition of the material from which the violin string is made. Using the earlier expressions for the x - and t -dependences of the wave, $A(x, t)$, we find that the wave's frequency and wavelength are related by the so-called dispersion equation:

$$v^2 = (c/\lambda)^2, \quad (1.17)$$

or

$$c = \lambda v. \quad (1.18)$$

This relationship implies, for example, that an instrument string made of a very stiff material (large c) will produce a higher frequency tone for a given wavelength (i.e., a given value of n) than will a string made of a softer material (smaller c).

For waves moving on the surface of, for example, a rectangular two-dimensional surface of lengths L_x and L_y , one finds

$$A(x, y, t) = \sin(n_x \pi x / L_x) \sin(n_y \pi y / L_y) \cos(2\pi \nu t). \quad (1.19)$$

Hence, the waves are quantized in two dimensions because their wavelengths must be constrained to cause $A(x, y, t)$ to vanish at $x = 0$ and $x = L_x$ as well as at $y = 0$ and $y = L_y$ for all times t . Let us now return to the issue of waves that describe electrons moving.

The pioneers of quantum mechanics examined functional forms similar to those shown above. For example, forms such as $A = \exp[\pm 2\pi i(\nu t - x/\lambda)]$ were considered because they correspond to periodic waves that evolve in x and t under no external x - or t -dependent forces. Noticing that

$$\frac{d^2 A}{dx^2} = -\left(\frac{2\pi}{\lambda}\right)^2 A \quad (1.20)$$

and using the de Broglie hypothesis $\lambda = h/p$ in the above equation, one finds

$$\frac{d^2 A}{dx^2} = -p^2 \left(\frac{2\pi}{h}\right)^2 A. \quad (1.21)$$

If A is supposed to relate to the motion of a particle of momentum p under no external forces (since the waveform corresponds to this case), p^2 can be related to the energy E of the particle by $E = p^2/2m$. So, the equation for A can be rewritten as

$$\frac{d^2 A}{dx^2} = -2mE \left(\frac{2\pi}{h}\right)^2 A, \quad (1.22)$$

or, alternatively,

$$-\left(\frac{h}{2\pi}\right)^2 \frac{d^2 A}{dx^2} = EA. \quad (1.23)$$

Returning to the time-dependence of $A(x, t)$ and using $v = E/h$, one can also show that

$$i \left(\frac{h}{2\pi} \right) \frac{dA}{dt} = EA, \quad (1.24)$$

which, using the first result, suggests that

$$i \left(\frac{h}{2\pi} \right) \frac{dA}{dt} = - \left(\frac{h}{2\pi} \right)^2 \frac{d^2 A}{dx^2}. \quad (1.25)$$

This is a primitive form of the Schrödinger equation that we will address in much more detail below. Briefly, what is important to keep in mind is that the use of the de Broglie and Planck/Einstein connections ($\lambda = h/p$ and $E = h\nu$), both of which involve the constant h , produces suggestive connections between

$$i \left(\frac{h}{2\pi} \right) \frac{d}{dt} \quad \text{and} \quad E \quad (1.26)$$

and between

$$p^2 \quad \text{and} \quad - \left(\frac{h}{2\pi} \right)^2 \frac{d^2}{dx^2} \quad (1.27)$$

or, alternatively, between

$$p \quad \text{and} \quad -i \left(\frac{h}{2\pi} \right) \frac{d}{dx}. \quad (1.28)$$

These connections between physical properties (energy E and momentum p) and differential operators are some of the unusual features of quantum mechanics.

The above discussion about waves and quantized wavelengths as well as the observations about the wave equation and differential operators are not meant to provide or even suggest a derivation of the Schrödinger equation. Again the scientists who invented quantum mechanics did not derive its working equations. Instead, the equations and rules of quantum mechanics have been postulated and designed to be consistent with laboratory observations. My students often find this to be disconcerting because they are hoping and searching for an underlying fundamental basis from which the basic laws of quantum mechanics follow logically. I try to remind them that this is not how theory works. Instead, one uses experimental observation to postulate a rule or equation or theory, and one then tests the theory by making predictions that can be tested by further experiments. If the theory fails, it must be “refined”, and this process continues until one has a better and better theory. In this sense, quantum mechanics, with all of its unusual mathematical constructs and rules, should be viewed as arising from the imaginations of scientists who tried to invent a theory that was consistent with experimental data and which could be used to predict things that could then be tested in the laboratory. Thus far, this theory has proven reliable, but, of course,

we are always searching for a “new and improved” theory that describes how small light particles move.

If it helps you to be more accepting of quantum theory, I should point out that the quantum description of particles will reduce to the classical Newton description under certain circumstances. In particular, when treating heavy particles (e.g., macroscopic masses and even heavier atoms), it is often possible to use Newton dynamics. Briefly, we will discuss in more detail how the quantum and classical dynamics sometimes coincide (in which case one is free to use the simpler Newton dynamics). So, let us now move on to look at this strange Schrödinger equation that we have been digressing about for so long.

1.2 The Schrödinger equation and its components

It has been well established that electrons moving in atoms and molecules do not obey the classical Newton equations of motion. People long ago tried to treat electronic motion classically, and found that features observed clearly in experimental measurements simply were not consistent with such a treatment. Attempts were made to supplement the classical equations with conditions that could be used to rationalize such observations. For example, early workers required that the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ be allowed to assume only integer multiples of $h/2\pi$ (which is often abbreviated as \hbar), which can be shown to be equivalent to the Bohr postulate $n\lambda = 2\pi r$. However, until scientists realized that a new set of laws, those of quantum mechanics, applied to light microscopic particles, a wide gulf existed between laboratory observations of molecule-level phenomena and the equations used to describe such behavior.

Quantum mechanics is cast in a language that is not familiar to most students of chemistry who are examining the subject for the first time. Its mathematical content and how it relates to experimental measurements both require a great deal of effort to master. With these thoughts in mind, I have organized this material in a manner that first provides a brief introduction to the two primary constructs of quantum mechanics – operators and wave functions that obey a Schrödinger equation. Next, I demonstrate the application of these constructs to several chemically relevant model problems. By learning the solutions of the Schrödinger equation for a few model systems, the student can better appreciate the treatment of the fundamental postulates of quantum mechanics as well as their relation to experimental measurement for which the wave functions of the known model problems offer important interpretations.

1.2.1 Operators

Each physically measurable quantity has a corresponding operator. The eigenvalues of the operator tell the only values of the corresponding physical property that can be observed.

Any experimentally measurable physical quantity F (e.g., energy, dipole moment, orbital angular momentum, spin angular momentum, linear momentum, kinetic energy) has a classical mechanical expression in terms of the Cartesian positions $\{q_i\}$ and momenta $\{p_i\}$ of the particles that comprise the system of interest. Each such classical expression is assigned a corresponding quantum mechanical operator \mathbf{F} formed by replacing the $\{p_i\}$ in the classical form by the differential operator $-i\hbar \partial/\partial q_j$ and leaving the coordinates q_j that appear in F untouched. For example, the classical kinetic energy of N particles (with masses m_i) moving in a potential field containing both quadratic and linear coordinate-dependence can be written as

$$F = \sum_{i=1,N} \left[p_i^2/2m_i + 1/2k (q_i - q_i^0)^2 + L (q_i - q_i^0) \right]. \quad (1.29)$$

The quantum mechanical operator associated with this F is

$$\mathbf{F} = \sum_{i=1,N} \left[\frac{-\hbar^2}{2m_i} \frac{\partial^2}{\partial q_i^2} + \frac{1}{2}k (q_i - q_i^0)^2 + L (q_i - q_i^0) \right]. \quad (1.30)$$

Such an operator would occur when, for example, one describes the sum of the kinetic energies of a collection of particles (the $\sum_{i=1,N} (p_i^2/2m_i)$ term), plus the sum of ‘‘Hooke’s Law’’ parabolic potentials (the $1/2 \sum_{i=1,N} k(q_i - q_i^0)^2$), and (the last term in F) the interactions of the particles with an externally applied field whose potential energy varies linearly as the particles move away from their equilibrium positions $\{q_i^0\}$.

Let us try more examples. The sum of the z -components of angular momenta (recall that vector angular momentum \mathbf{L} is defined as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$) of a collection of N particles has the following classical expression:

$$F = \sum_{j=1,N} (x_j p_{yj} - y_j p_{xj}), \quad (1.31)$$

and the corresponding operator is

$$\mathbf{F} = -i\hbar \sum_{j=1,N} \left(x_j \frac{\partial}{\partial y_j} - y_j \frac{\partial}{\partial x_j} \right). \quad (1.32)$$

If one transforms these Cartesian coordinates and derivatives into polar coordinates, the above expression reduces to

$$\mathbf{F} = -i\hbar \sum_{j=1,N} \frac{\partial}{\partial \phi_j}. \quad (1.33)$$

The x -component of the dipole moment for a collection of N particles has a classical form of

$$F = \sum_{j=1,N} Z_j e x_j, \quad (1.34)$$

for which the quantum operator is

$$\mathbf{F} = \sum_{j=1..N} Z_j e x_j, \quad (1.35)$$

where $Z_j e$ is the charge on the j th particle. Notice that in this case, classical and quantum forms are identical because F contains no momentum operators.

The mapping from F to \mathbf{F} is straightforward only in terms of Cartesian coordinates. To map a classical function F , given in terms of curvilinear coordinates (even if they are orthogonal), into its quantum operator is not at all straightforward. The mapping can always be done in terms of Cartesian coordinates after which a transformation of the resulting coordinates and differential operators to a curvilinear system can be performed.

The relationship of these quantum mechanical operators to experimental measurement lies in the eigenvalues of the quantum operators. Each such operator has a corresponding eigenvalue equation

$$\mathbf{F}\chi_j = \alpha_j \chi_j \quad (1.36)$$

in which the χ_j are called eigenfunctions and the (scalar numbers) α_j are called eigenvalues. All such eigenvalue equations are posed in terms of a given operator (\mathbf{F} in this case) and those functions $\{\chi_j\}$ that \mathbf{F} acts on to produce the function back again but multiplied by a constant (the eigenvalue). Because the operator \mathbf{F} usually contains differential operators (coming from the momentum), these equations are differential equations. Their solutions χ_j depend on the coordinates that \mathbf{F} contains as differential operators. An example will help clarify these points. The differential operator d/dy acts on what functions (of y) to generate the same function back again but multiplied by a constant? The answer is functions of the form $\exp(ay)$ since

$$\frac{d(\exp(ay))}{dy} = a \exp(ay). \quad (1.37)$$

So, we say that $\exp(ay)$ is an eigenfunction of d/dy and a is the corresponding eigenvalue.

As I will discuss in more detail shortly, the eigenvalues of the operator \mathbf{F} tell us the *only* values of the physical property corresponding to the operator \mathbf{F} that can be observed in a laboratory measurement. Some \mathbf{F} operators that we encounter possess eigenvalues that are discrete or quantized. For such properties, laboratory measurement will result in only those discrete values. Other \mathbf{F} operators have eigenvalues that can take on a continuous range of values; for these properties, laboratory measurement can give any value in this continuous range.

1.2.2 Wave functions

The eigenfunctions of a quantum mechanical operator depend on the coordinates upon which the operator acts. The particular operator that corresponds to the total energy of the system is called the Hamiltonian operator. The eigenfunctions of this particular operator are called wave functions.

A special case of an operator corresponding to a physically measurable quantity is the Hamiltonian operator H that relates to the total energy of the system. The energy eigenstates of the system Ψ are functions of the coordinates $\{q_j\}$ that H depends on and of time t . The function $|\Psi(q_j, t)|^2 = \Psi^* \Psi$ gives the probability density for observing the coordinates at the values q_j at time t . For a many-particle system such as the H_2O molecule, the wave function depends on many coordinates. For H_2O , it depends on the x , y , and z (or r , θ , and ϕ) coordinates of the ten electrons and the x , y , and z (or r , θ , and ϕ) coordinates of the oxygen nucleus and of the two protons; a total of 39 coordinates appear in Ψ .

In classical mechanics, the coordinates q_j and their corresponding momenta p_j are functions of time. The state of the system is then described by specifying $q_j(t)$ and $p_j(t)$. In quantum mechanics, the concept that q_j is known as a function of time is replaced by the concept of the probability density for finding q_j at a particular value at a particular time $|\Psi(q_j, t)|^2$. Knowledge of the corresponding momenta as functions of time is also relinquished in quantum mechanics; again, only knowledge of the probability density for finding p_j with any particular value at a particular time t remains.

The Hamiltonian eigenstates are especially important in chemistry because many of the tools that chemists use to study molecules probe the energy states of the molecule. For example, most spectroscopic methods are designed to determine which energy state a molecule is in. However, there are other experimental methods that measure other properties (e.g., the z -component of angular momentum or the total angular momentum).

As stated earlier, if the state of some molecular system is characterized by a wave function Ψ that happens to be an eigenfunction of a quantum mechanical operator \mathbf{F} , one can immediately say something about what the outcome will be if the physical property F corresponding to the operator \mathbf{F} is measured. In particular, since

$$\mathbf{F}\chi_j = \lambda_j\chi_j, \quad (1.38)$$

where λ_j is one of the eigenvalues of \mathbf{F} , we know that the value λ_j will be observed if the property F is measured while the molecule is described by the wave function $\Psi = \chi_j$. In fact, once a measurement of a physical quantity F has been carried out and a particular eigenvalue λ_j has been observed, the system's wave function Ψ becomes the eigenfunction χ_j that corresponds to that eigenvalue. That is, the

act of making the measurement causes the system's wave function to become the eigenfunction of the property that was measured.

What happens if some other property G , whose quantum mechanical operator is \mathbf{G} is measured in such a case? We know from what was said earlier that some eigenvalue μ_k of the operator \mathbf{G} will be observed in the measurement. But, will the molecule's wave function remain, after G is measured, the eigenfunction of F , or will the measurement of G cause Ψ to be altered in a way that makes the molecule's state no longer an eigenfunction of F ? It turns out that if the two operators \mathbf{F} and \mathbf{G} obey the condition

$$\mathbf{F}\mathbf{G} = \mathbf{G}\mathbf{F}, \quad (1.39)$$

then, when the property G is measured, the wave function $\Psi = \chi_j$ will remain unchanged. This property, that the order of application of the two operators does not matter, is called commutation; that is, we say the two operators commute if they obey this property. Let us see how this property leads to the conclusion about Ψ remaining unchanged if the two operators commute. In particular, we apply the \mathbf{G} operator to the above eigenvalue equation:

$$\mathbf{G}\mathbf{F}\chi_j = \mathbf{G}\lambda_j\chi_j. \quad (1.40)$$

Next, we use the commutation to re-write the left-hand side of this equation, and use the fact that λ_j is a scalar number to thus obtain

$$\mathbf{F}\mathbf{G}\chi_j = \lambda_j\mathbf{G}\chi_j. \quad (1.41)$$

So, now we see that $(\mathbf{G}\chi_j)$ itself is an eigenfunction of \mathbf{F} having eigenvalue λ_j . So, unless there are more than one eigenfunctions of \mathbf{F} corresponding to the eigenvalue λ_j (i.e., unless this eigenvalue is degenerate), $\mathbf{G}\chi_j$ must itself be proportional to χ_j . We write this proportionality conclusion as

$$\mathbf{G}\chi_j = \mu_j\chi_j, \quad (1.42)$$

which means that χ_j is also an eigenfunction of \mathbf{G} . This, in turn, means that measuring the property G while the system is described by the wave function $\Psi = \chi_j$ does not change the wave function; it remains χ_j .

So, when the operators corresponding to two physical properties commute, once one measures one of the properties (and thus causes the system to be an eigenfunction of that operator), subsequent measurement of the second operator will (if the eigenvalue of the first operator is not degenerate) produce a unique eigenvalue of the second operator and will not change the system wave function.

If the two operators do not commute, one simply can not reach the above conclusions. In such cases, measurement of the property corresponding to the first operator will lead to one of the eigenvalues of that operator and cause the system wave function to become the corresponding eigenfunction. However,

subsequent measurement of the second operator will produce an eigenvalue of that operator, but the system wave function will be changed to become an eigenfunction of the second operator and thus no longer the eigenfunction of the first.

1.2.3 The Schrödinger equation

This equation is an eigenvalue equation for the energy or Hamiltonian operator; its eigenvalues provide the only allowed energy levels of the system.

The time-dependent equation

If the Hamiltonian operator contains the time variable explicitly, one must solve the time-dependent Schrödinger equation.

Before moving deeper into understanding what quantum mechanics “means”, it is useful to learn how the wave functions Ψ are found by applying the basic equation of quantum mechanics, the Schrödinger equation, to a few exactly soluble model problems. Knowing the solutions to these “easy” yet chemically very relevant models will then facilitate learning more of the details about the structure of quantum mechanics.

The Schrödinger equation is a differential equation depending on time and on all of the spatial coordinates necessary to describe the system at hand (thirty-nine for the H_2O example cited above). It is usually written

$$\mathbf{H}\Psi = i\hbar \partial\Psi/\partial t, \quad (1.43)$$

where $\Psi(q_j, t)$ is the unknown wave function and \mathbf{H} is the operator corresponding to the total energy of the system. This operator is called the Hamiltonian and is formed, as stated above, by first writing down the classical mechanical expression for the total energy (kinetic plus potential) in Cartesian coordinates and momenta and then replacing all classical momenta p_j by their quantum mechanical operators $p_j = -i\hbar \partial/\partial q_j$.

For the H_2O example used above, the classical mechanical energy of all thirteen particles is

$$E = \sum_i \left\{ \frac{p_i^2}{2m_e} + 1/2 \sum_j \frac{e^2}{r_{i,j}} - \sum_a \frac{Z_a e^2}{r_{i,a}} \right\} + \sum_a \left\{ \frac{p_a^2}{2m_a} + 1/2 \sum_b \frac{Z_a Z_b e^2}{r_{a,b}} \right\}, \quad (1.44)$$

where the indices i and j are used to label the 10 electrons whose 30 Cartesian coordinates are $\{q_i\}$ and a and b label the three nuclei whose charges are denoted $\{Z_a\}$, and whose nine Cartesian coordinates are $\{q_a\}$. The electron and nuclear

masses are denoted m_e and $\{m_a\}$, respectively. The corresponding Hamiltonian operator is

$$\mathbf{H} = \sum_i \left\{ -\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \sum_j \frac{e^2}{r_{i,j}} - \sum_a \frac{Z_a e^2}{r_{i,a}} \right\} + \sum_a \left\{ -\frac{\hbar^2}{2m_a} \frac{\partial^2}{\partial q_a^2} + \frac{1}{2} \sum_b \frac{Z_a Z_b e^2}{r_{a,b}} \right\}. \quad (1.45)$$

Notice that \mathbf{H} is a second order differential operator in the space of the 39 Cartesian coordinates that describe the positions of the ten electrons and three nuclei. It is a second order operator because the momenta appear in the kinetic energy as p_j^2 and p_a^2 , and the quantum mechanical operator for each momentum $p = -i\hbar\partial/\partial q$ is of first order. The Schrödinger equation for the H_2O example at hand then reads:

$$\sum_i \left\{ -\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \sum_j \frac{e^2}{r_{i,j}} - \sum_a \frac{Z_a e^2}{r_{i,a}} \right\} \Psi + \sum_a \left\{ -\frac{\hbar^2}{2m_a} \frac{\partial^2}{\partial q_a^2} + \frac{1}{2} \sum_b \frac{Z_a Z_b e^2}{r_{a,b}} \right\} \Psi = i\hbar \frac{\partial \Psi}{\partial t}. \quad (1.46)$$

The Hamiltonian in this case contains t nowhere. An example of a case where H does contain t occurs when an oscillating electric field $\mathbf{E} \cos(\omega t)$ along the x -axis interacts with the electrons and nuclei and a term

$$\sum_a Z_a e X_a \mathbf{E} \cos(\omega t) - \sum_j e x_j \mathbf{E} \cos(\omega t) \quad (1.47)$$

is added to the Hamiltonian. Here, X_a and x_j denote the x coordinates of the a th nucleus and the j th electron, respectively.

The time-independent equation

If the Hamiltonian operator does not contain the time variable explicitly, one can solve the time-independent Schrödinger equation.

In cases where the classical energy, and hence the quantum Hamiltonian, do not contain terms that are explicitly time dependent (e.g., interactions with time varying external electric or magnetic fields would add to the above classical energy expression time dependent terms), the separations of variables techniques can be used to reduce the Schrödinger equation to a time-independent equation.

In such cases, \mathbf{H} is not explicitly time dependent, so one can assume that $\Psi(q_j, t)$ is of the form (n.b., this step is an example of the use of the separations of variables method to solve a differential equation)

$$\Psi(q_j, t) = \Psi(q_j)F(t). \quad (1.48)$$

Substituting this “ansatz” into the time-dependent Schrödinger equation gives

$$\Psi(q_j) i\hbar \partial F/\partial t = F(t) \mathbf{H} \Psi(q_j). \quad (1.49)$$

Dividing by $\Psi(q_j)F(t)$ then gives

$$F^{-1} (i\hbar \partial F/\partial t) = \Psi^{-1} [\mathbf{H} \Psi(q_j)]. \quad (1.50)$$

Since $F(t)$ is only a function of time t , and $\Psi(q_j)$ is only a function of the spatial coordinates $\{q_j\}$, and because the left-hand and right-hand sides must be equal for all values of t and of $\{q_j\}$, both the left- and right-hand sides must equal a constant. If this constant is called E , the two equations that are embodied in this separated Schrödinger equation read as follows:

$$\mathbf{H} \Psi(q_j) = E \Psi(q_j), \quad (1.51)$$

$$i\hbar dF(t)/dt = EF(t). \quad (1.52)$$

The first of these equations is called the time-independent Schrödinger equation; it is a so-called eigenvalue equation in which one is asked to find functions that yield a constant multiple of themselves when acted on by the Hamiltonian operator. Such functions are called eigenfunctions of \mathbf{H} and the corresponding constants are called eigenvalues of \mathbf{H} . For example, if \mathbf{H} were of the form $(-\hbar^2/2M)\partial^2/\partial\phi^2 = \mathbf{H}$, then functions of the form $\exp(im\phi)$ would be eigenfunctions because

$$\left\{ -\frac{\hbar^2}{2M} \frac{\partial^2}{\partial\phi^2} \right\} \exp(im\phi) = \left\{ \frac{m^2\hbar^2}{2M} \right\} \exp(im\phi). \quad (1.53)$$

In this case, $m^2\hbar^2/2M$ is the eigenvalue. In this example, the Hamiltonian contains the square of an angular momentum operator (recall earlier that we showed the z -component of angular momentum is to equal $-i\hbar d/d\phi$).

When the Schrödinger equation can be separated to generate a time-independent equation describing the spatial coordinate dependence of the wave function, the eigenvalue E must be returned to the equation determining $F(t)$ to find the time-dependent part of the wave function. By solving

$$i\hbar dF(t)/dt = EF(t) \quad (1.54)$$

once E is known, one obtains

$$F(t) = \exp(-iEt/\hbar), \quad (1.55)$$

and the full wave function can be written as

$$\Psi(q_j, t) = \Psi(q_j) \exp(-iEt/\hbar). \quad (1.56)$$

For the above example, the time dependence is expressed by

$$F(t) = \exp(-it\{m^2\hbar^2/2M\}/\hbar). \quad (1.57)$$

In summary, whenever the Hamiltonian does not depend on time explicitly, one can solve the time-independent Schrödinger equation first and then obtain the time dependence as $\exp(-iEt/\hbar)$ once the energy E is known. In the case of molecular structure theory, it is a quite daunting task even to approximately solve the full Schrödinger equation because it is a partial differential equation depending on all of the coordinates of the electrons and nuclei in the molecule. For this reason, there are various approximations that one usually implements when attempting to study molecular structure using quantum mechanics.

The Born–Oppenheimer approximation

One of the most important approximations relating to applying quantum mechanics to molecules is known as the Born–Oppenheimer (BO) approximation. The basic idea behind this approximation involves realizing that in the full electrons-plus-nuclei Hamiltonian operator introduced above,

$$\mathbf{H} = \sum_i \left\{ -\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \sum_j \frac{e^2}{r_{i,j}} - \sum_a \frac{Z_a e^2}{r_{i,a}} \right\} + \sum_a \left\{ -\frac{\hbar^2}{2m_a} \frac{\partial^2}{\partial q_a^2} + \frac{1}{2} \sum_b \frac{Z_a Z_b e^2}{r_{a,b}} \right\}. \quad (1.58)$$

the time scales with which the electrons and nuclei move are generally quite different. In particular, the heavy nuclei (i.e., even a H nucleus weighs nearly 2000 times what an electron weighs) move (i.e., vibrate and rotate) more slowly than do the lighter electrons. Thus, we expect the electrons to be able to “adjust” their motions to the much more slowly moving nuclei. This observation motivates us to solve the Schrödinger equation for the movement of the electrons in the presence of fixed nuclei as a way to represent the fully adjusted state of the electrons at any fixed positions of the nuclei.

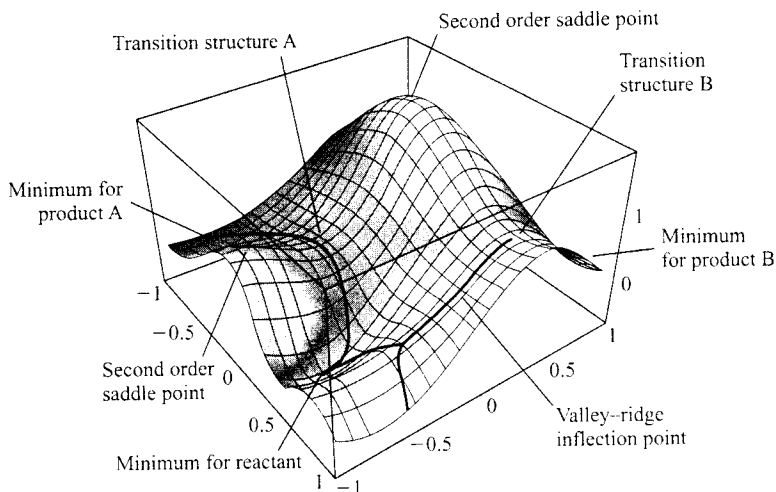
The electronic Hamiltonian that pertains to the motions of the electrons in the presence of so-called clamped nuclei,

$$\mathbf{H} = \sum_i \left\{ -\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \sum_j \frac{e^2}{r_{i,j}} - \sum_a \frac{Z_a e^2}{r_{i,a}} \right\}. \quad (1.59)$$

produces as its eigenvalues, through the equation

$$\mathbf{H}\psi_J(q_j | q_a) = E_J(q_a)\psi_J(q_j | q_a). \quad (1.60)$$

energies $E_K(q_a)$ that depend on where the nuclei are located (i.e., the $\{q_a\}$ coordinates). As its eigenfunctions, one obtains what are called electronic wave functions $\{\psi_K(q_i | q_a)\}$ which also depend on where the nuclei are located. The energies $E_K(q_a)$ are what we usually call potential energy surfaces. An example of such a surface is shown in Fig. 1.5. This surface depends on two geometrical coordinates $\{q_a\}$ and is a plot of one particular eigenvalue $E_J(q_a)$ vs. these two coordinates.

**Figure 1.5**

Two-dimensional potential energy surface showing local minima, transition states and paths connecting them.

Although this plot has more information on it than we shall discuss now, a few features are worth noting. There appear to be three minima (i.e., points where the derivatives of E_J with respect to both coordinates vanish and where the surface has positive curvature). These points correspond, as we will see toward the end of this introductory material, to geometries of stable molecular structures. The surface also displays two first order saddle points (labeled transition structures A and B) that connect the three minima. These points have zero first derivative of E_J with respect to both coordinates but have one direction of negative curvature. As we will show later, these points describe transition states that play crucial roles in the kinetics of transitions among the three stable geometries.

Keep in mind that Fig. 1.5 shows just one of the E_J surfaces; each molecule has a ground-state surface (i.e., the one that is lowest in energy) as well as an infinite number of excited-state surfaces. Let's now return to our discussion of the BO model and ask what one does once one has such an energy surface in hand.

The motions of the nuclei are subsequently, within the BO model, assumed to obey a Schrödinger equation in which $\sum_a \{ -(\hbar^2/2m_a)\partial^2/\partial q_a^2 + 1/2 \sum_b Z_a Z_b e^2/r_{a,b} \} + E_K(q_a)$ defines a rotation-vibration Hamiltonian for the particular energy state E_K of interest. The rotational and vibrational energies and wave functions belonging to each electronic state (i.e., for each value of the index K in $E_K(q_a)$) are then found by solving a Schrödinger equation with such a Hamiltonian.

This BO model forms the basis of much of how chemists view molecular structure and molecular spectroscopy. For example, as applied to formaldehyde $\text{H}_2\text{C}=\text{O}$, we speak of the singlet ground electronic state (with all electrons spin paired and occupying the lowest energy orbitals) and its vibrational states as

well as the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ electronic states and their vibrational levels. Although much more will be said about these concepts later in this text, the student should be aware of the concepts of electronic energy surfaces (i.e., the $\{E_k(q_a)\}$) and the vibration–rotation states that belong to each such surface.

Having been introduced to the concepts of operators, wave functions, the Hamiltonian and its Schrödinger equation, it is important to now consider several examples of the applications of these concepts. The examples treated below were chosen to provide the reader with valuable experience in solving the Schrödinger equation; they were also chosen because they form the most elementary chemical models of electronic motions in conjugated molecules and in atoms, rotations of linear molecules, and vibrations of chemical bonds.

1.3 Your first application of quantum mechanics – motion of a particle in one dimension

This is a very important problem whose solutions chemists use to model a wide variety of phenomena.

Let's begin by examining the motion of a single particle of mass m in one direction which we will call x while under the influence of a potential denoted $V(x)$. The classical expression for the total energy of such a system is $E = p^2/2m + V(x)$, where p is the momentum of the particle along the x -axis. To focus on specific examples, consider how this particle would move if $V(x)$ were of the forms shown in Fig. 1.6, where the total energy E is denoted by the position of the horizontal line.

1.3.1 Classical probability density

I would like you to imagine what the probability density would be for this particle moving with total energy E and with $V(x)$ varying as the three plots in Fig. 1.6

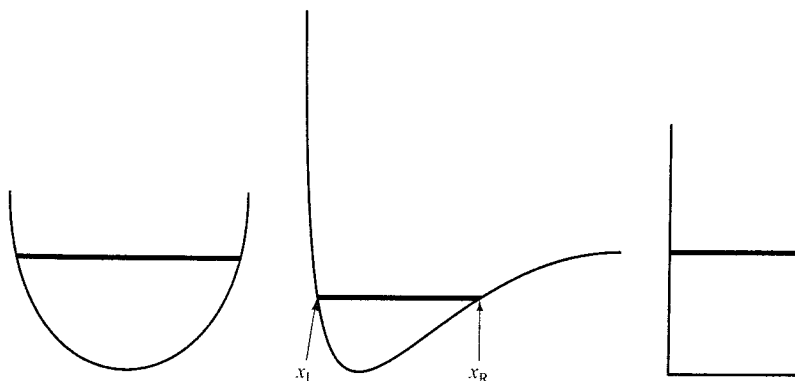


Figure 1.6 Three characteristic potentials showing left and right classical turning points at energies denoted by the horizontal lines.

illustrate. To conceptualize the probability density, imagine the particle to have a blinking lamp attached to it and think of this lamp blinking say 100 times for each time interval it takes for the particle to complete a full transit from its left turning point to its right turning point and back to the former. The turning points x_L and x_R are the positions at which the particle, if it were moving under Newton's laws, would reverse direction (as the momentum changes sign) and turn around. These positions can be found by asking where the momentum goes to zero:

$$0 = p = (2m(E - V(x)))^{1/2}. \quad (1.61)$$

These are the positions where all of the energy appears as potential energy $E = V(x)$ and correspond in the above figures to the points where the dark horizontal lines touch the $V(x)$ plots as shown in the central plot.

The probability density at any value of x represents the fraction of time the particle spends at this value of x (i.e., within x and $x + dx$). Think of forming this density by allowing the blinking lamp attached to the particle to shed light on a photographic plate that is exposed to this light for many oscillations of the particle between x_L and x_R . Alternatively, one can express this probability amplitude $P(x)$ by dividing the spatial distance dx by the velocity of the particle at the point x :

$$P(x) = (2m(E - V(x)))^{-1/2} m dx. \quad (1.62)$$

Because E is constant throughout the particle's motion, $P(x)$ will be small at x values where the particle is moving quickly (i.e., where V is low) and will be high where the particle moves slowly (where V is high). So, the photographic plate will show a bright region where V is high (because the particle moves slowly in such regions) and less brightness where V is low.

The bottom line is that the probability densities anticipated by analyzing the classical Newtonian dynamics of this one particle would appear as the histogram plots shown in Fig. 1.7 illustrate. Where the particle has high kinetic energy (and thus lower $V(x)$), it spends less time and $P(x)$ is small. Where the particle moves

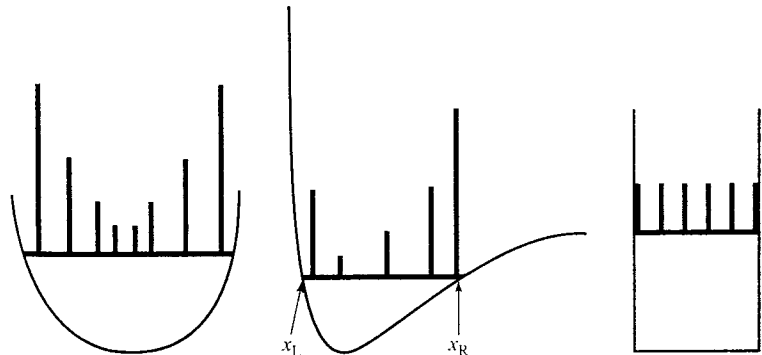


Figure 1.7 Classical probability plots for the three potentials shown in Fig. 1.6.

slowly, it spends more time and $P(x)$ is larger. For the plot on the right, $V(x)$ is constant within the “box”, so the speed is constant, hence $P(x)$ is constant for all x values within this one-dimensional box. I ask that you keep these plots in mind because they are very different from what one finds when one solves the Schrödinger equation for this same problem. Also please keep in mind that these plots represent what one expects if the particle were moving according to classical Newtonian dynamics (which we know it is not!).

1.3.2 Quantum treatment

To solve for the quantum mechanical wave functions and energies of this same problem, we first write the Hamiltonian operator as discussed above by replacing p by $-i\hbar d/dx$:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x). \quad (1.63)$$

We then try to find solutions $\psi(x)$ to $H\psi = E\psi$ that obey certain conditions. These conditions are related to the fact that $|\psi(x)|^2$ is supposed to be the probability density for finding the particle between x and $x + dx$. To keep things as simple as possible, let's focus on the “box” potential V shown in the right side of Fig. 1.7. This potential, expressed as a function of x is: $V(x) = \infty$ for $x < 0$ and for $x > L$; $V(x) = 0$ for x between 0 and L .

The fact that V is infinite for $x < 0$ and for $x > L$, and that the total energy E must be finite, says that ψ must vanish in these two regions ($\psi = 0$ for $x < 0$ and for $x > L$). This condition means that the particle can not access these regions where the potential is infinite. The second condition that we make use of is that $\psi(x)$ must be continuous; this means that the probability of the particle being at x can not be discontinuously related to the probability of it being at a nearby point.

1.3.3 Energies and wave functions

The second order differential equation

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi = E\psi \quad (1.64)$$

has two solutions (because it is a second order equation) in the region between $x = 0$ and $x = L$:

$$\psi = \sin(kx) \quad \text{and} \quad \psi = \cos(kx), \quad \text{where } k \text{ is defined as } k = (2mE/\hbar^2)^{1/2}. \quad (1.65)$$

Hence, the most general solution is some combination of these two:

$$\psi = A \sin(kx) + B \cos(kx). \quad (1.66)$$

The fact that ψ must vanish at $x = 0$ (n.b., ψ vanishes for $x < 0$ and is continuous, so it must vanish at the point $x = 0$) means that the weighting amplitude of the $\cos(kx)$ term must vanish because $\cos(kx) = 1$ at $x = 0$. That is,

$$B = 0. \quad (1.67)$$

The amplitude of the $\sin(kx)$ term is not affected by the condition that ψ vanish at $x = 0$, since $\sin(kx)$ itself vanishes at $x = 0$. So, now we know that ψ is really of the form:

$$\psi(x) = A \sin(kx). \quad (1.68)$$

The condition that ψ also vanish at $x = L$ has two possible implications. Either $A = 0$ or k must be such that $\sin(kL) = 0$. The option $A = 0$ would lead to an answer ψ that vanishes at all values of x and thus a probability that vanishes everywhere. This is unacceptable because it would imply that the particle is never observed anywhere.

The other possibility is that $\sin(kL) = 0$. Let's explore this answer because it offers the first example of energy quantization that you have probably encountered. As you know, the sine function vanishes at integral multiples of π . Hence kL must be some multiple of π ; let's call the integer n and write $Lk = n\pi$ (using the definition of k) in the form:

$$L(2mE/\hbar^2)^{1/2} = n\pi. \quad (1.69)$$

Solving this equation for the energy E , we obtain:

$$E = n^2 \pi^2 \hbar^2 / (2mL^2). \quad (1.70)$$

This result says that the only energy values that are capable of giving a wave function $\psi(x)$ that will obey the above conditions are these specific E values. In other words, not all energy values are "allowed" in the sense that they can produce ψ functions that are continuous and vanish in regions where $V(x)$ is infinite. If one uses an energy E that is not one of the allowed values and substitutes this E into $\sin(kx)$, the resultant function will not vanish at $x = L$. I hope the solution to this problem reminds you of the violin string that we discussed earlier. Recall that the violin string being tied down at $x = 0$ and at $x = L$ gave rise to quantization of the wavelength just as the conditions that ψ be continuous at $x = 0$ and $x = L$ gave energy quantization.

Substituting $k = n\pi/L$ into $\psi = A \sin(kx)$ gives

$$\psi(x) = A \sin(n\pi x/L). \quad (1.71)$$

The value of A can be found by remembering that $|\Psi|^2$ is supposed to represent the probability density for finding the particle at x . Such probability densities are

supposed to be normalized, meaning that their integral over all x values should amount to unity. So, we can find A by requiring that

$$1 = \int |\psi(x)|^2 dx = |A|^2 \int \sin^2(n\pi x/L) dx, \quad (1.72)$$

where the integral ranges from $x = 0$ to $x = L$. Looking up the integral of $\sin^2(ax)$ and solving the above equation for the so-called normalization constant A gives

$$A = (2/L)^{1/2} \quad (1.73)$$

and so

$$\psi(x) = (2/L)^{1/2} \sin(n\pi x/L). \quad (1.74)$$

The values that n can take on are $n = 1, 2, 3, \dots$; the choice $n = 0$ is unacceptable because it would produce a wave function $\psi(x)$ that vanishes at all x .

The full x - and t -dependent wave functions are then given as

$$\Psi(x, t) = \left(\frac{2}{L}\right)^{1/2} \sin \frac{n\pi x}{L} \exp \left[\frac{-itn^2\pi^2\hbar^2}{2mL^2/\hbar} \right]. \quad (1.75)$$

Notice that the spatial probability density $|\Psi(x, t)|^2$ is not dependent on time and is equal to $|\psi(x)|^2$ because the complex exponential disappears when $\Psi^*\Psi$ is formed. This means that the probability of finding the particle at various values of x is time-independent.

Another thing I want you to notice is that, unlike the classical dynamics case, not all energy values E are allowed. In the Newtonian dynamics situation, E could be specified and the particle's momentum at any x value was then determined to within a sign. In contrast, in quantum mechanics one must determine, by solving the Schrödinger equation, what the allowed values of E are. These E values are quantized, meaning that they occur only for discrete values $E = n^2\pi^2\hbar^2/(2mL^2)$ determined by a quantum number n , by the mass of the particle m , and by characteristics of the potential (L in this case).

1.3.4 Probability densities

Let's now look at some of the wave functions $\Psi(x)$ and compare the probability densities $|\Psi(x)|^2$ that they represent to the classical probability densities discussed earlier. The $n = 1$ and $n = 2$ wave functions are shown in the top of Fig. 1.8. The corresponding probability densities are shown below the wave functions in two formats (as x - y plots and shaded plots that could relate to the flashing light way of monitoring the particle's location that we discussed earlier). A more complete set of wave functions (for n ranging from 1 to 7) are shown in Fig. 1.9.

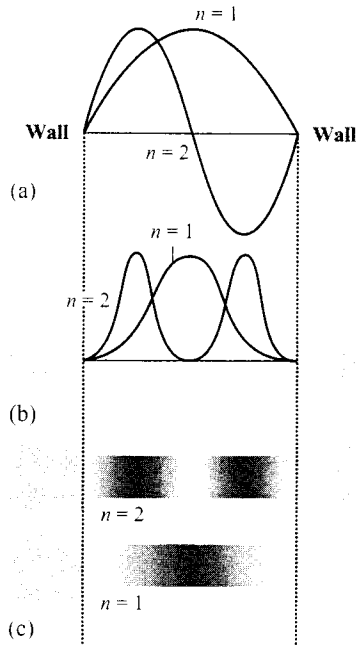


Figure 1.8 The two lowest wave functions and probability densities.

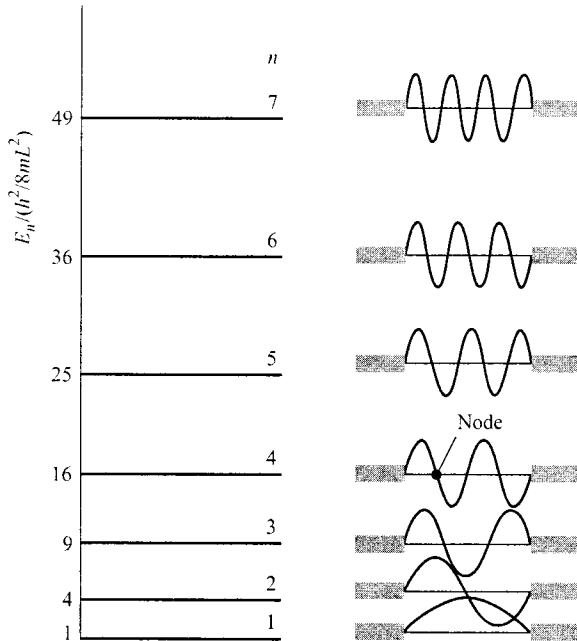


Figure 1.9 Seven lowest wave functions and energies.

Notice that as the quantum number n increases, the energy E also increases (quadratically with n in this case) and the number of nodes in Ψ also increases. Also notice that the probability densities are very different from what we encountered earlier for the classical case. For example, look at the $n = 1$ and $n = 2$ densities and compare them to the classical density illustrated in Fig. 1.10. The classical density is easy to understand because we are familiar with classical dynamics. In this case, we say that $P(x)$ is constant within the box because the fact that $V(x)$ is constant causes the kinetic energy and hence the speed of the particle to remain constant. In contrast, the $n = 1$ quantum wave function's $P(x)$ plot is peaked in the middle of the box and falls to zero at the walls. The $n = 2$ density has two peaks $P(x)$ (one to the left of the box midpoint, and one to the right), a node at the box midpoint, and falls to zero at the walls. One thing that students often ask me is “how does the particle get from being in the left peak to being in the right peak if it has zero chance of ever being at the midpoint where the node is?” The difficulty with this question is that it is posed in a terminology that asks for a classical dynamics answer. That is, by asking “how does the particle get . . .” one is demanding an answer that involves describing its motion (i.e., it moves from here at time t_1 to there at time t_2). Unfortunately, quantum mechanics does not deal with issues such as a particle's trajectory (i.e., where it is at various times) but only with its probability of being somewhere (i.e., $|\Psi|^2$). The next section will treat such paradoxical issues even further.

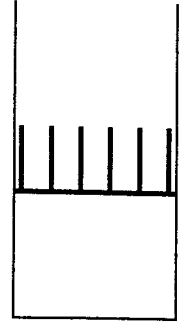


Figure 1.10 Classical probability density for potential shown.

1.3.5 Classical and quantum probability densities

As just noted, it is tempting for most beginning students of quantum mechanics to attempt to interpret the quantum behavior of a particle in classical terms. However, this adventure is full of danger and bound to fail because small, light particles simply do not move according to Newton's laws. To illustrate, let's try to “understand” what kind of (classical) motion would be consistent with the $n = 1$ or $n = 2$ quantum $P(x)$ plots shown in Fig. 1.8. However, as I hope you anticipate, this attempt at gaining classical understanding of a quantum result will not “work” in that it will lead to nonsensical results. My point in leading you to attempt such a classical understanding is to teach you that classical and quantum results are simply different and that you must resist the urge to impose a classical understanding on quantum results.

For the $n = 1$ case, we note that $P(x)$ is highest at the box midpoint and vanishes at $x = 0$ and $x = L$. In a classical mechanics world, this would mean that the particle moves slowly near $x = L/2$ and more quickly near $x = 0$ and $x = L$. Because the particle's total energy E must remain constant as it moves, in regions where it moves slowly, the potential it experiences must be high, and where it moves quickly, V must be small. This analysis (n.b., based on classical concepts) would lead us to conclude that the $n = 1$ $P(x)$ arises from the

particle moving in a potential that is high near $x = L/2$ and low as x approaches 0 or L .

A similar analysis of the $n = 2$ $P(x)$ plot would lead us to conclude that the particle for which this is the correct $P(x)$ must experience a potential that is high midway between $x = 0$ and $x = L/2$, high midway between $x = L/2$ and $x = L$, and very low near $x = L/2$ and near $x = 0$ and $x = L$. These conclusions are “crazy” because we know that the potential $V(x)$ for which we solved the Schrödinger equation to generate both of the wave functions (and both probability densities) is constant between $x = 0$ and $x = L$. That is, we know the same $V(x)$ applies to the particle moving in the $n = 1$ and $n = 2$ states, whereas the classical motion analysis offered above suggests that $V(x)$ is different for these two cases.

What is wrong with our attempt to understand the quantum $P(x)$ plots? The mistake we made was in attempting to apply the equations and concepts of classical dynamics to a $P(x)$ plot that did not arise from classical motion. Simply put, one can not ask how the particle is moving (i.e., what is its speed at various positions) when the particle is undergoing quantum dynamics. Most students, when first experiencing quantum wave functions and quantum probabilities, try to think of the particle moving in a classical way that is consistent with the quantum $P(x)$. This attempt to retain a degree of classical understanding of the particle’s movement is always met with frustration, as I illustrated with the above example and will illustrate later in other cases.

Continuing with this first example of how one solves the Schrödinger equation and how one thinks of the quantized E values and wave functions Ψ , let me offer a little more optimistic note than offered in the preceding discussion. If we examine the $\Psi(x)$ plot shown in Fig. 1.9 for $n = 7$, and think of the corresponding $P(x) = |\Psi(x)|^2$, we note that the $P(x)$ plot would look something like that shown in Fig. 1.11. It would have seven maxima separated by six nodes. If we were to plot

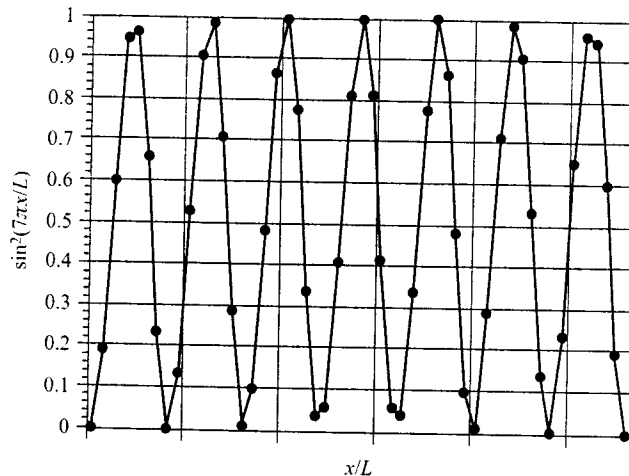


Figure 1.11 Quantum probability density for $n = 7$ showing seven peaks and six nodes.

$|\Psi(x)|^2$ for a very large n value such as $n = 55$, we would find a $P(x)$ plot having 55 maxima separated by 54 nodes, with the maxima separated approximately by distances of $(1/55L)$. Such a plot, when viewed in a “coarse grained” sense (i.e., focusing with somewhat blurred vision on the positions and heights of the maxima) looks very much like the classical $P(x)$ plot in which $P(x)$ is constant for all x . In fact, it is a general result of quantum mechanics that the quantum $P(x)$ distributions for large quantum numbers take on the form of the classical $P(x)$ for the same potential V that was used to solve the Schrödinger equation. It is also true that classical and quantum results agree when one is dealing with heavy particles. For example, a given particle-in-a-box energy $E_n = n^2\hbar^2/(2mL^2)$ would be achieved for a heavier particle at higher n -values than for a lighter particle. Hence, heavier particles, moving with a given energy E , have higher n and thus more classical probability distributions.

We will encounter this so-called quantum–classical correspondence principle again when we examine other model problems. It is an important property of solutions to the Schrödinger equation because it is what allows us to bridge the “gap” between using the Schrödinger equation to treat small, light particles and the Newton equations for macroscopic (big, heavy) systems.

Another thing I would like you to be aware of concerning the solutions ψ and E to this Schrödinger equation is that each pair of wave functions ψ_n and $\psi_{n'}$ belonging to different quantum numbers n and n' (and to different energies) display a property termed orthonormality. This property means that not only are ψ_n and $\psi_{n'}$ each normalized

$$1 = \int |\psi_n|^2 dx = \int |\psi_{n'}|^2 dx, \quad (1.76)$$

but they are also orthogonal to each other

$$0 = \int (\psi_n)^* \psi_{n'} dx, \quad (1.77)$$

where the complex conjugate $*$ of the first function appears only when the ψ solutions contain imaginary components (you have only seen one such case thus far – the $\exp(im\phi)$ eigenfunctions of the z -component of angular momentum). It is common to write the integrals displaying the normalization and orthogonality conditions in the following so-called Dirac notation:

$$1 = \langle \psi_n | \psi_n \rangle \quad 0 = \langle \psi_n | \psi_{n'} \rangle, \quad (1.78)$$

where the $| \rangle$ and $\langle |$ symbols represent ψ and ψ^* , respectively, and putting the two together in the $\langle | \rangle$ construct implies the integration over the variable that ψ depends upon.

The orthogonality condition can be viewed as similar to the condition of two vectors \mathbf{v}_1 and \mathbf{v}_2 being perpendicular, in which case their scalar (sometimes called “dot”) product vanishes $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. I want you to keep this property in mind

because you will soon see that it is a characteristic not only of these particle-in-a-box wave functions but of all wave functions obtained from any Schrödinger equation.

In fact, the orthogonality property is even broader than the above discussion suggests. It turns out that all quantum mechanical operators formed as discussed earlier (replacing Cartesian momenta p by the corresponding $-i\hbar\partial/\partial q$ operator and leaving all Cartesian coordinates as they are) can be shown to be so-called Hermitian operators. This means that they form Hermitian matrices when they are placed between pairs of functions and the coordinates are integrated over. For example, the matrix representation of an operator \mathbf{F} when acting on a set of functions denoted $\{\phi_J\}$ is

$$F_{I,J} = \langle \phi_I | \mathbf{F} | \phi_J \rangle = \int \phi_I^* \mathbf{F} \phi_J dq. \quad (1.79)$$

For all of the operators formed following the rules stated earlier, one finds that these matrices have the following property:

$$F_{I,J} = F_{J,I}^*, \quad (1.80)$$

which makes the matrices what we call Hermitian. If the functions upon which \mathbf{F} acts and \mathbf{F} itself have no imaginary parts (i.e., are real), then the matrices turn out to be symmetric:

$$F_{I,J} = F_{J,I}. \quad (1.81)$$

The importance of the Hermiticity or symmetry of these matrices lies in the fact that it can be shown that such matrices have all real (i.e., not complex) eigenvalues and have eigenvectors that are orthogonal.

So, all quantum mechanical operators, not just the Hamiltonian, have real eigenvalues (this is good since these eigenvalues are what can be measured in any experimental observation of that property) and orthogonal eigenfunctions. It is important to keep these facts in mind because we make use of them many times throughout this text.

1.3.6 Time propagation of wave functions

For a system that exists in an eigenstate $\Psi(x) = (2/L)^{1/2} \sin(n\pi x/L)$ having an energy $E_n = n^2\pi^2\hbar^2/(2mL^2)$, the time-dependent wave function is

$$\Psi(x, t) = \left(\frac{2}{L}\right)^{1/2} \sin \frac{n\pi x}{L} \exp\left(-\frac{itE_n}{\hbar}\right), \quad (1.82)$$

which can be generated by applying the so-called time evolution operator $U(t, 0)$ to the wave function at $t = 0$:

$$\Psi(x, t) = U(t, 0)\Psi(x, 0). \quad (1.83)$$

where an explicit form for $U(t, t')$ is

$$U(t, t') = \exp \left[-\frac{i(t - t')H}{\hbar} \right]. \quad (1.84)$$

The function $\Psi(x, t)$ has a spatial probability density that does not depend on time because

$$\Psi^*(x, t)\Psi(x, t) = \left(\frac{2}{L} \right) \sin^2 \left(\frac{n\pi x}{L} \right), \quad (1.85)$$

since $\exp(-itE_n/\hbar)\exp(itE_n/\hbar) = 1$. However, it is possible to prepare systems (even in real laboratory settings) in states that are not single eigenstates; we call such states superposition states. For example, consider a particle moving along the x -axis within the “box” potential but in a state whose wave function at some initial time $t = 0$ is

$$\Psi(x, 0) = 2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \sin \left(\frac{1\pi x}{L} \right) - 2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \sin \left(\frac{2\pi x}{L} \right). \quad (1.86)$$

This is a superposition of the $n = 1$ and $n = 2$ eigenstates. The probability density associated with this function is

$$|\Psi|^2 = \frac{1}{2} \left\{ \frac{2}{L} \sin^2 \left(\frac{1\pi x}{L} \right) + \frac{2}{L} \sin^2 \left(\frac{2\pi x}{L} \right) - 2 \left(\frac{2}{L} \right) \times \sin \left(\frac{1\pi x}{L} \right) \sin \left(\frac{2\pi x}{L} \right) \right\}. \quad (1.87)$$

The $n = 1$ and $n = 2$ components, the superposition Ψ , and the probability density at $t = 0$ $|\Psi|^2$ are shown in the first three panels of Fig. 1.12. It should be noted that the probability density associated with this superposition state is not symmetric about the $x = L/2$ midpoint even though the $n = 1$ and $n = 2$ component wave functions and densities are. Such a density describes the particle localized more strongly in the large- x region of the box than in the small- x region.

Now, let's consider the superposition wave function and its density at later times. Applying the time evolution operator $\exp(-itH/\hbar)$ to $\Psi(x, 0)$ generates this time-evolved function at time t :

$$\begin{aligned} \Psi(x, t) &= \exp \left(-\frac{itH}{\hbar} \right) \left\{ 2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \sin \left(\frac{1\pi x}{L} \right) - 2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \sin \left(\frac{2\pi x}{L} \right) \right\} \\ &= \left[2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \sin \left(\frac{1\pi x}{L} \right) \exp \left(-\frac{itE_1}{\hbar} \right) - 2^{-1/2} \left(\frac{2}{L} \right)^{1/2} \right. \\ &\quad \left. \times \sin \left(\frac{2\pi x}{L} \right) \exp \left(-\frac{itE_2}{\hbar} \right) \right]. \end{aligned} \quad (1.88)$$

The spatial probability density associated with this Ψ is

$$|\Psi(x, t)|^2 = \frac{1}{2} \left\{ \left(\frac{2}{L} \right) \sin^2 \left(\frac{1\pi x}{L} \right) + \left(\frac{2}{L} \right) \sin^2 \left(\frac{2\pi x}{L} \right) - 2 \left(\frac{2}{L} \right) \cos \left[(E_2 - E_1) \frac{t}{\hbar} \right] \sin \left(\frac{1\pi x}{L} \right) \sin \left(\frac{2\pi x}{L} \right) \right\}. \quad (1.89)$$

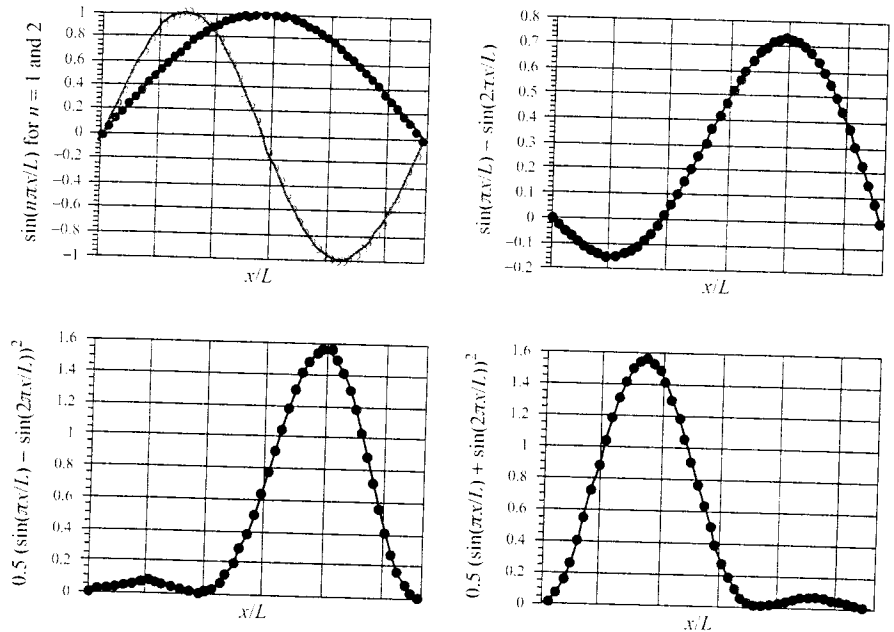


Figure 1.12 The $n = 1$ and $n = 2$ wave functions, their superposition, and the $t = 0$ and time-evolved probability densities of the superposition.

At $t = 0$, this function clearly reduces to that written earlier for $\Psi(x, 0)$. Notice that as time evolves, this density changes because of the $\cos[(E_2 - E_1)t/\hbar]$ factor it contains. In particular, note that as t moves through a period of length $\delta t = \pi\hbar/(E_2 - E_1)$, the cos factor changes sign. That is, for $t = 0$, the cos factor is $+1$; for $t = \pi\hbar/(E_2 - E_1)$, the cos factor is -1 ; for $t = 2\pi\hbar/(E_2 - E_1)$, it returns to $+1$. The result of this time variation in the cos factor is that $|\Psi|^2$ changes in form from that shown in the bottom left panel of Fig. 1.12 to that shown in the bottom right panel (at $t = \pi\hbar/(E_2 - E_1)$) and then back to the form in the bottom left panel (at $t = 2\pi\hbar/(E_2 - E_1)$). One can interpret this time variation as describing the particle's probability density (not its classical position!), initially localized toward the right side of the box, moving to the left and then back to the right. Of course, this time evolution will continue over more and more cycles as time evolves.

This example illustrates once again the difficulty with attempting to localize particles that are being described by quantum wave functions. For example, a particle that is characterized by the eigenstate $(2/L)^{1/2} \sin(\pi x/L)$ is more likely to be detected near $x = L/2$ than near $x = 0$ or $x = L$ because the square of this function is large near $x = L/2$. A particle in the state $(2/L)^{1/2} \sin(2\pi x/L)$ is most likely to be found near $x = L/4$ and $x = 3L/4$, but not near $x = 0$, $x = L/2$, or

$x = L$. The issue of how the particle in the latter state moves from being near $x = L/4$ to $x = 3L/4$ is not something quantum mechanics deals with. Quantum mechanics does not allow us to follow the particle's trajectory, which is what we need to know when we ask how it moves from one place to another. Nevertheless, superposition wave functions can offer, to some extent, the opportunity to follow the motion of the particle. For example, the superposition state written above as $2^{-1/2}(2/L)^{1/2} \sin(\pi x/L) - 2^{-1/2}(2/L)^{1/2} \sin(2\pi x/L)$ has a probability amplitude that changes with time as shown in the figure. Moreover, this amplitude's major peak does move from side to side within the box as time evolves. So, in this case, we can say with what frequency the major peak moves back and forth. In a sense, this allows us to "follow" the particle's movements, but only to the extent that we are satisfied with ascribing its location to the position of the major peak in its probability distribution. That is, we can not really follow its "precise" location, but we can follow the location of where it is very likely to be found. This is an important observation that I hope the student will keep fresh in mind. It is also an important ingredient in modern quantum dynamics in which localized wave packets, similar to superposed eigenstates, are used to detail the position and speed of a particle's main probability density peak.

The above example illustrates how one time-evolves a wave function that can be expressed as a linear combination (i.e., superposition) of eigenstates of the problem at hand. There is a large amount of current effort in the theoretical chemistry community aimed at developing efficient approximations to the $\exp(-itH/\hbar)$ evolution operator that do not require $\Psi(x, 0)$ to be explicitly written as a sum of eigenstates. This is important because, for most systems of direct relevance to molecules, one can not solve for the eigenstates; it is simply too difficult to do so. You can find a significantly more detailed treatment of this subject at the research-level in my TheoryPage web site and my *QMIC* textbook. However, let's spend a little time on a brief introduction to what is involved.

The problem is to express $\exp(-itH/\hbar)\Psi(q_j)$, where $\Psi(q_j)$ is some initial wave function but not an eigenstate, in a manner that does not require one to first find the eigenstates $\{\Psi_J\}$ of H and to expand Ψ in terms of these eigenstates:

$$\Psi = \sum_J C_J \Psi_J \quad (1.90)$$

after which the desired function is written as

$$\exp\left(-\frac{itH}{\hbar}\right)\Psi(q_j) = \sum_J C_J \Psi_J \exp\left(-\frac{itE_J}{\hbar}\right). \quad (1.91)$$

The basic idea is to break H into its kinetic T and potential V energy components and to realize that the differential operators appear in T only. The importance of this observation lies in the fact that T and V do not commute, which means that TV is not equal to VT (n.b., for two quantities to commute means that their

order of appearance does not matter). Why do they not commute? Because T contains second derivatives with respect to the coordinates $\{q_i\}$ that V depends on, so, for example, $d^2/dq^2[V(q)\Psi(q)]$ is not equal to $V(q)d^2/dq^2\Psi(q)$. The fact that T and V do not commute is important because the most common approach to approximating $\exp(-iH/\hbar)$ is to write this single exponential in terms of $\exp(-iT/\hbar)$ and $\exp(-iV/\hbar)$. However, the identity

$$\exp\left(-\frac{iH}{\hbar}\right) = \exp\left(-\frac{iV}{\hbar}\right) \exp\left(-\frac{iT}{\hbar}\right) \quad (1.92)$$

is not fully valid as one can see by expanding all three of the above exponential factors as $\exp(x) = 1 + x + x^2/2! + \dots$, and noting that the two sides of the above equation only agree if one can assume that $TV = VT$, which, as we noted is not true.

In most modern approaches to time propagation, one divides the time interval t into many (i.e., P of them) small time “slices” $\tau = t/P$. One then expresses the evolution operator as a product of P short-time propagators:

$$\begin{aligned} \exp\left(-\frac{iH}{\hbar}\right) &= \exp\left(-\frac{i\tau H}{\hbar}\right) \exp\left(-\frac{i\tau H}{\hbar}\right) \exp\left(-\frac{i\tau H}{\hbar}\right) \dots \\ &= \left[\exp\left(-\frac{i\tau H}{\hbar}\right) \right]^P. \end{aligned} \quad (1.93)$$

If one can develop an efficient means of propagating for a short time τ , one can then do so over and over again P times to achieve the desired full-time propagation.

It can be shown that the exponential operator involving H can better be approximated in terms of the T and V exponential operators as follows:

$$\exp\left(-\frac{i\tau H}{\hbar}\right) \approx \exp\left(-\tau^2 \frac{(TV - VT)}{\hbar^2}\right) \exp\left(-\frac{i\tau V}{\hbar}\right) \exp\left(-\frac{i\tau T}{\hbar}\right). \quad (1.94)$$

So, if one can be satisfied with propagating for very short time intervals (so that the τ^2 term can be neglected), one can indeed use

$$\exp\left(-\frac{i\tau H}{\hbar}\right) \approx \exp\left(-\frac{i\tau V}{\hbar}\right) \exp\left(-\frac{i\tau T}{\hbar}\right) \quad (1.95)$$

as an approximation for the propagator $U(\tau, 0)$.

To progress further, one then expresses $\exp(-i\tau T/\hbar)$ acting on the initial function $\Psi(q)$ in terms of the eigenfunctions of the kinetic energy operator T . Note that these eigenfunctions do not depend on the nature of the potential V , so this step is valid for any and all potentials. The eigenfunctions of $T = -\hbar^2/2md^2/dq^2$ are

$$\psi_p(q) = \left(\frac{1}{2\pi}\right)^{1/2} \exp\left(\frac{ipq}{\hbar}\right) \quad (1.96)$$

and they obey the following orthogonality

$$\int \psi_{p'}^*(q) \psi_p(q) dq = \delta(p' - p) \quad (1.97)$$

and completeness relations

$$\int \psi_p(q) \psi_p^*(q') dp = \delta(q - q'). \quad (1.98)$$

Writing $\Psi(q)$ as

$$\Psi(q) = \int \delta(q - q') \Psi(q') dq'. \quad (1.99)$$

and using the above expression for $\delta(q - q')$ gives

$$\Psi(q) = \iint \psi_p(q) \psi_p^*(q') \Psi(q') dq' dp. \quad (1.100)$$

Then inserting the explicit expressions for $\psi_p(q)$ and $\psi_p^*(q')$ in terms of $\psi_p(q) = (1/2\pi)^{1/2} \exp(ipq/\hbar)$ gives

$$\Psi(q) = \iint \left(\frac{1}{2\pi}\right)^{1/2} \exp\left(\frac{ipq}{\hbar}\right) \left(\frac{1}{2\pi}\right)^{1/2} \exp\left(-\frac{ipq'}{\hbar}\right) \Psi(q') dq' dp. \quad (1.101)$$

Now, allowing $\exp(-i\tau T/\hbar)$ to act on this form for $\Psi(q)$ produces

$$\begin{aligned} \exp\left(-\frac{i\tau T}{\hbar}\right) \Psi(q) &= \iint \exp\left(-\frac{i\tau p^2 \hbar^2}{2m\hbar}\right) \left(\frac{1}{2\pi}\right)^{1/2} \\ &\times \exp\left[\frac{ip(q - q')}{\hbar}\right] \left(\frac{1}{2\pi}\right)^{1/2} \Psi(q') dq' dp. \end{aligned} \quad (1.102)$$

The integral over p above can be carried out analytically and gives

$$\exp\left(-\frac{i\tau T}{\hbar}\right) \Psi(q) = \left(\frac{m}{2\pi\tau\hbar}\right)^{1/2} \int \exp\left[\frac{im(q - q')^2}{2\tau\hbar}\right] \Psi(q') dq'. \quad (1.103)$$

So, the final expression for the short-time propagated wave function is

$$\Psi(q, \tau) = \exp\left[-\frac{i\tau V(q)}{\hbar}\right] \left(\frac{m}{2\pi\tau\hbar}\right)^{1/2} \int \exp\left[\frac{im(q - q')^2}{2\tau\hbar}\right] \Psi(q') dq'. \quad (1.104)$$

which is the working equation one uses to compute $\Psi(q, \tau)$ knowing $\Psi(q)$. Notice that all one needs to know to apply this formula is the potential $V(q)$ at each point in space. One does not need to know any of the eigenfunctions of the Hamiltonian to apply this method. However, one does have to use this formula over and over again to propagate the initial wave function through many small time steps τ to achieve full propagation for the desired time interval $t = P\tau$.

Because this type of time propagation technique is a very active area of research in the theory community, it is likely to continue to be refined and improved. Further discussion of it is beyond the scope of this book, so I will not go further in this direction.

1.4 Free particle motions in more dimensions

The number of dimensions depends on the number of particles and the number of spatial (and other) dimensions needed to characterize the position and motion of each particle.

1.4.1 The Schrödinger equation

Consider an electron of mass m and charge e moving on a two-dimensional surface that defines the x, y plane (e.g., perhaps an electron is constrained to the surface of a solid by a potential that binds it tightly to a narrow region in the z -direction), and assume that the electron experiences a constant and not time-varying potential V_0 at all points in this plane. The pertinent time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x, y) + V_0 \psi(x, y) = E \psi(x, y). \quad (1.105)$$

The task at hand is to solve the above eigenvalue equation to determine the “allowed” energy states for this electron. Because there are no terms in this equation that couple motion in the x and y directions (e.g., no terms of the form $x^a y^b$ or $\partial/\partial x \partial/\partial y$ or $x \partial/\partial y$), separation of variables can be used to write ψ as a product $\psi(x, y) = A(x)B(y)$. Substitution of this form into the Schrödinger equation, followed by collecting together all x -dependent and all y -dependent terms, gives

$$-\frac{\hbar^2}{2m} \frac{1}{A} \frac{\partial^2 A}{\partial x^2} - \frac{\hbar^2}{2m} \frac{1}{B} \frac{\partial^2 B}{\partial y^2} = E - V_0. \quad (1.106)$$

Since the first term contains no y -dependence and the second contains no x -dependence, and because the right side of the equation is independent of both x and y , both terms on the left must actually be constant (these two constants are denoted E_x and E_y , respectively). This observation allows two separate Schrödinger equations to be written:

$$-\frac{\hbar^2}{2m} A^{-1} \frac{\partial^2 A}{\partial x^2} = E_x, \quad (1.107)$$

and

$$-\frac{\hbar^2}{2m} B^{-1} \frac{\partial^2 B}{\partial y^2} = E_y. \quad (1.108)$$

The total energy E can then be expressed in terms of these separate energies E_x and E_y as $E_x + E_y = E - V_0$. Solutions to the x - and y -Schrödinger equations are easily seen to be:

$$A(x) = \exp \left[ix \left(\frac{2mE_x}{\hbar^2} \right)^{1/2} \right] \quad \text{and} \quad \exp \left[-ix \left(\frac{2mE_x}{\hbar^2} \right)^{1/2} \right], \quad (1.109)$$

$$B(y) = \exp \left[iy \left(\frac{2mE_y}{\hbar^2} \right)^{1/2} \right] \quad \text{and} \quad \exp \left[-iy \left(\frac{2mE_y}{\hbar^2} \right)^{1/2} \right]. \quad (1.110)$$

Two independent solutions are obtained for each equation because the x - and y -space Schrödinger equations are both second order differential equations (i.e., a second order differential equation has two independent solutions).

1.4.2 Boundary conditions

The boundary conditions, not the Schrödinger equation, determine whether the eigenvalues will be discrete or continuous.

If the electron is entirely unconstrained within the x, y plane, the energies E_x and E_y can assume any values; this means that the experimenter can “inject” the electron onto the x, y plane with any total energy E and any components E_x and E_y along the two axes as long as $E_x + E_y = E$. In such a situation, one speaks of the energies along both coordinates as being “in the continuum” or “not quantized”.

In contrast, if the electron is constrained to remain within a fixed area in the x, y plane (e.g., a rectangular or circular region), then the situation is qualitatively different. Constraining the electron to any such specified area gives rise to boundary conditions that impose additional requirements on the above A and B functions. These constraints can arise, for example, if the potential $V_0(x, y)$ becomes very large for x, y values outside the region, in which case the probability of finding the electron outside the region is very small. Such a case might represent, for example, a situation in which the molecular structure of the solid surface changes outside the enclosed region in a way that is highly repulsive to the electron (e.g., as in the case of molecular corrals on metal surfaces). This case could then represent a simple model of so-called “corrals” in which the particle is constrained to a finite region of space.

For example, if motion is constrained to take place within a rectangular region defined by $0 \leq x \leq L_x; 0 \leq y \leq L_y$, then the continuity property that all wave functions must obey (because of their interpretation as probability densities, which must be continuous) causes $A(x)$ to vanish at 0 and at L_x . That is, because A must vanish for $x < 0$ and must vanish for $x > L_x$, and because A is continuous, it must vanish at $x = 0$ and at $x = L_x$. Likewise, $B(y)$ must vanish at 0 and at L_y . To implement these constraints for $A(x)$, one must linearly combine the above two solutions $\exp[ix(2mE_x/\hbar^2)^{1/2}]$ and $\exp[-ix(2mE_x/\hbar^2)^{1/2}]$ to achieve a function that vanishes at $x = 0$:

$$A(x) = \exp\left[ix\left(\frac{2mE_x}{\hbar^2}\right)^{1/2}\right] - \exp\left[-ix\left(\frac{2mE_x}{\hbar^2}\right)^{1/2}\right]. \quad (1.111)$$

One is allowed to linearly combine solutions of the Schrödinger equation that have the same energy (i.e., are degenerate) because Schrödinger equations are

linear differential equations. An analogous process must be applied to $B(y)$ to achieve a function that vanishes at $y = 0$:

$$B(y) = \exp\left[iy\left(\frac{2mE_y}{\hbar^2}\right)^{1/2}\right] - \exp\left[-iy\left(\frac{2mE_y}{\hbar^2}\right)^{1/2}\right]. \quad (1.112)$$

Further requiring $A(x)$ and $B(y)$ to vanish, respectively, at $x = L_x$ and $y = L_y$, gives equations that can be obeyed only if E_x and E_y assume particular values:

$$\exp\left[iL_x\left(\frac{2mE_x}{\hbar^2}\right)^{1/2}\right] - \exp\left[-iL_x\left(\frac{2mE_x}{\hbar^2}\right)^{1/2}\right] = 0, \quad (1.113)$$

and

$$\exp\left[iL_y\left(\frac{2mE_y}{\hbar^2}\right)^{1/2}\right] - \exp\left[-iL_y\left(\frac{2mE_y}{\hbar^2}\right)^{1/2}\right] = 0. \quad (1.114)$$

These equations are equivalent (i.e., using $\exp(ix) = \cos x + i \sin x$) to

$$\sin\left[L_x\left(\frac{2mE_x}{\hbar^2}\right)^{1/2}\right] = \sin\left[L_y\left(\frac{2mE_y}{\hbar^2}\right)^{1/2}\right] = 0. \quad (1.115)$$

Knowing that $\sin \theta$ vanishes at $\theta = n\pi$, for $n = 1, 2, 3, \dots$ (although the $\sin(n\pi)$ function vanishes for $n = 0$, this function vanishes for all x or y , and is therefore unacceptable because it represents zero probability density at all points in space), one concludes that the energies E_x and E_y can assume only values that obey

$$L_x\left(\frac{2mE_x}{\hbar^2}\right)^{1/2} = n_x\pi, \quad (1.116)$$

$$L_y\left(\frac{2mE_y}{\hbar^2}\right)^{1/2} = n_y\pi, \quad (1.117)$$

$$\text{or} \quad E_x = \frac{n_x^2\pi^2\hbar^2}{2mL_x^2}, \quad (1.118)$$

$$\text{and} \quad E_y = \frac{n_y^2\pi^2\hbar^2}{2mL_y^2}, \quad \text{with } n_x \text{ and } n_y = 1, 2, 3, \dots \quad (1.119)$$

It is important to stress that it is the imposition of boundary conditions, expressing the fact that the electron is spatially constrained, that gives rise to quantized energies. In the absence of spatial confinement, or with confinement only at $x = 0$ or L_x or only at $y = 0$ or L_y , quantized energies would *not* be realized.

In this example, confinement of the electron to a finite interval along both the x and y coordinates yields energies that are quantized along both axes. If the electron is confined along one coordinate (e.g., between $0 \leq x \leq L_x$) but not along the other (i.e., $B(y)$ is either restricted to vanish at $y = 0$ or at $y = L_y$ or at neither point), then the total energy E lies in the continuum; its E_x component is quantized but E_y is not. Analogs of such cases arise, for example, when a linear triatomic molecule has more than enough energy in one of its bonds to

rupture it but not much energy in the other bond; the first bond's energy lies in the continuum, but the second bond's energy is quantized.

Perhaps more interesting is the case in which the bond with the higher dissociation energy is excited to a level that is not enough to break it but that is in excess of the dissociation energy of the weaker bond. In this case, one has two degenerate states: (i) the strong bond having high internal energy and the weak bond having low energy (ψ_1), and (ii) the strong bond having little energy and the weak bond having more than enough energy to rupture it (ψ_2). Although an experiment may prepare the molecule in a state that contains only the former component (i.e., $\psi = C_1\psi_1 + C_2\psi_2$ with $C_1 = 1$, $C_2 = 0$), coupling between the two degenerate functions (induced by terms in the Hamiltonian \mathbf{H} that have been ignored in defining ψ_1 and ψ_2) usually causes the true wave function $\Psi = \exp(-it\mathbf{H}/\hbar)\psi$ to acquire a component of the second function as time evolves. In such a case, one speaks of internal vibrational energy relaxation (IVR) giving rise to unimolecular decomposition of the molecule.

1.4.3 Energies and wave functions for bound states

For discrete energy levels, the energies are specified functions that depend on quantum numbers, one for each degree of freedom that is quantized.

Returning to the situation in which motion is constrained along both axes, the resultant total energies and wave functions (obtained by inserting the quantum energy levels into the expressions for $A(x)B(y)$) are as follows:

$$E_x = \frac{n_x^2 \pi^2 \hbar^2}{2mL_x^2}, \quad (1.120)$$

$$\text{and } E_y = \frac{n_y^2 \pi^2 \hbar^2}{2mL_y^2}, \quad (1.121)$$

$$E = E_x + E_y + V_0, \quad (1.122)$$

$$\begin{aligned} \psi(x, y) &= \left(\frac{1}{2L_x}\right)^{\frac{1}{2}} \left(\frac{1}{2L_y}\right)^{\frac{1}{2}} \left[\exp\left(\frac{in_x \pi x}{L_x}\right) - \exp\left(\frac{-in_x \pi x}{L_x}\right) \right] \\ &\quad \times \left[\exp\left(\frac{in_y \pi y}{L_y}\right) - \exp\left(\frac{-in_y \pi y}{L_y}\right) \right], \\ &\text{with } n_x \text{ and } n_y = 1, 2, 3, \dots \end{aligned} \quad (1.123)$$

The two $(1/2L)^{1/2}$ factors are included to guarantee that ψ is normalized:

$$\int |\psi(x, y)|^2 dx dy = 1. \quad (1.124)$$

Normalization allows $|\psi(x, y)|^2$ to be properly identified as a probability density for finding the electron at a point x, y .

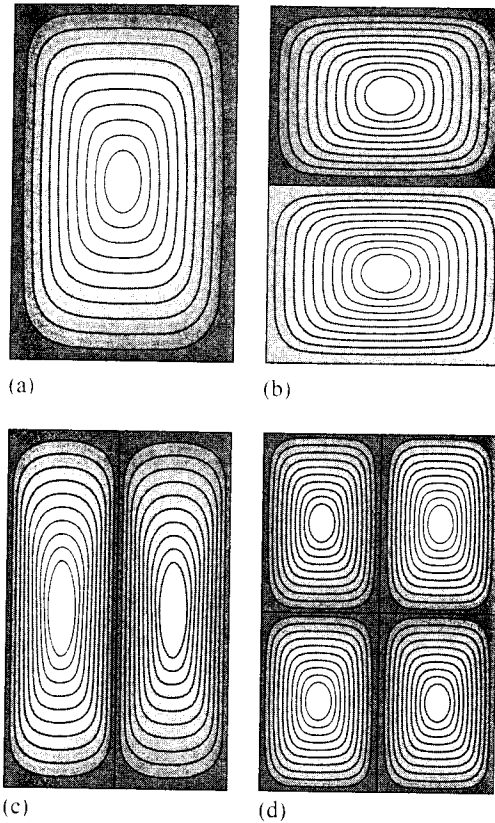


Figure 1.13 Plots of the (a) (1,1), (b) (2,1), (c) (1,2) and (d) (2,2) wave functions.

Shown in Fig. 1.13 are plots of four such two-dimensional wave functions for n_x and n_y values of (1,1), (2,1), (1,2) and (2,2), respectively. Note that the functions vanish on the boundaries of the box, and notice how the number of nodes (i.e., zeroes encountered as the wave function oscillates from positive to negative) is related to the n_x and n_y quantum numbers and to the energy. This pattern of more nodes signifying higher energy is one that we encounter again and again in quantum mechanics and is something the student should be able to use to “guess” the relative energies of wave functions when their plots are at hand. Finally, you should also notice that, as in the one-dimensional box case, any attempt to classically interpret the probabilities $P(x, y)$ corresponding to the above quantum wave functions will result in failure. As in the one-dimensional case, the classical $P(x, y)$ would be constant along slices of fixed x and varying y or slices of fixed y and varying x within the box because the speed is constant there. However, the quantum $P(x, y)$ plots, at least for small quantum numbers, are not constant. For large n_x and n_y values, the quantum $P(x, y)$ plots will again, via the quantum–classical correspondence principle, approach the (constant) classical $P(x, y)$ form.

1.4.4 Quantized action can also be used to derive energy levels

There is another approach that can be used to find energy levels and is especially straightforward to use for systems whose Schrödinger equations are separable. The so-called classical action (denoted S) of a particle moving with momentum \mathbf{p} along a path leading from initial coordinate \mathbf{q}_i at initial time t_i to a final coordinate \mathbf{q}_f at time t_f is defined by

$$S = \int_{\mathbf{q}_i; t_i}^{\mathbf{q}_f; t_f} \mathbf{p} \cdot d\mathbf{q}. \quad (1.125)$$

Here, the momentum vector \mathbf{p} contains the momenta along all coordinates of the system, and the coordinate vector \mathbf{q} likewise contains the coordinates along all such degrees of freedom. For example, in the two-dimensional particle-in-a-box problem considered above, $\mathbf{q} = (x, y)$ has two components as does $\mathbf{p} = (p_x, p_y)$, and the action integral is

$$S = \int_{x_i; y_i; t_i}^{x_f; y_f; t_f} (p_x dx + p_y dy). \quad (1.126)$$

In computing such actions, it is essential to keep in mind the sign of the momentum as the particle moves from its initial to its final positions. An example will help clarify these matters.

For systems such as the above particle-in-a-box example for which the Hamiltonian is separable, the action integral decomposes into a sum of such integrals, one for each degree of freedom. In this two-dimensional example, the additivity of H ,

$$\begin{aligned} H &= H_x + H_y = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + V(x) + V(y) \\ &= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial y^2} + V(y), \end{aligned} \quad (1.127)$$

means that p_x and p_y can be independently solved for in terms of the potentials $V(x)$ and $V(y)$ as well as the energies E_x and E_y associated with each separate degree of freedom:

$$p_x = \pm \sqrt{2m(E_x - V(x))}, \quad (1.128)$$

$$p_y = \pm \sqrt{2m(E_y - V(y))}; \quad (1.129)$$

the signs on p_x and p_y must be chosen to properly reflect the motion that the particle is actually undergoing. Substituting these expressions into the action integral yields

$$S = S_x + S_y \quad (1.130)$$

$$= \int_{x_i; t_i}^{x_f; t_f} \pm \sqrt{2m(E_x - V(x))} dx + \int_{y_i; t_i}^{y_f; t_f} \pm \sqrt{2m(E_y - V(y))} dy. \quad (1.131)$$

The relationship between these classical action integrals and the existence of quantized energy levels has been shown to involve equating the classical action for motion on a closed path to an integral multiple of Planck's constant:

$$S_{\text{closed}} = \int_{q_i, t_i}^{q_f=q_i, t_f} \mathbf{p} \cdot d\mathbf{q} = nh \quad (n = 1, 2, 3, 4, \dots). \quad (1.132)$$

Applied to each of the independent coordinates of the two-dimensional particle-in-a-box problem, this expression reads

$$n_x h = \int_{x=0}^{x=L_x} \sqrt{2m(E_x - V(x))} dx + \int_{x=L_x}^{x=0} -\sqrt{2m(E_x - V(x))} dx, \quad (1.133)$$

$$n_y h = \int_{y=0}^{y=L_y} \sqrt{2m(E_y - V(y))} dy + \int_{y=L_y}^{y=0} -\sqrt{2m(E_y - V(y))} dy. \quad (1.134)$$

Notice that the signs of the momenta are positive in each of the first integrals appearing above (because the particle is moving from $x = 0$ to $x = L_x$, and analogously for y -motion, and thus has positive momentum) and negative in each of the second integrals (because the motion is from $x = L_x$ to $x = 0$ (and analogously for y -motion) and thus the particle has negative momentum). Within the region bounded by $0 \leq x \leq L_x$; $0 \leq y \leq L_y$, the potential vanishes, so $V(x) = V(y) = 0$. Using this fact, and reversing the upper and lower limits, and thus the sign, in the second integrals above, one obtains

$$n_x h = 2 \int_{x=0}^{x=L_x} \sqrt{2mE_x} dx = 2\sqrt{2mE_x} L_x, \quad (1.135)$$

$$n_y h = 2 \int_{y=0}^{y=L_y} \sqrt{2mE_y} dy = 2\sqrt{2mE_y} L_y. \quad (1.136)$$

Solving for E_x and E_y , one finds

$$E_x = \frac{(n_x h)^2}{8mL_x^2}, \quad (1.137)$$

$$E_y = \frac{(n_y h)^2}{8mL_y^2}. \quad (1.138)$$

These are the same quantized energy levels that arose when the wave function boundary conditions were matched at $x = 0$, $x = L_x$ and $y = 0$, $y = L_y$. In this case, one says that the Bohr-Sommerfeld quantization condition,

$$nh = \int_{\mathbf{q}_i; t_i}^{\mathbf{q}_f=q_i; t_f} \mathbf{p} \cdot d\mathbf{q}, \quad (1.139)$$

has been used to obtain the result.

The use of action quantization as illustrated above has become a very important tool. It has allowed scientists to make great progress toward bridging the gap between classical and quantum descriptions of molecular dynamics. In particular,

by using classical concepts such as trajectories and then appending quantal action conditions, people have been able to develop so-called semi-classical models of molecular dynamics. In such models, one is able to retain a great deal of classical understanding while building in quantum effects such as energy quantization, zero-point energies, and interferences.

1.4.5 Quantized action does not always work

Unfortunately, the approach of quantizing the action does not always yield the correct expression for the quantized energies. For example, when applied to the so-called harmonic oscillator problem that we will study in quantum form later, which serves as the simplest reasonable model for vibration of a diatomic molecule AB, one expresses the total energy as

$$E = \frac{p^2}{2\mu} + \frac{k}{2}x^2 \quad (1.140)$$

where $\mu = m_A m_B / (m_A + m_B)$ is the reduced mass of the AB diatom, k is the force constant describing the bond between A and B, x is the bond-length displacement, and p is the momentum along the bond length. The quantized action requirement then reads

$$nh = \int p \, dx = \int \left[2\mu \left(E - \frac{k}{2}x^2 \right) \right]^{1/2} dx. \quad (1.141)$$

This integral is carried out between $x = -(2E/k)^{1/2}$ and $(2E/k)^{1/2}$, the left and right turning points of the oscillatory motion, and back again to form a closed path. Carrying out this integral and equating it to nh gives the following expression for the energy E :

$$E = n \frac{h}{2\pi} \left(\frac{k}{\mu} \right)^{1/2}, \quad (1.142)$$

where the quantum number n is allowed to assume integer values ranging from $n = 0, 1, 2,$ to infinity. The problem with this result is that it is wrong! As experimental data clearly show, the lowest energy levels for the vibrations of a molecule do not have $E = 0$; they have a “zero-point” energy that is approximately equal to $1/2(h/2\pi)(k/\mu)^{1/2}$. So, although the action quantization condition yields energies whose spacings are reasonably in agreement with laboratory data for low-energy states (e.g., such states have approximately constant spacings), it fails to predict the zero-point energy content of such vibrations. As we will see later, a proper quantum mechanical treatment of the harmonic oscillator yields energies of the form

$$E = \left(n + \frac{1}{2} \right) \left(\frac{h}{2\pi} \right) \left(\frac{k}{\mu} \right)^{1/2} \quad (1.143)$$

which differs from the action-based result by the proper zero-point energy.

Even with such difficulties known, much progress has been made in extending the most elementary action-based methods to more and more systems by introducing, for example, rules that allow the quantum number n to assume half-integer as well as integer values. Clearly, if n were allowed to equal $1/2, 3/2, 5/2, \dots$, the earlier action integral would have produced the correct result. However, how does one know when to allow n to assume only integer or only half-integer or both integer and half-integer values? The answers to this question are beyond the scope of this text and constitute an active area of research. For now, it is enough for the student to be aware that one can often find energy levels by using action integrals, but one must be careful in doing so because sometimes the answers are wrong.

Before leaving this section, it is worth noting that the appearance of half-integer quantum numbers does not only occur in the harmonic oscillator case. To illustrate, let us consider the L_z angular momentum operator discussed earlier. As we showed, this operator, when computed as the z -component of $\mathbf{r} \times \mathbf{p}$, can be written in polar (r, θ, ϕ) coordinates as

$$L_z = -i\hbar d/d\phi. \quad (1.144)$$

The eigenfunctions of this operator have the form $\exp(ia\phi)$, and the eigenvalues are $a\hbar$. Because geometries with azimuthal angles equal to ϕ or equal to $\phi + 2\pi$ are exactly the same geometries, the function $\exp(ia\phi)$ should be exactly the same as $\exp(ia(\phi + 2\pi))$. This can only be the case if a is an integer. Thus, one concludes that only integral multiples of \hbar can be “allowed” values of the z -component of angular momentum. Experimentally, one measures the z -component of an angular momentum by placing the system possessing the angular momentum in a magnetic field of strength B and observing how many z -component energy states arise. For example, a boron atom with its 2p orbital has one unit of orbital angular momentum, so one finds three separate z -component values which are usually denoted $m = -1, m = 0$, and $m = 1$. Another example is offered by the scandium atom with one unpaired electron in a d orbital; this atom’s states split into five ($m = -2, -1, 0, 1, 2$) z -component states. In each case, one finds $2L + 1$ values of the m quantum number, and, because L is an integer, $2L + 1$ is an odd integer. Both of these observations are consistent with the expectation that only integer values can occur for L_z eigenvalues.

However, it has been observed that some species do not possess 3 or 5 or 7 or 9 z -component states but an even number of such states. In particular, when electrons, protons, or neutrons are subjected to the kind of magnetic field experiment mentioned above, these particles are observed to have only two z -component eigenvalues. Because, as we discuss later in this text, all angular momenta have z -component eigenvalues that are separated from one another by unit multiples of \hbar , one is forced to conclude that these three fundamental building-block particles have z -component eigenvalues of $1/2\hbar$ and $-1/2\hbar$. The appearance of

half-integer angular momenta is not consistent with the observation made earlier that ϕ and $\phi + 2\pi$ correspond to exactly the same physical point in coordinate space, which, in turn, implies that only full-integer angular momenta are possible.

The resolution of the above paradox (i.e., how can half-integer angular momenta exist?) involved realizing that some angular momenta correspond not to the $\mathbf{r} \times \mathbf{p}$ angular momenta of a physical mass rotating, but, instead, are intrinsic properties of certain particles. That is, the intrinsic angular momenta of electrons, protons, and neutrons can not be viewed as arising from rotation of some mass that comprises these particles. Instead, such intrinsic angular momenta are fundamental “built in” characteristics of these particles. For example, the two $1/2\hbar$ and $-1/2\hbar$ angular momentum states of an electron, usually denoted α and β , respectively, are two internal states of the electron that are degenerate in the absence of a magnetic field but which represent two distinct states of the electron. Analogously, a proton has $1/2\hbar$ and $-1/2\hbar$ states, as do neutrons. All such half-integer angular momentum states can not be accounted for using classical mechanics but are known to arise in quantum mechanics.