A Brief Review of Elementary Quantum Chemistry

C. David Sherrill
School of Chemistry and Biochemistry
Georgia Institute of Technology

Last Revised on 27 January 2001

Contents

1 The Motivation for Quantum Mechanics

Physicists at the end of the nineteenth century believed that most of the fundamental physical laws had been worked out. They expected only minor refinements to get "an extra decimal place" of accuracy. As it turns out, the field of physics was transformed profoundly in the early twentieth century by Einstein's discovery of relativity and by the development of quantum mechanics. While relativity has had fairly little impact on chemistry, all of theoretical chemistry is founded upon quantum mechanics.

The development of quantum mechanics was initially motivated by two observations which demonstrated the inadeqacy of classical physics. These are the "ultraviolet catastrophe" and the photoelectric effect.

1.1 The Ultraviolet Catastrophe

A *blackbody* is an idealized object which absorbs and emits all frequencies. Classical physics can be used to derive an equation which describes the intensity of blackbody radiation as a function of frequency for a fixed temperature—the result is known as the Rayleigh-Jeans law. Although the Rayleigh-Jeans law works for low frequencies, it diverges as ν^2; this divergence for high frequencies is called the ultraviolet catastrophe.

Max Planck explained the blackbody radiation in 1900 by assuming that the energies of the oscillations of electrons which gave rise to the radiation must be proportional to integral multiples of the frequency, i.e.,

$$E = nh\nu \tag{1}$$

Using statistical mechanics, Planck derived an equation similar to the Rayleigh-Jeans equation, but with the adjustable parameter h. Planck found that for $h = 6.626 \times 10^{-34}$ J s, the experimental data could be reproduced. Nevertheless, Planck could not offer a good justification for his assumption of energy quantization.

Physicicsts did not take this energy quantization idea seriously until Einstein invoked a similar assumption to explain the photoelectric effect.

1.2 The Photoelectric Effect

In 1886 and 1887, Heinrich Hertz discovered that ultraviolet light can cause electrons to be ejected from a metal surface. According to the classical wave theory of light, the intensity of the light determines the amplitude of the wave, and so a greater light intensity should cause the electrons on the metal to oscillate more violently and to be ejected with a greater kinetic energy. In contrast, the experiment showed that the kinetic energy of the ejected electrons depends on the *frequency* of the light. The light intensity affects only the number of ejected electrons and not their kinetic energies.

Einstein tackled the problem of the photoelectric effect in 1905. Instead of assuming that the electronic oscillators had energies given by Planck's formula (1), Einstein assumed that the radiation itself consisted of packets of energy $E = h\nu$, which are now called photons. Einstein successfully explained the photoelectric effect using this assumption, and he calculated a value of h close to that obtained by Planck.

Two years later, Einstein showed that not only is light quantized, but so are atomic vibrations. Classical physics predicts that the molar heat capacity at constant volume (C_v) of a crystal is $3R$, where R is the molar gas constant. This works well for high temperatures, but for low temperatures C_v actually falls to zero. Einstein was able to explain this result by assuming that the oscillations of atoms about their equilibrium positions are quantized according to $E = nh\nu$, Planck's quantization condition for electronic oscillators. This demonstrated that the energy quantization concept was important even for a system of atoms in a crystal, which should be well-modeled by a system of masses and springs (i.e., by classical mechanics).

1.3 Quantization of Electronic Angular Momentum

Rutherford proposed that electrons orbit about the nucleus of an atom. One problem with this model is that, classically, orbiting electrons experience a centripetal acceleration, and accelerating charges lose energy by radiating; a stable electronic orbit is classically forbidden. Bohr nevertheless assumed stable electronic orbits with the electronic angular momentum quantized as

$$l = mvr = n\hbar \tag{2}$$

Quantization of angular momentum means that the radius of the orbit and the energy will be quantized as well. Bohr assumed that the discrete lines seen in the spectrum of the hydrogen atom were due to transitions of an electron from one allowed orbit/energy to another. He further assumed that the energy for a transition is acquired or released in the form of a photon as proposed by Einstein, so that

$$\Delta E = h\nu \tag{3}$$

This is known as the *Bohr frequency condition.* This condition, along with Bohr's expression for the allowed energy levels, gives a good match to the observed hydrogen atom spectrum. However, it works only for atoms with one electron.

1.4 Wave-Particle Duality

Einstein had shown that the momentum of a photon is

$$p = \frac{h}{\lambda} \tag{4}$$

This can be easily shown as follows. Assuming $E = h\nu$ for a photon and $\lambda\nu = c$ for an electromagnetic wave, we obtain

$$E = \frac{hc}{\lambda} \tag{5}$$

Now we use Einstein's relativity result $E = mc^2$ to find

$$\lambda = \frac{h}{mc} \tag{6}$$

which is equivalent to equation (4). Note that m refers to the relativistic mass, not the rest mass, since the rest mass of a photon is zero. Since light can behave both as a wave (it can be diffracted, and it has a wavelength), and as a particle (it contains packets of energy $h\nu$), de Broglie reasoned in 1924 that matter also can exhibit this *wave-particle duality.* He further reasoned that matter would obey the same equation (4) as light. In 1927, Davisson and Germer observed diffraction patterns by bombarding metals with electrons, confirming de Broglie's proposition.

de Broglie's equation offers a justification for Bohr's assumption (2). If we think of an electron as a wave, then for the electron orbit to be stable the wave must complete an integral number of wavelengths during its orbit. Otherwise, it would interfere destructively with itself. This condition may be written as

$$2\pi r = n\lambda \tag{7}$$

If we use the de Broglie relation (4), this can be rewritten as

$$mvr = n\hbar \tag{8}$$

which is identical to Bohr's equation (2).

Although de Broglie's equation justifies Bohr's quantization assumption, it also demonstrates a deficiency of Bohr's model. Heisenberg showed that the wave-particle duality leads to the famous uncertainty principle

$$\Delta x \Delta p \approx h \tag{9}$$

One result of the uncertainty principle is that if the orbital radius of an electron in an atom r is known exactly, then the angular momentum must be completely unknown. The problem with Bohr's model is that it specifies r exactly and it also specifies that the orbital angular momentum must be an integral multiple of $\hbar$. Thus the stage was set for a new quantum theory which was consistent with the uncertainty principle.

2 The Schrödinger Equation

In 1925, Erwin Schrödinger and Werner Heisenberg independently developed the new quantum theory. Schrödinger's method involves partial differential equations, whereas Heisenberg's method employs matrices; however, a year later the two methods were shown to be mathematically equivalent. Most textbooks begin with Schrödinger's equation, since it seems to have a better physical interpretation via the classical wave equation. Indeed, the Schrödinger equation can be viewed as a form of the wave equation applied to matter waves.

2.1 The Time-Independent Schrödinger Equation

Here we follow the treatment of McQuarrie [1], Section 3-1. We start with the one-dimensional classical wave equation,

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} \tag{10}$$

By introducing the separation of variables

$$u(x,t) = \psi(x)f(t) \tag{11}$$

we obtain

$$f(t)\frac{d^2\psi(x)}{dx^2} = \frac{1}{v^2}\psi(x)\frac{d^2 f(t)}{dt^2} \tag{12}$$

If we introduce one of the standard wave equation solutions for $f(t)$ such as $e^{i\omega t}$ (the constant can be taken care of later in the normalization), we obtain

$$\frac{d^2\psi(x)}{dx^2} = \frac{-\omega^2}{v^2}\psi(x) \tag{13}$$

Now we have an ordinary differential equation describing the spatial amplitude of the matter wave as a function of position. The energy of a particle is the sum of kinetic and potential parts

$$E = \frac{p^2}{2m} + V(x) \tag{14}$$

which can be solved for the momentum, p, to obtain

$$p = \{2m[E - V(x)]\}^{1/2} \tag{15}$$

Now we can use the de Broglie formula (4) to get an expression for the wavelength

$$\lambda = \frac{h}{p} = \frac{h}{\{2m[E - V(x)]\}^{1/2}} \tag{16}$$

The term ω^2/v^2 in equation (13) can be rewritten in terms of λ if we recall that $\omega = 2\pi\nu$ and $\nu\lambda = v$.

$$\frac{\omega^2}{v^2} = \frac{4\pi^2\nu^2}{v^2} = \frac{4\pi^2}{\lambda^2} = \frac{2m[E - V(x)]}{\hbar^2} \tag{17}$$

When this result is substituted into equation (13) we obtain the famous *time-independent Schrödinger equation*

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}[E - V(x)]\psi(x) = 0 \tag{18}$$

which is almost always written in the form

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \tag{19}$$

This single-particle one-dimensional equation can easily be extended to the case of three dimensions, where it becomes

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{20}$$

A two-body problem can also be treated by this equation if the mass m is replaced with a reduced mass μ.

It is important to point out that this analogy with the classical wave equation only goes so far. We cannot, for instance, derive the time-*dependent* Schrödinger equation in an analogous fashion (for instance, that equation involves the partial first derivative with respect to time instead of the partial second derivative). In fact, Schrödinger presented his time-independent equation first, and then went back and postulated the more general time-dependent equation.

2.2 The Time-Dependent Schrödinger Equation

We are now ready to consider the time-dependent Schrödinger equation. Although we were able to derive the single-particle time-independent Schrödinger equation starting from the classical wave equation and the de Broglie relation, the time-dependent Schrödinger equation cannot be derived using elementary methods and is generally given as a postulate of quantum mechanics. It is possible to show that the time-dependent equation is at least *reasonable* if not derivable, but the arguments are rather involved (cf. Merzbacher [2], Section 3.2; Levine [3], Section 1.4).

The single-particle three-dimensional time-dependent Schrödinger equation is

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}, t) + V(\mathbf{r}) \psi(\mathbf{r}, t) \tag{21}$$

where V is assumed to be a real function and represents the potential energy of the system (a complex function V will act as a source or sink for probability, as shown in Merzbacher [2], problem 4.1). *Wave Mechanics* is the branch of quantum mechanics with equation (21) as its dynamical law. Note that equation (21) does not yet account for spin or relativistic effects.

Of course the time-dependent equation can be used to derive the time-independent equation. If we write the wavefunction as a product of spatial and temporal terms, $\psi(\mathbf{r}, t) = \psi(\mathbf{r}) f(t)$, then equation (21) becomes

$$\psi(\mathbf{r}) i\hbar \frac{df(t)}{dt} = f(t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) \tag{22}$$

or

$$\frac{i\hbar}{f(t)} \frac{df}{dt} = \frac{1}{\psi(\mathbf{r})} \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) \tag{23}$$

Since the left-hand side is a function of t only and the right hand side is a function of $\mathbf{r}$ only, the two sides must equal a constant. If we tentatively designate this constant E (since the right-hand side clearly must have the dimensions of energy), then we extract two ordinary differential equations, namely

$$\frac{1}{f(t)}\frac{df(t)}{dt} = -\frac{iE}{\hbar} \tag{24}$$

and

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{25}$$

The latter equation is once again the time-independent Schrödinger equation. The former equation is easily solved to yield

$$f(t) = e^{-iEt/\hbar} \tag{26}$$

The Hamiltonian in equation (25) is a Hermitian operator, and the eigenvalues of a Hermitian operator must be real, so E is real. This means that the solutions $f(t)$ are purely oscillatory, since $f(t)$ never changes in magnitude (recall Euler's formula $e^{\pm i\theta} = cos\theta \pm i\ sin\theta$). Thus if

$$\psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-iEt/\hbar} \tag{27}$$

then the total wave function $\psi(\mathbf{r},t)$ differs from $\psi(\mathbf{r})$ only by a phase factor of constant magnitude. There are some interesting consequences of this. First of all, the quantity $|\psi(\mathbf{r},t)|^2$ is time independent, as we can easily show:

$$|\psi(\mathbf{r},t)|^2 = \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) = e^{iEt/\hbar}\psi^*(\mathbf{r})e^{-iEt/\hbar}\psi(\mathbf{r}) = \psi^*(\mathbf{r})\psi(\mathbf{r}) \tag{28}$$

Secondly, the expectation value for any time-independent operator is also time-independent, if $\psi(\mathbf{r},t)$ satisfies equation (27). By the same reasoning applied above,

$$< A >= \int \psi^*(\mathbf{r},t)\hat{A}\psi(\mathbf{r},t) = \int \psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r}) \tag{29}$$

For these reasons, wave functions of the form (27) are called *stationary states.* The state $\psi(\mathbf{r},t)$ is "stationary," but the particle it describes is not!

Of course equation (27) represents a particular solution to equation (21). The general solution to equation (21) will be a linear combination of these particular solutions, i.e.

$$\psi(\mathbf{r},t) = \sum_i c_i e^{-iE_it/\hbar}\psi_i(\mathbf{r}) \tag{30}$$

3 Mathematical Background

3.1 Operators

Levine [3] defines an *operator* as “a rule that transforms a given function into another function” (p. 33). The differentation operator d/dx is an example—it transforms a differentiable function $f(x)$ into another function $f'(x)$. Other examples include integration, the square root, and so forth. Numbers can also be considered as operators (they multiply a function). McQuarrie [1] gives an even more general definition for an operator: “An *operator* is a symbol that tells you to do something with whatever follows the symbol” (p. 79). Perhaps this definition is more appropriate if we want to refer to the $\hat{C}_3$ operator acting on NH_3, for example.

3.1.1 Operators and Quantum Mechanics

In quantum mechanics, physical observables (e.g., energy, momentum, position, etc.) are represented mathematically by operators. For instance, the operator corresponding to energy is the Hamiltonian operator

$$\hat{H} = -\frac{\hbar^2}{2}\sum_i \frac{1}{m_i}\nabla_i^2 + V \tag{31}$$

where i is an index over all the particles of the system. We have already encountered the single-particle Hamiltonian in equation (25). The average value of an observable A represented by an operator $\hat{A}$ for a quantum molecular state $\psi(\mathbf{r})$ is given by the “expectation value” formula

$$< A >= \int \psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r})d\mathbf{r} \tag{32}$$

3.1.2 Basic Properties of Operators

Most of the properties of operators are obvious, but they are summarized below for completeness.

- The **sum** and **difference** of two operators $\hat{A}$ and $\hat{B}$ are given by

$$(\hat{A}+\hat{B})f = \hat{A}f+\hat{B}f \tag{33}$$

$$(\hat{A}-\hat{B})f = \hat{A}f-\hat{B}f \tag{34}$$

- The **product** of two operators is defined by

$$\hat{A}\hat{B}f \equiv \hat{A}[\hat{B}f] \tag{35}$$

- Two operators are **equal** if

$$\hat{A}f = \hat{B}f \tag{36}$$

  for all functions f.

- The **identity operator** $\hat{1}$ does nothing (or multiplies by 1)

$$\hat{1}f = f \tag{37}$$

  A common mathematical trick is to write this operator as a sum over a complete set of states (more on this later).

$$\sum_i |i\rangle\langle i|f = f \tag{38}$$

- The **associative law** holds for operators

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \tag{39}$$

- The **commutative law** does *not* generally hold for operators. In general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. It is convenient to define the quantity

$$[\hat{A},\hat{B}] \equiv \hat{A}\hat{B}-\hat{B}\hat{A} \tag{40}$$

  which is called the **commutator** of $\hat{A}$ and $\hat{B}$. Note that the order matters, so that $[\hat{A},\hat{B}] = -[\hat{B},\hat{A}]$. If $\hat{A}$ and $\hat{B}$ happen to commute, then $[\hat{A},\hat{B}] = 0$.

- The **n-th power** of an operator $\hat{A}^n$ is defined as n successive applications of the operator, e.g.

$$\hat{A}^2 f = \hat{A}\hat{A} f \tag{41}$$

- The **exponential** of an operator $e^{\hat{A}}$ is defined via the power series

$$e^{\hat{A}} = \hat{1} + \hat{A} + \frac{\hat{A}^2}{2!} + \frac{\hat{A}^3}{3!} + \cdots \tag{42}$$

3.1.3 Linear Operators

Almost all operators encountered in quantum mechanics are *linear operators*. A linear operator is an operator which satisfies the following two conditions:

$$\begin{aligned} \hat{A}(f+g) &= \hat{A}f + \hat{A}g & (43) \\ \hat{A}(cf) &= c\hat{A}f & (44) \end{aligned}$$

where c is a constant and f and g are functions. As an example, consider the operators d/dx and $()^2$. We can see that d/dx is a linear operator because

$$\begin{aligned} (d/dx)[f(x)+g(x)] &= (d/dx)f(x) + (d/dx)g(x) & (45) \\ (d/dx)[cf(x)] &= c\,(d/dx)f(x) & (46) \end{aligned}$$

However, $()^2$ is not a linear operator because

$$(f(x)+g(x))^2 \neq (f(x))^2 + (g(x))^2 \tag{47}$$

The only other category of operators relevant to quantum mechanics is the set of *antilinear* operators, for which

$$\hat{A}(\lambda f + \mu g) = \lambda^* \hat{A} f + \mu^* \hat{A} g \tag{48}$$

Time-reversal operators are antilinear (cf. Merzbacher [2], section 16-11).

3.1.4 Eigenfunctions and Eigenvalues

An *eigenfunction* of an operator $\hat{A}$ is a function f such that the application of $\hat{A}$ on f gives f again, times a constant.

$$\hat{A}f = kf \tag{49}$$

where k is a constant called the *eigenvalue.* It is easy to show that if $\hat{A}$ is a linear operator with an eigenfunction g, then any multiple of g is also an eigenfunction of $\hat{A}$.

When a system is in an *eigenstate* of observable A (i.e., when the wavefunction is an eigenfunction of the operator $\hat{A}$) then the expectation value of A is the eigenvalue of the wavefunction. Thus if

$$\hat{A}\psi(\mathbf{r}) = a\psi(\mathbf{r}) \tag{50}$$

then

$$\begin{aligned} < A > &= \int \psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r})d\mathbf{r} \\ &= \int \psi^*(\mathbf{r})a\psi(\mathbf{r})d\mathbf{r} \\ &= a\int \psi^*(\mathbf{r})\psi(\mathbf{r})d\mathbf{r} \\ &= a \end{aligned} \tag{51}$$

assuming that the wavefunction is normalized to 1, as is generally the case. In the event that $\psi(\mathbf{r})$ is not or cannot be normalized (free particle, etc.) then we may use the formula

$$< A >= \frac{\int \psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r})}{d}\mathbf{r}\int \psi^*(\mathbf{r})\psi(\mathbf{r})d\mathbf{r} \tag{52}$$

What if the wavefunction is a combination of eigenstates? Let us assume that we have a wavefunction which is a linear combination of two eigenstates of $\hat{A}$ with eigenvalues a and b.

$$\psi = c_a\psi_a + c_b\psi_b \tag{53}$$

where $\hat{A}\psi_a = a\psi_a$ and $\hat{A}\psi_b = b\psi_b$. Then what is the expectation value of A?

$$\begin{aligned} < A > &= \int \psi^* \hat{A} \psi \qquad (54) \\ &= \int [c_a\psi_a + c_b\psi_b]^* \hat{A} [c_a\psi_a + c_b\psi_b] \\ &= \int [c_a\psi_a + c_b\psi_b]^* [ac_a\psi_a + bc_b\psi_b] \\ &= a|c_a|^2 \int \psi_a^*\psi_a + bc_a^*c_b \int \psi_a^*\psi_b + ac_b^*c_a \int \psi_b^*\psi_a + b|c_b|^2 \int \psi_b^*\psi_b \\ &= a|c_a|^2 + b|c_b|^2 \end{aligned}$$

assuming that ψ_a and ψ_b are orthonormal (shortly we will show that eigenvectors of Hermitian operators are orthogonal). Thus the average value of A is a weighted average of eigenvalues, with the weights being the squares of the coefficients of the eigenvectors in the overall wavefunction.

3.1.5 Hermitian Operators

As mentioned previously, the expectation value of an operator $\hat{A}$ is given by

$$< A >= \int \psi^*(\mathbf{r}) \hat{A} \psi(\mathbf{r}) d\mathbf{r} \qquad (55)$$

and all physical observables are represented by such expectation values. Obviously, the value of a physical observable such as energy or density must be real, so we require $< A >$ to be real. This means that we must have $< A >=< A >^*$, or

$$\int \psi^*(\mathbf{r}) \hat{A} \psi(\mathbf{r}) d\mathbf{r} = \int (\hat{A}\psi(\mathbf{r}))^* \psi(\mathbf{r}) d\mathbf{r} \qquad (56)$$

Operators $\hat{A}$ which satisfy this condition are called *Hermitian*. One can also show that for a Hermitian operator,

$$\int \psi_1^*(\mathbf{r}) \hat{A} \psi_2(\mathbf{r}) d\mathbf{r} = \int (\hat{A}\psi_1(\mathbf{r}))^* \psi_2(\mathbf{r}) d\mathbf{r} \qquad (57)$$

for any two states ψ_1 and ψ_2.

An important property of Hermitian operators is that their eigenvalues are real. We can see this as follows: if we have an eigenfunction of $\hat{A}$ with eigenvalue

a, i.e. $\hat{A}\psi_a = a\psi_a$, then for a Hermitian operator $\hat{A}$

$$\begin{aligned} \int \psi_a^* \hat{A} \psi_a &= \int \psi_a (\hat{A}\psi_a)^* \\ a \int \psi_a^* \psi_a &= a^* \int \psi_a \psi_a^* \\ (a - a^*) \int |\psi_a|^2 &= 0 \end{aligned} \tag{58}$$

Since $|\psi_a|^2$ is never negative, we must have either $a = a^*$ or $\psi_a = 0$. Since $\psi_a = 0$ is not an acceptable wavefunction, $a = a^*$, so a is real.

Another important property of Hermitian operators is that their eigenvectors are orthogonal (or can be chosen to be so). Suppose that ψ_a and ψ_b are eigenfunctions of $\hat{A}$ with eigenvalues a and b, with $a \neq b$. If $\hat{A}$ is Hermitian then

$$\begin{aligned} \int \psi_a^* \hat{A} \psi_b &= \int \psi_b (\hat{A}\psi_a)^* \\ b \int \psi_a^* \psi_b &= a^* \int \psi_b \psi_a^* \\ (b - a) \int \psi_a^* \psi_b &= 0 \end{aligned} \tag{59}$$

since $a = a^*$ as shown above. Because we assumed $b \neq a$, we must have $\int \psi_a^* \psi_b = 0$, i.e. ψ_a and ψ_b are orthogonal. Thus we have shown that eigenfunctions of a Hermitian operator with different eigenvalues are orthogonal. In the case of degeneracy (more than one eigenfunction with the same eigenvalue), we can *choose* the eigenfunctions to be orthogonal. We can easily show this for the case of two eigenfunctions of $\hat{A}$ with the same eigenvalue. Suppose we have

$$\begin{aligned} \hat{A}\psi_j &= j\psi_j \\ \hat{A}\psi_k &= j\psi_k \end{aligned} \tag{60}$$

We now want to take linear combinations of ψ_j and ψ_k to form two new eigenfunctions $\psi_{j'}$ and $\psi_{k'}$, where $\psi_{j'} = \psi_j$ and $\psi_{k'} = \psi_k + c\psi_j$. Now we want $\psi_{j'}$ and $\psi_{k'}$ to be orthogonal, so

$$\begin{aligned} \int \psi_{j'}^* \psi_{k'} &= 0 \\ \int \psi_j^* (\psi_k + c\psi_j) &= 0 \\ \int \psi_j^* \psi_k + c \int \psi_j^* \psi_j &= 0 \end{aligned} \tag{61}$$

Thus we merely need to choose

$$c = -\frac{\int \psi_j^* \psi_k}{\int \psi_j^* \psi_j} \tag{62}$$

and we obtain orthogonal eigenfunctions. This Schmidt-orthogonalization procedure can be extended to the case of n-fold degeneracy, so we have shown that for a Hermitian operator, the eigenvectors can be made orthogonal.

3.1.6 Unitary Operators

A linear operator whose inverse is its adjoint is called *unitary.* These operators can be thought of as generalizations of complex numbers whose absolue value is 1.

$$\begin{aligned} U^{-1} &= U^\dagger \\ UU^\dagger = U^\dagger U &= I \end{aligned} \tag{63}$$

A unitary operator preserves the "lengths" and "angles" between vectors, and it can be considered as a type of rotation operator in abstract vector space. Like Hermitian operators, the eigenvectors of a unitary matrix are orthogonal. However, its eigenvalues are not necessarily real.

3.2 Commutators in Quantum Mechanics

The *commutator*, defined in section 3.1.2, is very important in quantum mechanics. Since a definite value of observable A can be assigned to a system only if the system is in an eigenstate of $\hat{A}$, then we can simultaneously assign definite values to two observables A and B only if the system is in an eigenstate of both $\hat{A}$ and $\hat{B}$. Suppose the system has a value of A_i for observable A and B_j for observable B. The we require

$$\begin{aligned} \hat{A}\psi_{A_i,B_j} &= A_i \psi_{A_i,B_j} \\ \hat{B}\psi_{A_i,B_j} &= B_j \psi_{A_i,B_j} \end{aligned} \tag{64}$$

If we multiply the first equation by $\hat{B}$ and the second by $\hat{A}$ then we obtain

$$\begin{aligned} \hat{B}\hat{A}\psi_{A_i,B_j} &= \hat{B}A_i\psi_{A_i,B_j} \\ \hat{A}\hat{B}\psi_{A_i,B_j} &= \hat{A}B_j\psi_{A_i,B_j} \end{aligned} \tag{65}$$

and, using the fact that ψ_{A_i,B_j} is an eigenfunction of $\hat{A}$ and $\hat{B}$, this becomes

$$\begin{aligned} \hat{B}\hat{A}\psi_{A_i,B_j} &= A_iB_j\psi_{A_i,B_j} \\ \hat{A}\hat{B}\psi_{A_i,B_j} &= B_jA_i\psi_{A_i,B_j} \end{aligned} \tag{66}$$

so that if we subtract the first equation from the second, we obtain

$$(\hat{A}\hat{B} - \hat{B}\hat{A})\psi_{A_i,B_j} = 0 \tag{67}$$

For this to hold for general eigenfunctions, we must have $\hat{A}\hat{B} = \hat{B}\hat{A}$, or $[\hat{A}, \hat{B}] = 0$. That is, for two physical quantities to be simultaneously observable, their operator representations must commute.

Section 8.8 of Merzbacher [2] contains some useful rules for evaluating commutators. They are summarized below.

$$[\hat{A}, \hat{B}] + [\hat{B}, \hat{A}] = 0 \tag{68}$$

$$[\hat{A}, \hat{A}] = 0 \tag{69}$$

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \tag{70}$$

$$[\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}] \tag{71}$$

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}] \tag{72}$$

$$[\hat{A}\hat{B}, \hat{C}] = [\hat{A}, \hat{C}]\hat{B} + \hat{A}[\hat{B}, \hat{C}] \tag{73}$$

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0 \tag{74}$$

If $\hat{A}$ and $\hat{B}$ are two operators which commute with their commutator, then

$$[\hat{A}, \hat{B}^n] = n\hat{B}^{n-1}[\hat{A}, \hat{B}] \tag{75}$$

$$[\hat{A}^n, \hat{B}] = n\hat{A}^{n-1}[\hat{A}, \hat{B}] \tag{76}$$

We also have the identity (useful for coupled-cluster theory)

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \cdots \tag{77}$$

Finally, if $[\hat{A}, \hat{B}] = i\hat{C}$ then the uncertainties in A and B, defined as $\Delta A^2 =< A^2 > - < A >^2$, obey the relation[1]

$$(\Delta A)(\Delta B) \geq \frac{1}{2}| < C > | \tag{78}$$

This is the famous Heisenberg uncertainty principle. It is easy to derive the well-known relation

$$(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2} \tag{79}$$

from this generalized rule.

3.3 Linear Vector Spaces in Quantum Mechanics

We have observed that most operators in quantum mechanics are linear operators. This is fortunate because it allows us to represent quantum mechanical operators as matrices and wavefunctions as vectors in some linear vector space. Since computers are particularly good at performing operations common in linear algebra (multiplication of a matrix times a vector, etc.), this is quite advantageous from a practical standpoint.

In an n-dimensional space we may expand any vector Ψ as a linear combination of basis vectors

$$\Psi = \sum_{i=1}^{n} a_i \Psi_i \tag{80}$$

For a general vector space, the coefficients a_i may be complex; thus one should not be too quick to draw parallels to the expansion of vectors in three-dimensional Euclidean space. The coefficients a_i are referred to as the "components" of the state vector Ψ, and for a given basis, the components of a vector specify it completely.

---

[1] Assuming that the quantum covariance $< (\hat{A}\hat{B} + \hat{B}\hat{A})/2 - < \hat{A} >< \hat{B} >>$ is zero.

The components of the sum of two vectors are the sums of the components. If $\Psi_a = \sum a_i \Psi_i$ and $\Psi_b = \sum b_i \Psi_i$ then

$$\Psi_a + \Psi_b = \sum_i (a_i + b_i)\Psi_i \tag{81}$$

and similarly

$$\lambda \Psi_a = \sum_i (\lambda a_i)\Psi_i \tag{82}$$

The *scalar product* of two vectors is a complex number denoted by

$$(\Psi_b, \Psi_a) = (\Psi_a, \Psi_b)^* \tag{83}$$

where we have used the standard linear-algebra notation. If we also require that

$$(\Psi_a, \lambda \Psi_b) = \lambda (\Psi_a, \Psi_b) \tag{84}$$

then it follows that

$$(\lambda \Psi_a, \Psi_b) = \lambda^* (\Psi_a, \Psi_b) \tag{85}$$

We also require that

$$(\Psi_a, \Psi_b + \Psi_c) = (\Psi_a, \Psi_b) + (\Psi_a, \Psi_c) \tag{86}$$

If the scalar product vanishes (and if neither vector in the product is the null vector) then the two vectors are orthogonal.

Generally the basis is chosen to be orthonormal, such that

$$(\hat{\Psi}_i, \hat{\Psi}_j) = \delta_{ij} \tag{87}$$

In this case, we can write the scalar product of two arbitrary vectors as

$$\begin{aligned} (\Psi_a, \Psi_b) &= (\sum_i a_i \hat{\Psi}_i, \sum_j b_j \hat{\Psi}_j) \\ &= \sum_i \sum_j a_i^* b_j (\hat{\Psi}_i, \hat{\Psi}_j) \\ &= \sum_i a_i^* b_i \end{aligned} \tag{88}$$

This can also be written in vector notation as

$$(\Psi_a, \Psi_b) = (a_1^* a_2^* \cdots a_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{89}$$

It is useful at this point to introduce Dirac's bra-ket notation. We define a "bra" as

$$\langle \Psi_a | = (a_1^* a_2^* \cdots a_n^*) \tag{90}$$

and a "ket" as

$$| \Psi_a \rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{91}$$

A bra to the left of a ket implies a scalar product, so

$$\langle \Psi_a | \Psi_b \rangle = (\Psi_a, \Psi_b) \tag{92}$$

Sometimes in superficial treatments of Dirac notation, the symbol $\langle \Psi_a | \Psi_b \rangle$ is defined alternatively as

$$\langle \Psi_a | \Psi_b \rangle = \int \Psi_a^*(x) \Psi_b(x) dx \tag{93}$$

This is equivalent to the above definition if we make the connections $a_i = \Psi_a(x)$ and $b_i = \Psi_b(x)$. This means that our basis vectors are *every possible value of* x. Since x is continuous, the sum is replaced by an integral (see Szabo and Ostlund [4] , exercise 1.17). Often only the subscript of the vector is used to denote a bra or ket; we may have written the above equation as

$$\langle a | b \rangle = \int \Psi_a^*(x) \Psi_b(x) dx \tag{94}$$

Now we turn our attention to matrix representations of operators. An operator $\hat{A}$ can be characterized by its effect on the basis vectors. The action of $\hat{A}$ on a basis vector $\hat{\Psi}_j$ yields some new vector Ψ'_j which can be expanded in terms of the basis vectors so long as we have a complete basis set.

$$\hat{A}\hat{\Psi}_j = \Psi'_j = \sum_i^n \hat{\Psi}_i A_{ij} \tag{95}$$

If we know the effect of $\hat{A}$ on the basis vectors, then we know the effect of $\hat{A}$ on any arbitrary vector because of the linearity of $\hat{A}$.

$$\begin{aligned} \Psi_b = \hat{A}\Psi_a = \hat{A}\sum_j a_j \hat{\Psi}_j &= \sum_j a_j \hat{A}\hat{\Psi}_j = \sum_j \sum_i a_j \hat{\Psi}_i A_{ij} \\ &= \sum_i \hat{\Psi}_i (\sum_j A_{ij} a_j) \end{aligned} \tag{96}$$

or

$$b_i = \sum_j A_{ij} a_j \tag{97}$$

This may be written in matrix notation as

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{98}$$

We can obtain the coefficients A_{ij} by taking the inner product of both sides of equation 95 with $\hat{\Psi}_i$, yielding

$$\begin{aligned} (\hat{\Psi}_i, \hat{A}\hat{\Psi}_j) &= (\hat{\Psi}_i, \sum_k^n \hat{\Psi}_k A_{kj}) \\ &= \sum_k^n A_{kj} (\hat{\Psi}_i, \hat{\Psi}_k) \\ &= A_{ij} \end{aligned} \tag{99}$$

since $(\hat{\Psi}_i, \hat{\Psi}_k) = \delta_{ik}$ due to the orthonormality of the basis. In bra-ket notation, we may write

$$A_{ij} = \langle i|\hat{A}|j\rangle \tag{100}$$

where i and j denote two basis vectors. This use of bra-ket notation is consistent with its earlier use if we realize that $\hat{A}|j\rangle$ is just another vector $|j'\rangle$.

It is easy to show that for a linear operator $\hat{A}$, the inner product $(\Psi_a, \hat{A}\Psi_b)$ for two general vectors (not necessarily basis vectors) Ψ_a and Ψ_b is given by

$$(\Psi_a, \hat{A}\Psi_b) = \sum_i \sum_j a_i^* A_{ij} b_j \tag{101}$$

or in matrix notation

$$(\Psi_a, \hat{A}\Psi_b) = (a_1^* a_2^* \cdots a_n^*) \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{102}$$

By analogy to equation (93), we may generally write this inner product in the form

$$(\Psi_a, \hat{A}\Psi_b) = \langle a|\hat{A}|b\rangle = \int \Psi_a^*(x)\hat{A}\Psi_b(x)dx \tag{103}$$

Previously, we noted that $(\Psi_a, \Psi_b) = (\Psi_b, \Psi_a)^*$, or $\langle a|b\rangle = \langle b|a\rangle^*$. Thus we can see also that

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}\Psi_b, \Psi_a)^* \tag{104}$$

We now define the *adjoint* of an operator $\hat{A}$, denoted by $\hat{A}^\dagger$, as that linear operator for which

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}^\dagger\Psi_a, \Psi_b) \tag{105}$$

That is, we can make an operator act *backwards* into "bra" space if we take it's adjoint. With this definition, we can further see that

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}\Psi_b, \Psi_a)^* = (\Psi_b, \hat{A}^\dagger\Psi_a)^* = (\hat{A}^\dagger\Psi_a, \Psi_b) \tag{106}$$

or, in bra-ket notation,

$$\langle a|\hat{A}|b\rangle = \langle \hat{A}b|a\rangle^* = \langle b|\hat{A}^\dagger|a\rangle^* = \langle \hat{A}^\dagger a|b\rangle \tag{107}$$

If we pick $\Psi_a = \hat{\Psi}_i$ and $\Psi_b = \hat{\Psi}_j$ (i.e., if we pick two basis vectors), then we obtain

$$\begin{aligned} (\hat{A}\hat{\Psi}_i, \hat{\Psi}_j) &= (\hat{\Psi}_i, \hat{A}^\dagger \hat{\Psi}_j) \\ (\hat{\Psi}_j, \hat{A}\hat{\Psi}_i)^* &= (\hat{\Psi}_i, \hat{A}^\dagger \hat{\Psi}_j) \\ A_{ji}^* &= A_{ij}^\dagger \end{aligned} \tag{108}$$

But this is precisely the condition for the elements of a matrix and its adjoint! Thus the adjoint of the matrix representation of $\hat{A}$ is the same as the matrix representation of $\hat{A}^\dagger$.

This correspondence between operators and their matrix representations goes quite far, although of course the specific matrix representation depends on the choice of basis. For instance, we know from linear algebra that if a matrix and its adjoint are the same, then the matrix is called Hermitian. The same is true of the operators; if

$$\hat{A} = \hat{A}^\dagger \tag{109}$$

then $\hat{A}$ is a Hermitian operator, and all of the special properties of Hermitian operators apply to $\hat{A}$ or its matrix representation.

4 Postulates of Quantum Mechanics

In this section, we will present six postulates of quantum mechanics. Again, we follow the presentation of McQuarrie [1], with the exception of postulate 6, which McQuarrie does not include. A few of the postulates have already been discussed in section 3.

> **Postulate 1**. The state of a quantum mechanical system is completely specified by a function $\Psi(\mathbf{r}, t)$ that depends on the coordinates of the particle(s) and on time. This function, called the wave function or state function, has the important property that $\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)d\tau$ is the probability that the particle lies in the volume element $d\tau$ located at $\mathbf{r}$ at time t.

The wavefunction must satisfy certain mathematical conditions because of this probabilistic interpretation. For the case of a single particle, the probability of finding it *somewhere* is 1, so that we have the normalization condition

$$\int_{-\infty}^{\infty} \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)d\tau = 1 \tag{110}$$

It is customary to also normalize many-particle wavefunctions to 1.[2] The wavefunction must also be single-valued, continuous, and finite.

> **Postulate 2**. To every observable in classical mechanics there corresponds a linear, Hermitian operator in quantum mechanics.

This postulate comes about because of the considerations raised in section 3.1.5: if we require that the expectation value of an operator $\hat{A}$ is real, then $\hat{A}$ must be a Hermitian operator. Some common operators occuring in quantum mechanics are collected in Table 1.

---

[2] In some cases, such as the free-particle, one must use special tricks to normalize the wavefunction. See Merzbacher [2], section 8.1.

Table 1: Physical observables and their corresponding quantum operators (single particle)

Observable Name	Observable Symbol	Operator Symbol	Operator Operation
Position	$\mathbf{r}$	$\hat{\mathbf{r}}$	Multiply by $\mathbf{r}$
Momentum	$\mathbf{p}$	$\hat{\mathbf{p}}$	$-i\hbar\left(\hat{i}\frac{\partial}{\partial x}+\hat{j}\frac{\partial}{\partial y}+\hat{k}\frac{\partial}{\partial z}\right)$
Kinetic energy	T	$\hat{T}$	$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)$
Potential energy	$V(\mathbf{r})$	$\hat{V}(\mathbf{r})$	Multiply by $V(\mathbf{r})$
Total energy	E	$\hat{H}$	$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(\mathbf{r})$
Angular momentum	l_x	$\hat{l}_x$	$-i\hbar\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)$
	l_y	$\hat{l}_y$	$-i\hbar\left(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z}\right)$
	l_z	$\hat{l}_z$	$-i\hbar\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)$

Postulate 3. In any measurement of the observable associated with operator $\hat{A}$, the only values that will ever be observed are the eigenvalues a, which satisfy the eigenvalue equation

$$\hat{A}\Psi = a\Psi \tag{111}$$

This postulate captures the central point of quantum mechanics—the values of dynamical variables can be quantized (although it is still possible to have a continuum of eigenvalues in the case of unbound states). If the system is in an eigenstate of $\hat{A}$ with eigenvalue a, then any measurement of the quantity A will yield a.

Although measurements must always yield an eigenvalue, the state does not have to be an eigenstate of $\hat{A}$ *initially*. An arbitrary state can be expanded in the complete set of eigenvectors of $\hat{A}$ ($\hat{A}\Psi_i = a_i\Psi_i$) as

$$\Psi = \sum_i^n c_i\Psi_i \tag{112}$$

where n may go to infinity. In this case we only know that the measurement of A will yield *one* of the values a_i, but we don't know which one. However, we do know the *probability* that eigenvalue a_i will occur—it is the absolute value squared of the coefficient, $|c_i|^2$ (cf. section 3.1.4), leading to the fourth postulate below.

An important second half of the third postulate is that, after measurement of Ψ yields some eigenvalue a_i, the wavefunction immediately "collapses" into the corresponding eigenstate Ψ_i (in the case that a_i is degenerate, then Ψ becomes the projection of Ψ onto the degenerate subspace). Thus, measurement affects the state of the system. This fact is used in many elaborate experimental tests of quantum mechanics.

> **Postulate 4**. If a system is in a state described by a normalized wave function Ψ, then the average value of the observable corresponding to $\hat{A}$ is given by

$$< A >= \int_{-\infty}^{\infty} \Psi^* \hat{A} \Psi d\tau \tag{113}$$

> **Postulate 5**. The wavefunction or state function of a system evolves in time according to the time-dependent Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}, t) = i\hbar \frac{\partial \Psi}{\partial t} \tag{114}$$

The central equation of quantum mechanics must be accepted as a postulate, as discussed in section 2.2.

> **Postulate 6**. The total wavefunction must be antisymmetric with respect to the interchange of all coordinates of one fermion with those of another. Electronic spin must be included in this set of coordinates.

The Pauli exclusion principle is a direct result of this *antisymmetry principle.* We will later see that Slater determinants provide a convenient means of enforcing this property on electronic wavefunctions.

5 Some Analytically Soluble Problems

Quantum chemists are generally concerned with solving the time-independent Schrödinger equation (25). This equation can be solved analytically only in a few special cases. In this section we review the results of some of these analytically soluble problems.

5.1 The Particle in a Box

Consider a particle constrained to move in a single dimension, under the influence of a potential $V(x)$ which is zero for $0 \leq x \leq a$ and infinite elsewhere. Since the wavefunction is not allowed to become infinite, it must have a value of zero where $V(x)$ is infinite, so $\psi(x)$ is nonzero only within $[0, a]$. The Schrödinger equation is thus

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi(x) \quad 0 \leq x \leq a \tag{115}$$

It is easy to show that the eigenvectors and eigenvalues of this problem are

$$\psi_n(x) = \sqrt{\frac{2}{a}} sin\left(\frac{n\pi x}{a}\right) \qquad 0 \leq x \leq a \qquad n = 1, 2, 3, \ldots \tag{116}$$

$$E_n = \frac{h^2 n^2}{8ma^2} \qquad n = 1, 2, \ldots \tag{117}$$

Extending the problem to three dimensions is rather straightforward; see McQuarrie [1], section 6.1.

5.2 The Harmonic Oscillator

Now consider a particle subject to a restoring force $F = -kx$, as might arise for a mass-spring system obeying Hooke's Law. The potential is then

$$\begin{aligned} V(x) &= -\int_{-\infty}^{\infty}(-kx)dx \\ &= V_0 + \frac{1}{2}kx^2 \end{aligned} \tag{118}$$

If we choose the energy scale such that $V_0 = 0$ then $V(x) = (1/2)kx^2$. This potential is also appropriate for describing the interaction of two masses connected by an ideal spring. In this case, we let x be the distance between the masses, and for the mass m we substitute the reduced mass μ. Thus the harmonic oscillator is the simplest model for the vibrational motion of the atoms in a diatomic molecule, if we consider the two atoms as point masses and the bond between them as a spring. The one-dimensional Schrödinger equation becomes

$$-\frac{\hbar^2}{2\mu}\frac{d^2\psi}{dx^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x) \tag{119}$$

After some effort, the eigenfunctions are

$$\psi_n(x) = N_n H_n(\alpha^{1/2}x)e^{-\alpha x^2/2} \qquad n = 0, 1, 2, \ldots \tag{120}$$

where H_n is the Hermite polynomial of degree n, and α and N_n are defined by

$$\alpha = \sqrt{\frac{k\mu}{\hbar^2}} \qquad N_n = \frac{1}{\sqrt{2^n n!}}\left(\frac{\alpha}{\pi}\right)^{1/4} \tag{121}$$

The eigenvalues are

$$E_n = \hbar\omega(n + 1/2) \tag{122}$$

with $\omega = \sqrt{k/\mu}$.

5.3 The Rigid Rotor

The rigid rotor is a simple model of a rotating diatomic molecule. We consider the diatomic to consist of two point masses at a fixed internuclear distance. We then reduce the model to a one-dimensional system by considering the rigid rotor to have one mass fixed at the origin, which is orbited by the reduced mass μ, at a distance r. The Schrödinger equation is (cf. McQuarrie [1], section 6.4 for a clear explanation)

$$-\frac{\hbar^2}{2I}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi(r) = E\psi(r) \tag{123}$$

After a little effort, the eigenfunctions can be shown to be the spherical harmonics $Y_J^M(\theta, \phi)$, defined by

$$Y_J^M(\theta, \phi) = \left[\frac{(2J+1)}{4\pi} \frac{(J-|M|)!}{(J+|M|)!} \right]^{1/2} P_J^{|M|}(cos\theta) e^{iM\phi} \tag{124}$$

where $P_J^{|M|}(x)$ are the associated Legendre functions. The eigenvalues are simply

$$E_J = \frac{\hbar^2}{2I} J(J+1) \tag{125}$$

Each energy level E_J is $2J+1$-fold degenerate in M, since M can have values $-J, -J+1, \ldots, J-1, J$.

5.4 The Hydrogen Atom

Finally, consider the hydrogen atom as a proton fixed at the origin, orbited by an electron of reduced mass μ. The potential due to electrostatic attraction is

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r} \tag{126}$$

in SI units. The kinetic energy term in the Hamiltonian is

$$\hat{T} = -\frac{\hbar^2}{2\mu} \nabla^2 \tag{127}$$

so we write out the Schrödinger equation in spherical polar coordinates as

$$-\frac{\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) \frac{1}{r^2 sin\theta} \frac{\partial}{\partial \theta} \left(sin\theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 sin^2\theta} \frac{\partial^2 \psi}{\partial \phi^2} \right] - \frac{e^2}{4\pi\epsilon_0 r} \psi(r, \theta, \phi) = E\psi(r, \theta, \phi) \tag{128}$$

It happens that we can factor $\psi(r, \theta, \phi)$ into $R(r)_{nl} Y_l^m(\theta, \phi)$, where $Y_l^m(\theta, \phi)$ are again the spherical harmonics. The radial part $R(r)$ then can be shown to obey the equation

$$-\frac{\hbar^2}{2\mu r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) - E \right] R(r) = 0 \tag{129}$$

which is called the *radial equation* for the hydrogen atom. Its (messy) solutions are

$$R_{nl}(r) = -\left[\frac{(n-l-1)!}{2n[(n+l)!]^3}\right]^{1/2}\left(\frac{2}{na_0}\right)^{l+3/2} r^l e^{-r/na_0} L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right) \tag{130}$$

where $0 \leq l \leq n-1$, and a_0 is the Bohr radius, $\epsilon_0 h^2/\pi\mu e^2$. The functions $L_{n+l}^{2l+1}(2r/na_0)$ are the associated Laguerre functions. The hydrogen atom eigenvalues are

$$E_n = -\frac{e^2}{8\pi\epsilon_0 a_0 n^2} \qquad n = 1, 2, \ldots \tag{131}$$

There are relatively few other interesting problems that can be solved analytically. For molecular systems, one must resort to approximate solutions.

6 Approximate Methods

The problems discussed in the previous section (harmonic oscillator, rigid rotator, etc.) are some of the few quantum mechanics problems which can be solved analytically. For the vast majority of chemical applications, the Schrödinger equation must be solved by approximate methods. The two primary approximation techniques are the variational method and perturbation theory.

6.1 Perturbation Theory

The basic idea of perturbation theory is very simple: we split the Hamiltonian into a piece we know how to solve (the "reference" or "unperturbed" Hamiltonian) and a piece we don't know how to solve (the "perturbation"). As long as the perburbation is small compared to the unperturbed Hamiltonian, perturbation theory tells us how to correct the solutions to the unperturbed problem to approximately account for the influence of the perturbation. For example, perturbation theory can be used to approximately solve an anharmonic oscillator problem with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3. \tag{132}$$

Here, since we know how to solve the harmonic oscillator problem (see 5.2), we make that part the unperturbed Hamiltonian (denoted $\hat{H}^{(0)}$), and the new, anharmonic term is the perturbation (denoted $\hat{H}^{(1)}$):

$$\hat{H}^{(0)} = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2, \tag{133}$$

$$\hat{H}^{(1)} = +\frac{1}{6}\gamma x^3. \tag{134}$$

Perturbation theory solves such a problem in two steps. First, obtain the eigenfunctions and eigenvalues of the unperturbed Hamiltonian, $\hat{H}^{(0)}$:

$$\hat{H}^{(0)}\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)}. \tag{135}$$

Second, correct these eigenvalues and/or eigenfunctions to account for the perturbation's influence. Perturbation theory gives these corrections as an infinite series of terms, which become smaller and smaller for well-behaved systems:

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots \tag{136}$$

$$\Psi_n = \Psi_n^{(0)} + \Psi_n^{(1)} + \Psi_n^{(2)} + \cdots \tag{137}$$

Quite frequently, the corrections are only taken through first or second order (i.e., superscripts (1) or (2)). According to perturbation theory, the first-order correction to the energy is

$$E_n^{(1)} = \int \Psi_n^{(0)*} \hat{H}^{(1)} \Psi_n^{(0)}, \tag{138}$$

and the second-order correction is

$$E_n^{(2)} = \int \Psi_n^{(0)*} \hat{H}^{(1)} \Psi_n^{(1)}. \tag{139}$$

One can see that the first-order correction to the wavefunction, $\Psi_n^{(1)}$, seems to be needed to compute the second-order energy correction. However, it turns out that the correction $\Psi_n^{(1)}$ can be written in terms of the zeroth-order wavefunction as

$$\Psi_n^{(1)} = \sum_{i \neq n} \Psi_i^{(0)} \frac{\int \Psi_i^{(0)*} \hat{H}^{(1)} \Psi_n^{(0)}}{E_n^{(0)} - E_i^{(0)}}. \tag{140}$$

Substituting this in the expression for $E_n^{(2)}$, we obtain

$$E_n^{(2)} = \sum_{i \neq n} \frac{| \int \Psi_n^{(0)*} \hat{H}^{(1)} \Psi_i^{(0)} |^2}{E_n^{(0)} - E_i^{(0)}}. \tag{141}$$

Going back to the anharmonic oscillator example, the ground state wavefunction for the unperturbed problem is just (from section 5.2)

$$E_0^{(0)} = \frac{1}{2}\hbar\omega, \tag{142}$$

$$\Psi_0^{(0)}(x) = N_0 H_0(\alpha^{1/2} x) e^{-\alpha x^2/2} \tag{143}$$

$$= \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2/2}. \tag{144}$$

The first-order correction to the ground state energy would be

$$E_0^{(1)} = \left(\frac{\alpha}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} \frac{1}{6}\gamma x^3 e^{-\alpha x^2} dx. \tag{145}$$

It turns out in this case that $E_0^{(1)} = 0$, since the integrand is odd. Does this mean that the anharmonic energy levels are the same as for the harmonic oscillator? No, because there are higher-order corrections such as $E_0^{(2)}$ which are not necessarily zero.

6.2 The Variational Method

The variational method is the other main approximate method used in quantum mechanics. Compared to perturbation theory, the variational method can be more robust in situations where it's hard to determine a good unperturbed Hamiltonian (i.e., one which makes the perturbation small but is still solvable). On the other hand, in cases where there is a good unperturbed Hamiltonian, perturbation theory can be more efficient than the variational method.

The basic idea of the variational method is to guess a "trial" wavefunction for the problem, which consists of some adjustable parameters called "variational parameters." These parameters are adjusted until the energy of the trial wavefunction is minimized. The resulting trial wavefunction and its corresponding energy are variational method approximations to the exact wavefunction and energy.

Why would it make sense that the best approximate trial wavefunction is the one with the lowest energy? This results from the Variational Theorem, which states that the energy of any trial wavefunction E is always an upper bound to the exact ground state energy $\mathcal{E}_0$. This can be proven easily. Let the trial wavefunction be denoted Φ. Any trial function can formally be expanded as a linear combination of the exact eigenfunctions Ψ_i. Of course, in practice, we don't know the Ψ_i, since we're assuming that we're applying the variational method to a problem we can't solve analytically. Nevertheless, that doesn't prevent us from using the exact eigenfunctions in our proof, since they certainly exist and form a complete set, even if we don't happen to know them. So, the trial wavefunction can be written

$$\Phi = \sum_i c_i \Psi_i, \tag{146}$$

and the approximate energy corresponding to this wavefunction is

$$E[\Phi] = \frac{\int \Phi^* \hat{H} \Phi}{\int \Phi^* \Phi}. \tag{147}$$

Substituting the expansion over the exact wavefuntions,

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j \int \Psi_i^* \hat{H} \Psi_j}{\sum_{ij} c_i^* c_j \int \Psi_i^* \Psi_j}. \tag{148}$$

Since the functions Ψ_j are the exact eigenfunctions of $\hat{H}$, we can use $\hat{H}\Psi_j = \mathcal{E}_j \Psi_j$ to obtain

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j \mathcal{E}_j \int \Psi_i^* \Psi_j}{\sum_{ij} c_i^* c_j \int \Psi_i^* \Psi_j}. \tag{149}$$

Now using the fact that eigenfunctions of a Hermitian operator form an orthonormal set (or can be made to do so),

$$E[\Phi] = \frac{\sum_i c_i^* c_i \mathcal{E}_i}{\sum_i c_i^* c_i}. \tag{150}$$

We now subtract the exact ground state energy $\mathcal{E}_0$ from both sides to obtain

$$E[\Phi] - \mathcal{E}_0 = \frac{\sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0)}{\sum_i c_i^* c_i}. \tag{151}$$

Since every term on the right-hand side is greater than or equal to zero, the left-hand side must also be greater than or equal to zero, or

$$E[\Phi] \geq \mathcal{E}_0. \tag{152}$$

In other words, the energy of any approximate wavefunction is always greater than or equal to the exact ground state energy $\mathcal{E}_0$. This explains the strategy of the variational method: since the energy of any approximate trial function is always above the true energy, then any variations in the trial function which lower its energy are necessarily making the approximate energy closer to the exact answer. (The trial wavefunction is also a better approximation to the true ground state wavefunction as the energy is lowered, although not necessarily in every possible sense unless the limit $\Phi = \Psi_0$ is reached).

One example of the variational method would be using the Gaussian function $\phi(r) = e^{-\alpha r^2}$ as a trial function for the hydrogen atom ground state. This problem could be solved by the variational method by obtaining the energy of $\phi(r)$ as a function of the variational parameter α, and then minimizing $E(\alpha)$ to find the optimum value α_{min}. The variational theorem's approximate wavefunction and energy for the hydrogen atom would then be $\phi(r) = e^{-\alpha_{min} r^2}$ and $E(\alpha_{min})$.

Frequently, the trial function is written as a linear combination of basis functions, such as

$$\Phi = \sum_i c_i \phi_i. \tag{153}$$

This leads to the *linear variation method*, and the variational parameters are the expansion coefficients c_i. The energy for this approximate wavefunction is just

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j \int \phi_i^* \hat{H} \phi_j}{\sum_{ij} c_i^* c_j \int \phi_i^* \phi_j}, \tag{154}$$

which can be simplified using the notation

$$\begin{aligned} H_{ij} &= \int \phi_i^* \hat{H} \phi_j, & (155) \\ S_{ij} &= \int \phi_i^* \phi_j, & (156) \end{aligned}$$

to yield

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j H_{ij}}{\sum_{ij} c_i^* c_j S_{ij}}. \tag{157}$$

Differentiating this energy with respect to the expansion coefficients c_i yields a non-trivial solution only if the following "secular determinant" equals 0.

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0. \tag{158}$$

If an orthonormal basis is used, the secular equation is greatly simplified because S_{ij} is 1 for $i = j$ and 0 for $i \neq j$. In this case, the secular determinant is

$$\begin{vmatrix} H_{11} - E & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} - E & \cdots & H_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} - E \end{vmatrix} = 0. \tag{159}$$

In either case, the secular determinant for N basis functions gives an N-th order polynomial in E which is solved for N different roots, each of which approximates a different eigenvalue.

The variational method lies behind Hartree-Fock theory and the configuration interaction method for the electronic structure of atoms and molecules.

7 Molecular Quantum Mechanics

In this section, we discuss the quantum mechanics of atomic and molecular systems. We begin by writing the Hamiltonian for a collection of nuclei and electrons, and then we introduce the Born-Oppenheimer approximation, which allows us to separate the nuclear and electronic degrees of freedom.

7.1 The Molecular Hamiltonian

We have noted before that the kinetic energy for a system of particles is

$$\hat{T} = -\frac{\hbar^2}{2}\sum_i \frac{1}{m_i}\nabla^2 \tag{160}$$

The potential energy for a system of charged particles is

$$\hat{V}(\mathbf{r}) = \sum_{i>j} \frac{Z_i Z_j e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{161}$$

For a molecule, it is reasonable to split the kinetic energy into two summations—one over electrons, and one over nuclei. Similarly, we can split the potential energy into terms representing interactions between nuclei, between electrons, or between electrons and nuclei. Using i and j to index electrons, and A and B to index nuclei, we have (in atomic units)

$$\hat{H} = -\sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_i \frac{1}{2}\nabla_i^2 + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} - \sum_{Ai} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}} \tag{162}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $R_{Ai} = |\mathbf{r}_A - \mathbf{r}_i|$, and $R_{AB} = |\mathbf{r}_A - \mathbf{r}_B|$. This is known as the "exact" nonrelativistic Hamiltonian in field-free space. However, it is important to remember that this Hamiltonian neglects at least two effects. Firstly, although the speed of an electron in a hydrogen atom is less than 1% of the speed of light, relativistic mass corrections can become appreciable for the inner electrons of heavier atoms. Secondly, we have neglected the spin-orbit effects. From the point of view of an electron, it is being orbited by a nucleus which produces a magnetic field (proportional to L); this field interacts with the electron's magnetic moment (proportional to S), giving rise to a spin-orbit interaction (proportional to $\mathbf{L} \cdot \mathbf{S}$ for a diatomic.) Although spin-orbit effects can be important, they are generally neglected in quantum chemical calculations.

7.2 The Born-Oppenheimer Approximation

We know that if a Hamiltonian is separable into two or more terms, then the total eigenfunctions are products of the individual eigenfunctions of the separated Hamiltonian terms, and the total eigenvalues are sums of individual eigenvalues of the separated Hamiltonian terms.

Consider, for example, a Hamiltonian which is separable into two terms, one involving coordinate q_1 and the other involving coordinate q_2.

$$\hat{H} = \hat{H}_1(q_1) + \hat{H}_2(q_2) \tag{163}$$

with the overall Schrödinger equation being

$$\hat{H}\psi(q_1, q_2) = E\psi(q_1, q_2) \tag{164}$$

If we assume that the total wavefunction can be written in the form $\psi(q_1, q_2) = \psi_1(q_1)\psi_2(q_2)$, where $\psi_1(q_1)$ and $\psi_2(q_2)$ are eigenfunctions of $\hat{H}_1$ and $\hat{H}_2$ with eigenvalues E_1 and E_2, then

$$\begin{aligned}
\hat{H}\psi(q_1, q_2) &= (\hat{H}_1 + \hat{H}_2)\psi_1(q_1)\psi_2(q_2) \qquad (165)\\
&= \hat{H}_1\psi_1(q_1)\psi_2(q_2) + \hat{H}_2\psi_1(q_1)\psi_2(q_2)\\
&= E_1\psi_1(q_1)\psi_2(q_2) + E_2\psi_1(q_1)\psi_2(q_2)\\
&= (E_1 + E_2)\psi_1(q_1)\psi_2(q_2)\\
&= E\psi(q_1, q_2)
\end{aligned}$$

Thus the eigenfunctions of $\hat{H}$ are products of the eigenfunctions of $\hat{H}_1$ and $\hat{H}_2$, and the eigenvalues are the sums of eigenvalues of $\hat{H}_1$ and $\hat{H}_2$.

If we examine the nonrelativistic Hamiltonian (162), we see that the term

$$\sum_{Ai} \frac{Z_A}{r_{Ai}} \tag{166}$$

prevents us from cleanly separating the electronic and nuclear coordinates and writing the total wavefunction as $\psi(\mathbf{r}, \mathbf{R}) = \psi_e(\mathbf{r})\psi_N(\mathbf{R})$, where $\mathbf{r}$ represents the set of all electronic coordinates, and $\mathbf{R}$ represents the set of all nuclear coordinates. The Born-Oppenheimer approximation is to assume that this separation is nevertheless *approximately* correct.

Qualitatively, the Born-Oppenheimer approximation rests on the fact that the nuclei are much more massive than the electrons. This allows us to say that the nuclei are nearly fixed with respect to electron motion. We can fix $\mathbf{R}$, the nuclear configuration, at some value $\mathbf{R_a}$, and solve for $\psi_e(\mathbf{r}; \mathbf{R_a})$; the electronic wavefunction depends only parametrically on $\mathbf{R}$. If we do this for a range of $\mathbf{R}$, we obtain the potential energy curve along which the nuclei move.

We now show the mathematical details. Let us abbreviate the molecular Hamiltonian as

$$\hat{H} = \hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{eN}(\mathbf{r}, \mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) \tag{167}$$

where the meaning of the individual terms should be obvious. Initially, $\hat{T}_N(\mathbf{R})$ can be neglected since $\hat{T}_N$ is smaller than $\hat{T}_e$ by a factor of M_A/m_e, where m_e is the mass of an electron. Thus for a *fixed* nuclear configuration, we have

$$\hat{H}_{el} = \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r}; \mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) \tag{168}$$

such that

$$\hat{H}_{el}\phi_e(\mathbf{r}; \mathbf{R}) = E_{el}\phi_e(\mathbf{r}; \mathbf{R}) \tag{169}$$

This is the "clamped-nuclei" Schrödinger equation. Quite frequently $\hat{V}_{NN}(\mathbf{R})$ is neglected in the above equation, which is justified since in this case $\mathbf{R}$ is just a parameter so that $\hat{V}_{NN}(\mathbf{R})$ is just a constant and shifts the eigenvalues only by some constant amount. Leaving $\hat{V}_{NN}(\mathbf{R})$ out of the electronic Schrödinger equation leads to a similar equation,

$$\hat{H}_e = \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r}; \mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) \tag{170}$$

$$\hat{H}_e \phi_e(\mathbf{r};\mathbf{R}) = E_e \phi_e(\mathbf{r};\mathbf{R}) \tag{171}$$

where we have used a new subscript "e" on the electronic Hamiltonian and energy to distinguish from the case where $\hat{V}_{NN}$ is included.

We now consider again the original Hamiltonian (167). If we insert a wavefunction of the form $\phi_T(\mathbf{r},\mathbf{R}) = \phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R})$, we obtain

$$\hat{H}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) = E_{tot}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) \tag{172}$$

$$\{\hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r},\mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) = E_{tot}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) \tag{173}$$

Since $\hat{T}_e$ contains no $\mathbf{R}$ dependence,

$$\hat{T}_e\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) = \phi_N(\mathbf{R})\hat{T}_e\phi_e(\mathbf{r};\mathbf{R}) \tag{174}$$

However, we may not immediately assume

$$\hat{T}_N\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) = \phi_e(\mathbf{r};\mathbf{R})\hat{T}_N\phi_N(\mathbf{R}) \tag{175}$$

(this point is tacitly assumed by most introductory textbooks). By the chain rule,

$$\nabla_A^2\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) = \phi_e(\mathbf{r};\mathbf{R})\nabla_A^2\phi_N(\mathbf{R}) + 2\nabla_A\phi_e(\mathbf{r};\mathbf{R})\nabla_A\phi_N(\mathbf{R}) + \phi_N(\mathbf{R})\nabla_A^2\phi_e(\mathbf{r};\mathbf{R}) \tag{176}$$

Using these facts, along with the electronic Schrödinger equation,

$$\{\hat{T}_e + \hat{V}_{eN}(\mathbf{r};\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\}\phi_e(\mathbf{r};\mathbf{R}) = \hat{H}_e\phi_e(\mathbf{r};\mathbf{R}) = E_e\phi_e(\mathbf{r};\mathbf{R}) \tag{177}$$

we simplify (173) to

$$\begin{aligned} &\phi_e(\mathbf{r};\mathbf{R})\hat{T}_N\phi_N(\mathbf{R}) + \phi_N(\mathbf{R})\phi_e(\mathbf{r};\mathbf{R})(E_e + \hat{V}_{NN}) \qquad (178) \\ &- \left\{ \sum_A \frac{1}{2M_A}(2\nabla_A\phi_e(\mathbf{r};\mathbf{R})\nabla_A\phi_N(\mathbf{R}) + \phi_N(\mathbf{R})\nabla_A^2\phi_e(\mathbf{r};\mathbf{R})) \right\} \\ &= E_{tot}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) \end{aligned}$$

We must now estimate the magnitude of the last term in brackets. Following Steinfeld [5], a typical contribution has the form $1/(2M_A)\nabla_A^2\phi_e(\mathbf{r};\mathbf{R})$, but

$\nabla_A \phi_e(\mathbf{r};\mathbf{R})$ is of the same order as $\nabla_i \phi_e(\mathbf{r};\mathbf{R})$ since the derivatives operate over approximately the same dimensions. The latter is $\phi_e(\mathbf{r};\mathbf{R})p_e$, with p_e the momentum of an electron. Therefore $1/(2M_A)\nabla_A^2 \phi_e(\mathbf{r};\mathbf{R}) \approx p_e^2/(2M_A) = (m/M_A)E_e$. Since $m/M_A \sim 1/10000$, the term in brackets can be dropped, giving

$$\phi_e(\mathbf{r};\mathbf{R})\hat{T}_N \phi_N(\mathbf{R}) + \phi_N(\mathbf{R})E_e\phi_e(\mathbf{r};\mathbf{R}) + \phi_N(\mathbf{R})\hat{V}_{NN}\phi_e(\mathbf{r};\mathbf{R}) = E_{tot}\phi_e(\mathbf{r};\mathbf{R})\phi_N(\mathbf{R}) \tag{179}$$

$$\{\hat{T}_N + E_e + \hat{V}_{NN}\}\phi_N(\mathbf{R}) = E_{tot}\phi_N(\mathbf{R}) \tag{180}$$

This is the nuclear Shrodinger equation we anticipated—the nuclei move in a potential set up by the electrons.

To summarize, the large difference in the relative masses of the electrons and nuclei allows us to approximately separate the wavefunction as a product of nuclear and electronic terms. The electronic wavefucntion $\phi_e(\mathbf{r};\mathbf{R})$ is solved for a given set of nuclear coordinates,

$$\hat{H}_e \phi_e(\mathbf{r};\mathbf{R}) = \left\{ -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{A,i} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}} \right\} \phi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R})\phi_e(\mathbf{r};\mathbf{R}) \tag{181}$$

and the electronic energy obtained contributes a potential term to the motion of the nuclei described by the nuclear wavefunction $\phi_N(\mathbf{R})$.

$$\hat{H}_N \phi_N(\mathbf{R}) = \left\{ -\sum_A \frac{1}{2M_A}\nabla_A^2 + E_e(\mathbf{R}) + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} \right\} \phi_N(\mathbf{R}) = E_{tot}\phi_N(\mathbf{R}) \tag{182}$$

As a final note, many textbooks, including Szabo and Ostlund [4], mean total energy *at fixed geometry* when they use the term "total energy" (i.e., they neglect the nuclear kinetic energy). This is just E_{el} of equation (169), which is also E_e plus the nuclear-nuclear repulsion. A somewhat more detailed treatment of the Born-Oppenheimer approximation is given elsewhere [6].

7.3 Separation of the Nuclear Hamiltonian

The nuclear Schrödinger equation can be approximately factored into translational, rotational, and vibrational parts. McQuarrie [1] explains how to do this for a diatomic in section 10-13. The rotational part can be cast into the form of the rigid rotor model, and the vibrational part can be written as a system of harmonic oscillators. Time does not allow further comment on the nuclear Schrödinger equation, although it is central to molecular spectroscopy.

8 Solving the Electronic Eigenvalue Problem

Once we have invoked the Born-Oppenheimer approximation, we attempt to solve the electronic Schrödinger equation (171), i.e.

$$\left[-\frac{1}{2} \sum_i \nabla_i^2 - \sum_{iA} \frac{Z_A}{r_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} \right] \psi_e(\mathbf{r}; \mathbf{R}) = E_e \psi_e(\mathbf{r}; \mathbf{R}) \tag{183}$$

But, as mentioned previously, this equation is quite difficult to solve!

8.1 The Nature of Many-Electron Wavefunctions

Let us consider the nature of the electronic wavefunctions $\psi_e(\mathbf{r}; \mathbf{R})$. Since the electronic wavefunction depends only parametrically on $\mathbf{R}$, we will suppress $\mathbf{R}$ in our notation from now on. What do we require of $\psi_e(\mathbf{r})$? Recall that $\mathbf{r}$ represents the set of all electronic coordinates, i.e., $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N\}$. So far we have left out one important item—we need to include the *spin* of each electron. We can define a new variable $\mathbf{x}$ which represents the set of all four coordinates associated with an electron: three spatial coordinates $\mathbf{r}$, and one spin coordinate ω, i.e., $\mathbf{x} = \{\mathbf{r}, \omega\}$.

Thus we write the electronic wavefunction as $\psi_e(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$. Why have we been able to avoid including spin until now? Because the non-relativistic Hamiltonian does not include spin. Nevertheless, spin must be included so that the electronic wavefunction can satisfy a very important requirement, which is the *antisymmetry principle* (see Postulate 6 in Section 4). This principle states that for a system of fermions, the wavefunction must be antisymmetric with respect to the interchange of all (space *and* spin) coordinates of one fermion with those of another. That is,

$$\psi_e(\mathbf{x}_1, \dots, \mathbf{x}_a, \dots, \mathbf{x}_b, \dots, \mathbf{x}_N) = -\psi_e(\mathbf{x}_1, \dots, \mathbf{x}_b, \dots, \mathbf{x}_a, \dots, \mathbf{x}_N) \tag{184}$$

The Pauli exclusion principle is a direct consequence of the antisymmetry principle.

A very important step in simplifying $\psi_e(\mathbf{x})$ is to expand it in terms of a set of one-electron functions, or "orbitals." This makes the electronic Schrödinger equation considerably easier to deal with.[3] A *spin orbital* is a function of the space and spin coordinates of a single electron, while a *spatial* orbital is a function of a single electron's spatial coordinates only. We can write a spin orbital as a product of a spatial orbital one of the two spin functions

$$\chi(\mathbf{x}) = \psi(\mathbf{r})|\alpha\rangle \tag{185}$$

or

$$\chi(\mathbf{x}) = \psi(\mathbf{r})|\beta\rangle \tag{186}$$

Note that for a given spatial orbital $\psi(\mathbf{r})$, we can form *two* spin orbitals, one with α spin, and one with β spin. The spatial orbital will be doubly occupied. It is possible (although sometimes frowned upon) to use one set of spatial orbitals for spin orbitals with α spin and another set for spin orbitals with β spin.[4]

Where do we get the one-particle spatial orbitals $\psi(\mathbf{r})$? That is beyond the scope of the current section, but we briefly itemize some of the more common possibilities:

- Orbitals centered on each atom (atomic orbitals).
- Orbitals centered on each atom but also symmetry-adapted to have the correct point-group symmetry species (symmetry orbitals).
- Molecular orbitals obtained from a Hartree-Fock procedure.

We now explain how an N-electron function $\psi_e(\mathbf{x})$ can be constructed from spin orbitals, following the arguments of Szabo and Ostlund [4] (p. 60). Assume we have a complete set of functions of a single variable $\{\chi_i(x)\}$. Then any function of a single variable can be expanded exactly as

$$\Phi(x_1) = \sum_i a_i \chi_i(x_1). \tag{187}$$

---

[3] It is not completely *necessary* to do this, however; for example, the Hylleras treatment of the Helium atom uses two-particle basis functions which are not further expanded in terms of single-particle functions.

[4] This is the procedure of the Unrestricted Hartree Fock (UHF) method.

How can we expand a function of *two* variables, e.g. $\Phi(x_1, x_2)$?

If we hold x_2 fixed, then

$$\Phi(x_1, x_2) = \sum_i a_i(x_2)\chi_i(x_1). \tag{188}$$

Now note that each expansion coefficient $a_i(x_2)$ is a function of a single variable, which can be expanded as

$$a_i(x_2) = \sum_j b_{ij}\chi_j(x_2). \tag{189}$$

Substituting this expression into the one for $\Phi(x_1, x_2)$, we now have

$$\Phi(x_1, x_2) = \sum_{ij} b_{ij}\chi_i(x_1)\chi_j(x_2) \tag{190}$$

a process which can obviously be extended for $\Phi(x_1, x_2, \ldots, x_N)$.

We can extend these arguments to the case of having a complete set of functions of the variable $\mathbf{x}$ (recall $\mathbf{x}$ represents x, y, and z and also ω). In that case, we obtain an analogous result,

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \sum_{ij} b_{ij}\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) \tag{191}$$

Now we must make sure that the antisymmetry principle is obeyed. For the two-particle case, the requirement

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = -\Phi(\mathbf{x}_2, \mathbf{x}_1) \tag{192}$$

implies that $b_{ij} = -b_{ji}$ and $b_{ii} = 0$, or

$$\begin{aligned} \Phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{j>i} b_{ij}[\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)] \\ &= \sum_{j>i} b_{ij}|\chi_i\chi_j\rangle \end{aligned} \tag{193}$$

where we have used the symbol $|\chi_i\chi_j\rangle$ to represent a *Slater determinant*, which in the genreral case is written

$$|\chi_1\chi_2\dots\chi_N\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \dots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \dots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \dots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{194}$$

We can extend the reasoning applied here to the case of N electrons; any N-electron wavefunction can be expressed exactly as a linear combination of all possible N-electron Slater determinants formed from a complete set of spin orbitals $\{\chi_i(\mathbf{x})\}$.

8.2 Matrix Mechanics

As we mentioned previously in section 2, Heisenberg's matrix mechanics, although little-discussed in elementary textbooks on quantum mechanics, is nevertheless formally equivalent to Schrödinger's wave equations. Let us now consider how we might solve the time-independent Schrödinger equation in matrix form.

If we want to solve $\hat{H}\psi_e(\mathbf{x}) = E_e\psi_e(\mathbf{x})$ as a matrix problem, we need to find a suitable linear vector space. Now $\psi_e(\mathbf{x})$ is an N-electron function that must be antisymmetric with respect to interchange of electronic coordinates. As we just saw in the previous section, any such N-electron function can be expressed *exactly* as a linear combination of Slater determinants, within the space spanned by the set of orbitals $\{\chi(\mathbf{x})\}$. If we denote our Slater determinant basis functions as $|\Phi_i\rangle$, then we can express the eigenvectors as

$$|\Psi_i\rangle = \sum_j^I c_{ij}|\Phi_j\rangle \tag{195}$$

for I possible N-electron basis functions (I will be infinite if we actually have a complete set of one electron functions χ). Similarly, we construct the matrix $\mathbf{H}$ in this basis by $H_{ij} = \langle\Phi_i|H|\Phi_j\rangle$.

If we solve this matrix equation, $\mathbf{H}|\Psi_n\rangle = E_n|\Psi_n\rangle$, in the space of all possible Slater determinants as just described, then the procedure is called *full configuration-interaction*, or full CI. A full CI constitues the *exact* solution to the time-independent Schrödinger equation within the given space of the spin orbitals χ. If we restrict the N-electron basis set in some way, then we will solve Schrödinger's equation *approximately.* The method is then called "configuration interaction," where we have dropped the prefix "full." For more information on configuration interaction, see the lecture notes by the present author [7] or one of the available review articles [8, 9].

References

[1] D. A. McQuarrie, *Quantum Chemistry.* University Science Books, Mill Valey, CA, 1983.

[2] E. Merzbacher, *Quantum Mechanics.* Wiley, New York, 2nd edition, 1970.

[3] I. N. Levine, *Quantum Chemistry.* Prentice Hall, Englewood Cliffs, NJ, 4th edition, 1991.

[4] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* McGraw-Hill, New York, 1989.

[5] J. I. Steinfeld, *Molecules and Radiation: An Introduction to Modern Molecular Spectroscopy.* MIT Press, Cambridge, MA, second edition, 1985.

[6] C. D. Sherrill, The Born-Oppenheimer approximation, 1997, http://vergil.chemistry.gatech.edu/notes/bo/bo.html.

[7] C. D. Sherrill, An introduction to configuration interaction theory, 1995, http://vergil.chemistry.gatech.edu/notes/ci/.

[8] I. Shavitt, The method of configuration interaction, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 189–275. Plenum Press, New York, 1977.

[9] C. D. Sherrill and H. F. Schaefer, The configuration interaction method: Advances in highly correlated approaches, in *Advances in Quantum Chemistry*, edited by P.-O. Löwdin, volume 34, pages 143–269. Academic Press, New York, 1999.