

The Variation Method

8.1 The Variation Theorem

To deal with the time-independent Schrödinger equation for systems (such as atoms or molecules) that contain interacting particles, we must use approximation methods. This chapter discusses the variation method, which allows us to approximate the ground-state energy of a system without solving the Schrödinger equation. The variation method is based on the following theorem:

THE VARIATION THEOREM

Given a system whose Hamiltonian operator $\hat{H}$ is time independent and whose lowest-energy eigenvalue is E_1 , if ϕ is any normalized, well-behaved function of the coordinates of the system's particles that satisfies the boundary conditions of the problem, then

$$\int \phi^* \hat{H} \phi d\tau \geq E_1, \quad \phi \text{ normalized} \quad (8.1)$$

The variation theorem allows us to calculate an upper bound for the system's ground-state energy.

To prove (8.1), we expand ϕ in terms of the complete, orthonormal set of eigenfunctions of $\hat{H}$, the stationary-state eigenfunctions ψ_k :

$$\phi = \sum_k a_k \psi_k \quad (8.2)$$

where

$$\hat{H}\psi_k = E_k \psi_k \quad (8.3)$$

Note that the expansion (8.2) requires that ϕ obey the same boundary conditions as the ψ_k 's. Substitution of (8.2) into the left side of (8.1) gives

$$\int \phi^* \hat{H} \phi d\tau = \int \sum_k a_k^* \psi_k^* \hat{H} \sum_j a_j \psi_j d\tau = \int \sum_k a_k^* \psi_k^* \sum_j a_j \hat{H} \psi_j d\tau$$

Using the eigenvalue equation (8.3) and assuming the validity of interchanging the integration and the infinite summations, we get

$$\begin{aligned} \int \phi^* \hat{H} \phi d\tau &= \int \sum_k a_k^* \psi_k^* \sum_j a_j E_j \psi_j d\tau = \sum_k \sum_j a_k^* a_j E_j \int \psi_k^* \psi_j d\tau \\ &= \sum_k \sum_j a_k^* a_j E_j \delta_{kj} \end{aligned}$$

where the orthonormality of the eigenfunctions ψ_k was used. We perform the sum over j , and, as usual, the Kronecker delta makes all terms zero except the one with $j = k$, giving

$$\int \phi^* \hat{H} \phi d\tau = \sum_k a_k^* a_k E_k = \sum_k |a_k|^2 E_k \quad (8.4)$$

Since E_1 is the lowest-energy eigenvalue of $\hat{H}$, we have $E_k \geq E_1$. Since $|a_k|^2$ is never negative, we can multiply the inequality $E_k \geq E_1$ by $|a_k|^2$ without changing the direction of the inequality sign to get $|a_k|^2 E_k \geq |a_k|^2 E_1$. Therefore, $\sum_k |a_k|^2 E_k \geq \sum_k |a_k|^2 E_1$, and use of (8.4) gives

$$\int \phi^* \hat{H} \phi d\tau = \sum_k |a_k|^2 E_k \geq \sum_k |a_k|^2 E_1 = E_1 \sum_k |a_k|^2 \quad (8.5)$$

Because ϕ is normalized, we have $\int \phi^* \phi d\tau = 1$. Substitution of the expansion (8.2) into the normalization condition gives

$$\begin{aligned} 1 = \int \phi^* \phi d\tau &= \int \sum_k a_k^* \psi_k^* \sum_j a_j \psi_j d\tau = \sum_k \sum_j a_k^* a_j \int \psi_k^* \psi_j d\tau = \sum_k \sum_j a_k^* a_j \delta_{kj} \\ 1 &= \sum_k |a_k|^2 \end{aligned} \quad (8.6)$$

[Note that in deriving Eqs. (8.4) and (8.6) we essentially repeated the derivations of Eqs. (7.70) and (7.69), respectively.]

Use of (8.6) in (8.5) gives the variation theorem (8.1):

$$\int \phi^* \hat{H} \phi d\tau \geq E_1, \quad \phi \text{ normalized} \quad (8.7)$$

Suppose we have a function ϕ that is not normalized. To apply the variation theorem, we multiply ϕ by a normalization constant N so that $N\phi$ is normalized. Replacing ϕ by $N\phi$ in (8.7), we have

$$|N|^2 \int \phi^* \hat{H} \phi d\tau \geq E_1 \quad (8.8)$$

N is determined by $\int (N\phi)^* N\phi d\tau = |N|^2 \int \phi^* \phi d\tau = 1$; so $|N|^2 = 1 / \int \phi^* \phi d\tau$ and (8.8) becomes

$$\boxed{\frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} \geq E_1} \quad (8.9)$$

where ϕ is any well-behaved function (not necessarily normalized) that satisfies the boundary conditions of the problem.

The function ϕ is called a **trial variation function**, and the integral in (8.1) [or the ratio of integrals in (8.9)] is called the **variational integral**. To arrive at a good approximation to the ground-state energy E_1 , we try many trial variation functions and look for the one that gives the lowest value of the variational integral. From (8.1), the lower the value of the variational integral, the better the approximation we have to E_1 . One way to disprove quantum mechanics would be to find a trial variation function that made the variational integral less than E_1 for some system where E_1 is known.

Let ψ_1 be the true ground-state wave function:

$$\hat{H}\psi_1 = E_1\psi_1 \quad (8.10)$$

If we happened to be lucky enough to hit upon a variation function that was equal to ψ_1 , then, using (8.10) in (8.1), we see that the variational integral will be equal to E_1 . Thus the ground-state wave function gives the minimum value of the variational integral. We therefore expect that the lower the value of the variational integral, the closer the trial variation function will approach the true ground-state wave function. However, it turns out that the variational integral approaches E_1 a lot faster than the trial variation function approaches ψ_1 , and it is possible to get a rather good approximation to E_1 using a rather poor ϕ .

In practice, one usually puts several parameters into the trial function ϕ and then varies these parameters so as to minimize the variational integral. Successful use of the variation method depends on the ability to make a good choice for the trial function.

Let us look at some examples of the variation method. Although the real utility of the method is for problems to which we do not know the true solutions, we will consider problems that are exactly solvable so that we can judge the accuracy of our results.

EXAMPLE

Devise a trial variation function for the particle in a one-dimensional box of length l .

The wave function is zero outside the box and the boundary conditions require that $\psi = 0$ at $x = 0$ and at $x = l$. The variation function ϕ must meet these boundary conditions of being zero at the ends of the box. As noted after Eq. (4.57), the ground-state ψ has no nodes interior to the boundary points, so it is desirable that ϕ have no interior nodes. A simple function that has these properties is the parabolic function

$$\phi = x(l - x) \quad \text{for } 0 \leq x \leq l \quad (8.11)$$

and $\phi = 0$ outside the box. Since we have not normalized ϕ , we use Eq. (8.9). Inside the box the Hamiltonian is $-(\hbar^2/2m) d^2/dx^2$. For the numerator and denominator of (8.9), we have

$$\int \phi^* \hat{H} \phi d\tau = -\frac{\hbar^2}{2m} \int_0^l (lx - x^2) \frac{d^2}{dx^2} (lx - x^2) dx = \frac{\hbar^2}{m} \int_0^l (lx - x^2) dx = \frac{\hbar^2 l^3}{6m} \quad (8.12)$$

$$\int \phi^* \phi d\tau = \int_0^l x^2 (l - x)^2 dx = \frac{l^5}{30} \quad (8.13)$$

Substituting in the variation theorem (8.9), we get

$$E_1 \leq \frac{5\hbar^2}{4\pi^2 ml^2} = 0.1266515 \frac{\hbar^2}{ml^2}$$

From Eq. (2.20), $E_1 = \hbar^2/8ml^2 = 0.125\hbar^2/ml^2$, and the energy error is 1.3%.

Since $\int |\phi|^2 d\tau = l^5/30$, the normalized form of (8.11) is $(30/l^5)^{1/2} x(l - x)$.

Figure 7.3 shows that this function rather closely resembles the true ground-state particle-in-a-box wave function.

EXERCISE A one-particle, one-dimensional system has the potential energy function $V = V_0$ for $0 \leq x \leq l$ and $V = \infty$ elsewhere (where V_0 is a constant). (a) Use the variation function $\phi = \sin(\pi x/l)$ for $0 \leq x \leq l$ and $\phi = 0$ elsewhere to estimate the ground-state energy of this system. (b) Explain why the result of (a) is the exact ground-state energy. *Hint:* See one of the Chapter 4 problems. (Answer: (a) $V_0 + \hbar^2/8ml^2$.)

The preceding example did not have a parameter in the trial function. The next example does.

EXAMPLE

For the one-dimensional harmonic oscillator, devise a variation function with a parameter and find the optimum value of the parameter. Estimate the ground-state energy.

The variation function ϕ must be quadratically integrable and so must go to zero as x goes to $\pm\infty$. The function e^{-x} has the proper behavior at $+\infty$ but becomes infinite at $-\infty$. The function e^{-x^2} has the proper behavior at $\pm\infty$. However, it is dimensionally unsatisfactory, since the power to which we raise e must be dimensionless. This can be seen from the Taylor series $e^z = 1 + z + z^2/2! + \cdots$ [Eq. (4.44)]. Since all the terms in this series must have the same dimensions, z must have the same dimensions as 1; that is, z in e^z must be dimensionless. Hence we modify e^{-x^2} to e^{-cx^2} , where c has units of length^{-2} . We shall take c as a variational parameter. The true ground-state ψ must have no nodes. Also, since $V = \frac{1}{2}kx^2$ is an even function, the ground-state ψ must have definite parity and must be an even function, since an odd function has a node at the origin. The trial function e^{-cx^2} has the desired properties of having no nodes and of being an even function.

Use of (4.30) for $\hat{H}$ and Appendix integrals gives (Prob. 8.3)

$$\begin{aligned}\int \phi^* \hat{H} \phi d\tau &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} e^{-cx^2} \frac{d^2 e^{-cx^2}}{dx^2} dx + 2\pi^2 v^2 m \int_{-\infty}^{\infty} x^2 e^{-2cx^2} dx \\ &= \frac{\hbar^2}{m} \left(\frac{\pi c}{8} \right)^{1/2} + v^2 m \left(\frac{\pi^5}{8} \right)^{1/2} c^{-3/2} \\ \int \phi^* \phi d\tau &= \int_{-\infty}^{\infty} e^{-2cx^2} dx = 2 \int_0^{\infty} e^{-2cx^2} dx = \left(\frac{\pi}{2c} \right)^{1/2}\end{aligned}$$

The variational integral W is

$$W \equiv \frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} = \frac{\hbar^2 c}{2m} + \frac{\pi^2 v^2 m}{2c} \quad (8.14)$$

We now vary c to minimize the variational integral (8.14). A necessary condition that W be minimized is that

$$\begin{aligned}\frac{dW}{dc} = 0 &= \frac{\hbar^2}{2m} - \frac{\pi^2 v^2 m}{2c^2} \\ c &= \pm \pi v m / \hbar\end{aligned} \quad (8.15)$$

The negative root $c = -\pi v m / \hbar$ is rejected, since it would make $\phi = e^{-cx^2}$ not quadratically integrable. Substitution of $c = \pi v m / \hbar$ into (8.14) gives $W = \frac{1}{2} \hbar v$. This is the exact ground-state harmonic-oscillator energy. With $c = \pi v m / \hbar$ the variation function ϕ is the same (except for being unnormalized) as the harmonic-oscillator ground-state wave function (4.53) and (4.31).

For the normalized harmonic-oscillator variation function $\phi = (2c/\pi)^{1/4} e^{-cx^2}$, a large value of c makes ϕ fall off very rapidly from its maximum value at $x = 0$. This makes the probability density large only near $x = 0$. The potential energy $V = \frac{1}{2} kx^2$ is low near $x = 0$, so a large c means a low $\langle V \rangle = \langle \phi | V | \phi \rangle$. [Note also that $\langle V \rangle$ equals the second term on the right side of (8.14).] However, because a large c makes ϕ fall off very rapidly from its maximum, it makes $|d\phi/dx|$ large in the region near

$x = 0$. From Prob. 7.7b, a large $|d\phi/dx|$ means a large value of $\langle T \rangle$ [which equals the first term on the right side of (8.14)]. The optimum value of c minimizes the sum $\langle T \rangle + \langle V \rangle = W$. In atoms and molecules, the true wave function is a compromise between the tendency to minimize $\langle V \rangle$ by confining the electrons to regions of low V (near the nuclei) and the tendency to minimize $\langle T \rangle$ by allowing the electron probability density to spread out over a large region.

EXERCISE Consider a one-particle, one-dimensional system with $V = 0$ for $-\frac{1}{2}l \leq x \leq \frac{1}{2}l$ and $V = b\hbar^2/ml^2$ elsewhere (where b is a positive constant) (Fig. 2.5 with $V_0 = b\hbar^2/ml^2$ and the origin shifted). (a) For the variation function $\phi = (x - c)(x + c) = x^2 - c^2$ for $-c \leq x \leq c$ and $\phi = 0$ elsewhere, where the variational parameter c satisfies $c > \frac{1}{2}l$, one finds that the variational integral W is given by

$$W = \frac{\hbar^2}{ml^2} \left[\frac{5l^2}{4c^2} + b \left(1 - \frac{15l}{16c} + \frac{5l^3}{32c^3} - \frac{3l^5}{256c^5} \right) \right]$$

Sketch ϕ and V on the same plot. Find the equation satisfied by the value of c that minimizes W . (b) Find the optimum c and W for $V_0 = 20\hbar^2/ml^2$ and compare with the true ground-state energy $2.814\hbar^2/ml^2$ (Prob. 4.31c). (*Hint*: You may want to use the Solver in a spreadsheet or a programmable calculator to find c/l .) (*Answer*: (a) $48t^4 - 24t^2 - 128t^3/b + 3 = 0$, where $t \equiv c/l$. (b) $c = 0.6715l$, $W = 3.454\hbar^2/ml^2$.)

8.2 Extension of the Variation Method

The variation method as presented in the last section gives information about only the *ground-state* energy and wave function. We now discuss extension of the variation method to excited states. (See also Section 8.5.)

Consider how we might extend the variation method to estimate the energy of the first excited state. We number the stationary states of the system 1, 2, 3, ... in order of increasing energy:

$$E_1 \leq E_2 \leq E_3 \leq \dots$$

We showed that for a normalized variational function ϕ [Eqs. (8.4) and (8.6)]

$$\int \phi^* \hat{H} \phi \, d\tau = \sum_{k=1}^{\infty} |a_k|^2 E_k \quad \text{and} \quad \int \phi^* \phi \, d\tau = \sum_{k=1}^{\infty} |a_k|^2 = 1$$

where the a_k 's are the expansion coefficients in $\phi = \sum_k a_k \psi_k$ [Eq. (8.2)]. We have $a_k = \langle \psi_k | \phi \rangle$ [Eq. (7.40)]. Let us restrict ourselves to normalized functions ϕ that are orthogonal to the true ground-state wave function ψ_1 . Then we have $a_1 = \langle \psi_1 | \phi \rangle = 0$ and

$$\int \phi^* \hat{H} \phi \, d\tau = \sum_{k=2}^{\infty} |a_k|^2 E_k \quad \text{and} \quad \int \phi^* \phi \, d\tau = \sum_{k=2}^{\infty} |a_k|^2 = 1 \quad (8.16)$$

For $k \geq 2$, we have $E_k \geq E_2$ and $|a_k|^2 E_k \geq |a_k|^2 E_2$. Hence

$$\sum_{k=2}^{\infty} |a_k|^2 E_k \geq \sum_{k=2}^{\infty} |a_k|^2 E_2 = E_2 \sum_{k=2}^{\infty} |a_k|^2 = E_2 \quad (8.17)$$

Combining (8.16) and (8.17), we have the desired result:

$$\int \phi^* \hat{H} \phi \, d\tau \geq E_2 \quad \text{if} \quad \int \psi_1^* \phi \, d\tau = 0 \quad \text{and} \quad \int \phi^* \phi \, d\tau = 1 \quad (8.18)$$

The inequality (8.18) allows us to get an upper bound to the energy E_2 of the first excited state. However, the restriction $\langle \psi_1 | \phi \rangle = 0$ makes this method troublesome to apply.

For certain systems, it is possible to be sure that $\langle \psi_1 | \phi \rangle = 0$ even though we do not know the true ground-state wave function. An example is a one-dimensional problem for which V is an even function of x . In this case the ground-state wave function is always an even function, while the first excited-state wave function is odd. (All the wave functions must be of definite parity. The ground-state wave function is nodeless, and, since an odd function vanishes at the origin, the ground-state wave function must be even. The first excited-state wave function has one node and must be odd.) Therefore, for odd trial functions, it must be true that $\langle \psi_1 | \phi \rangle = 0$; the even function ψ_1 times the odd function ϕ gives an odd integrand whose integral from $-\infty$ to ∞ is zero.

Another example is a particle moving in a central field (Section 6.1). The form of the potential energy might be such that we could not solve for the radial factor $R(r)$ in the eigenfunction. However, the angular factor in ψ is a spherical harmonic [Eq. (6.16)], and spherical harmonics with different values of l are orthogonal. Thus we can get an upper bound to the energy of the lowest state with any given angular momentum l by using the factor Y_l^m in the trial function. This result depends on the extension of (8.18) to higher excited states:

$$\frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} \geq E_{k+1} \quad \text{if} \quad \int \psi_1^* \phi d\tau = \int \psi_2^* \phi d\tau = \cdots = \int \psi_k^* \phi d\tau = 0 \quad (8.19)$$

8.3 Determinants

Section 8.5 discusses a kind of variation function that gives rise to an equation involving a determinant. Therefore, we now discuss determinants.

A **determinant** is a square array of n^2 quantities (called **elements**); the value of the determinant is calculated from its elements in a manner to be given shortly. The number n is the **order** of the determinant. Using a_{ij} to represent a typical element, we write the n th-order determinant as

$$\det(a_{ij}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \quad (8.20)$$

The vertical lines in (8.20) have nothing to do with absolute value. Before considering how the value of the n th-order determinant is defined, we consider determinants of first, second, and third orders.

A first-order determinant has one element, and its value is simply the value of that element. Thus

$$|a_{11}| = a_{11} \quad (8.21)$$

where the vertical lines indicate a determinant and not an absolute value.

A second-order determinant has four elements, and its value is defined by

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (8.22)$$

The value of a third-order determinant is defined by

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad (8.23)$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \quad (8.24)$$

A third-order determinant is evaluated by writing down the elements of the top row with alternating plus and minus signs and then multiplying each element by a certain second-order determinant; the second-order determinant that multiplies a given element is found by crossing out the row and column of the third-order determinant in which that element appears. The $(n - 1)$ -order determinant obtained by striking out the i th row and the j th column of the n th-order determinant is called the **minor** of the element a_{ij} . We define the **cofactor** of a_{ij} as the minor of a_{ij} times the factor $(-1)^{i+j}$. Thus (8.23) states that a third-order determinant is evaluated by multiplying each element of the top row by its cofactor and adding up the three products. [Note that (8.22) conforms to this evaluation by means of cofactors, since the cofactor of a_{11} in (8.22) is a_{22} , and the cofactor of a_{12} is $-a_{21}$.] A numerical example is

$$\begin{vmatrix} 5 & 10 & 2 \\ 0.1 & 3 & 1 \\ 0 & 4 & 4 \end{vmatrix} = 5 \begin{vmatrix} 3 & 1 \\ 4 & 4 \end{vmatrix} - 10 \begin{vmatrix} 0.1 & 1 \\ 0 & 4 \end{vmatrix} + 2 \begin{vmatrix} 0.1 & 3 \\ 0 & 4 \end{vmatrix} \\ = 5(8) - 10(0.4) + 2(0.4) = 36.8$$

Denoting the minor of a_{ij} by M_{ij} and the cofactor of a_{ij} by C_{ij} , we have

$$C_{ij} = (-1)^{i+j}M_{ij} \quad (8.25)$$

The expansion (8.23) of the third-order determinant can be written as

$$\det(a_{ij}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \quad (8.26)$$

A third-order determinant can be expanded using the elements of any row and the corresponding cofactors. For example, using the second row to expand the third-order determinant, we have

$$\det(a_{ij}) = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \quad (8.27)$$

$$\det(a_{ij}) = -a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \quad (8.28)$$

and expansion of the second-order determinants shows that (8.28) is equal to (8.24). We may also use the elements of any column and the corresponding cofactors to expand the determinant, as can be readily verified. Thus for the third-order determinant, we can write

$$\det(a_{ij}) = a_{k1}C_{k1} + a_{k2}C_{k2} + a_{k3}C_{k3} = \sum_{l=1}^3 a_{kl}C_{kl}, \quad k = 1 \text{ or } 2 \text{ or } 3 \\ \det(a_{ij}) = a_{1k}C_{1k} + a_{2k}C_{2k} + a_{3k}C_{3k} = \sum_{l=1}^3 a_{lk}C_{lk}, \quad k = 1 \text{ or } 2 \text{ or } 3$$

The first expansion uses one of the rows; the second uses one of the columns.

We define determinants of higher order by an analogous row (or column) expansion. For an n th-order determinant,

$$\det(a_{ij}) = \sum_{l=1}^n a_{kl} C_{kl} = \sum_{l=1}^n a_{lk} C_{lk}, \quad k = 1 \text{ or } 2 \text{ or } \dots \text{ or } n \quad (8.29)$$

Some theorems on determinants are as follows (for proofs, see *Sokolnikoff and Redheffer*, pp. 702–707):

- I. If every element of a row (or column) of a determinant is zero, the value of the determinant is zero.
- II. Interchanging any two rows (or columns) multiplies the value of a determinant by -1 .
- III. If any two rows (or columns) of a determinant are identical, the determinant has the value zero.
- IV. Multiplication of each element of any one row (or any one column) by some constant k multiplies the value of the determinant by k .
- V. Addition to each element of one row of the same constant multiple of the corresponding element of another row leaves the value of the determinant unchanged. This theorem also applies to the addition of a multiple of one column to another column.
- VI. The interchange of all corresponding rows and columns leaves the value of the determinant unchanged. (This interchange means that column one becomes row one, column two becomes row two, etc.)

EXAMPLE

Use Theorem V to evaluate

$$B = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} \quad (8.30)$$

Addition of -2 times the elements of row one to the corresponding elements of row four changes row four to $2 + (-2)1 = 0$, $3 + (-2)(2) = -1$, $4 + (-2)3 = -2$, and $1 + (-2)4 = -7$. Then, addition of -3 times row one to row three and -4 times row one to row two gives

$$B = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -7 & -10 & -13 \\ 0 & -2 & -8 & -10 \\ 0 & -1 & -2 & -7 \end{vmatrix} = 1 \begin{vmatrix} -7 & -10 & -13 \\ -2 & -8 & -10 \\ -1 & -2 & -7 \end{vmatrix} \quad (8.31)$$

where we expanded B in terms of elements of the first column. Subtracting twice row three from row two and seven times row three from row one, we have

$$B = \begin{vmatrix} 0 & 4 & 36 \\ 0 & -4 & 4 \\ -1 & -2 & -7 \end{vmatrix} = (-1) \begin{vmatrix} 4 & 36 \\ -4 & 4 \end{vmatrix} = -(16 + 144) = -160 \quad (8.32)$$

The diagonal of a determinant that runs from the top left to the lower right is the **principal diagonal**. A **diagonal determinant** is a determinant all of whose elements are zero except those on the principal diagonal. For a diagonal determinant,

The procedure just outlined fails if the determinant of the inhomogeneous system of $n - 1$ equations in $n - 1$ unknowns [(8.38) with the last equation omitted] happens to be zero. Cramer's rule then has a zero in the denominator and is of no use. We could try to get around this difficulty by initially assigning the arbitrary value to another of the unknowns rather than to x_n . We could also try discarding some other equation of (8.38), rather than the last one. What we are looking for is a nonvanishing determinant of order $n - 1$ formed from the determinant of the coefficients of the system (8.37) by striking out a row and a column. If such a determinant exists, then by the procedure given, with the right choice of the equation to be discarded and the right choice of the unknown to be assigned an arbitrary value, we can solve the system and will get solutions of the form (8.39). If no such determinant exists, we must assign arbitrary values to two of the unknowns and attempt to proceed from there. Thus the solution to (8.37) might contain two (or even more) arbitrary constants.

An efficient way to solve a system of linear homogeneous equations is to do Gauss–Jordan elimination on the equations. If only the trivial solution exists, the final set of equations obtained will be $x_1 = 0, x_2 = 0, \dots, x_n = 0$. If a nontrivial solution exists, at least one equation will be reduced to the form $0 = 0$; if m equations of the form $0 = 0$ are obtained, we assign arbitrary constants to m of the unknowns and express the remaining unknowns in terms of these m unknowns.

EXAMPLE

Use Gauss–Jordan elimination to solve the set of equations in the preceding example.

In doing Gaussian or Gauss–Jordan elimination on a set of n inhomogeneous or homogeneous equations, we can eliminate needless writing by omitting the variables $x_1, \dots, x_n$ and writing down only the n -row, $(n + 1)$ -column array of coefficients and constant terms (including any zero coefficients); we then produce the next array by operating on the numbers of each row as if that row were the equation it represents.

To eliminate one set of divisions, we interchange the first and second equations so that we start with $a_{11} = 1$. Detaching the coefficients and proceeding with Gauss–Jordan elimination, we have

$$\begin{array}{cccc|cccc|cccc|cccc} 1 & 3 & -2 & 0 & 1 & 3 & -2 & 0 & 1 & 3 & -2 & 0 & 1 & 0 & \frac{11}{5} & 0 \\ 3 & 4 & 1 & 0 & \rightarrow & 0 & -5 & 7 & 0 & \rightarrow & 0 & 1 & -\frac{7}{5} & 0 & \rightarrow & 0 & 1 & -\frac{7}{5} & 0 \\ 1 & -2 & 5 & 0 & & 0 & -5 & 7 & 0 & & 0 & -5 & 7 & 0 & & 0 & 0 & 0 & 0 \end{array}$$

The first array is the original set of equations with the first and second equations interchanged. To eliminate x_1 from the second and third equations, we subtract 3 times row one from row two and 1 times row one from row three, thereby producing the second array. Division of row two by -5 produces the third array. To eliminate x_2 from the first and third equations, we subtract 3 times row two from row one and -5 times row two from row three, thereby producing the fourth array. Because the fourth array has the x_3 coefficient in row three equal to zero, we cannot use row three to eliminate x_3 from rows one and two (as would be the last step in the Gauss–Jordan algorithm). Discarding the last equation, which reads $0 = 0$, we assign $x_3 = k$, where k is an arbitrary constant. The first and second equations in the last array read $x_1 + \frac{11}{5}x_3 = 0$ and $x_2 - \frac{7}{5}x_3 = 0$, or $x_1 = -\frac{11}{5}x_3$, $x_2 = \frac{7}{5}x_3$. The general solution is $x_1 = -\frac{11}{5}k$, $x_2 = \frac{7}{5}k$, $x_3 = k$.

8.5 Linear Variation Functions

A special kind of variation function widely used in the study of molecules is the linear variation function. A **linear variation function** is a linear combination of n linearly independent functions $f_1, f_2, \dots, f_n$:

$$\phi = c_1 f_1 + c_2 f_2 + \dots + c_n f_n = \sum_{j=1}^n c_j f_j \quad (8.40)$$

where ϕ is the trial variation function and the coefficients c_j are parameters to be determined by minimizing the variational integral. The functions f_j (which are called **basis functions**) must satisfy the boundary conditions of the problem. We shall restrict ourselves to *real* ϕ so that the c_j 's and f_j 's are all real. In (8.40), the functions f_j are known functions.

We now apply the variation theorem (8.9). For the real linear variation function, we have

$$\int \phi^* \phi \, d\tau = \int \sum_{j=1}^n c_j f_j \sum_{k=1}^n c_k f_k \, d\tau = \sum_{j=1}^n \sum_{k=1}^n c_j c_k \int f_j f_k \, d\tau \equiv \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} \quad (8.41)$$

where we defined the **overlap integral** S_{jk} as

$$S_{jk} \equiv \int f_j^* f_k \, d\tau \quad (8.42)$$

Note that S_{jk} is not necessarily equal to δ_{jk} , since there is no reason to suppose that the functions f_j are mutually orthogonal. They are not necessarily the eigenfunctions of any operator. The numerator in (8.9) is

$$\begin{aligned} \int \phi^* \hat{H} \phi \, d\tau &= \int \sum_{j=1}^n c_j f_j \hat{H} \sum_{k=1}^n c_k f_k \, d\tau \\ &= \sum_{j=1}^n \sum_{k=1}^n c_j c_k \int f_j \hat{H} f_k \, d\tau \equiv \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk} \end{aligned}$$

where we defined H_{jk} as

$$H_{jk} \equiv \int f_j^* \hat{H} f_k \, d\tau \quad (8.43)$$

The variational integral W is

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} = \frac{\sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk}}{\sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk}} \quad (8.44)$$

$$W \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk} \quad (8.45)$$

We now minimize W so as to approach as closely as we can to E_1 ($W \geq E_1$). The variational integral W is a function of the n independent variables $c_1, c_2, \dots, c_n$:

$$W = W(c_1, c_2, \dots, c_n)$$

A necessary condition for a minimum in a function W of several variables is that its partial derivatives with respect to each of the variables must be zero at the minimum:

$$\frac{\partial W}{\partial c_i} = 0, \quad i = 1, 2, \dots, n \quad (8.46)$$

We now differentiate (8.45) partially with respect to each c_i to obtain n equations:

$$\frac{\partial W}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} + W \frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk}, \quad i = 1, 2, \dots, n \quad (8.47)$$

Now

$$\frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{j=1}^n \sum_{k=1}^n \left[\frac{\partial}{\partial c_i} (c_j c_k) \right] S_{jk} = \sum_{j=1}^n \sum_{k=1}^n \left(c_k \frac{\partial c_j}{\partial c_i} + c_j \frac{\partial c_k}{\partial c_i} \right) S_{jk}$$

The c_j 's are independent variables, and therefore

$$\begin{aligned} \frac{\partial c_j}{\partial c_i} &= 0 \quad \text{if } j \neq i, & \frac{\partial c_j}{\partial c_i} &= 1 \quad \text{if } j = i \\ \frac{\partial c_j}{\partial c_i} &= \delta_{ij} \end{aligned} \quad (8.48)$$

We then have

$$\frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{k=1}^n \sum_{j=1}^n c_k \delta_{ij} S_{jk} + \sum_{j=1}^n \sum_{k=1}^n c_j \delta_{ik} S_{jk} = \sum_{k=1}^n c_k S_{ik} + \sum_{j=1}^n c_j S_{ji}$$

where we evaluated one of the sums in each double summation using Eq. (7.32). Use of (7.4) gives

$$S_{ji} = S_{ij}^* = S_{ij} \quad (8.49)$$

where the last equality follows because we are dealing with real functions. Hence,

$$\frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{k=1}^n c_k S_{ik} + \sum_{j=1}^n c_j S_{ij} = \sum_{k=1}^n c_k S_{ik} + \sum_{k=1}^n c_k S_{ik} = 2 \sum_{k=1}^n c_k S_{ik} \quad (8.50)$$

where the fact that j is a dummy variable was used.

By replacing S_{jk} by H_{jk} in each of these manipulations, we get

$$\frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk} = 2 \sum_{k=1}^n c_k H_{ik} \quad (8.51)$$

This result depends on the fact that

$$H_{ji} = H_{ij}^* = H_{ij} \quad (8.52)$$

which is true because $\hat{H}$ is Hermitian, and f_i , f_j , and $\hat{H}$ are real.

Substitution of Eqs. (8.46), (8.50), and (8.51) into (8.47) gives

$$\begin{aligned} 2W \sum_{k=1}^n c_k S_{ik} &= 2 \sum_{k=1}^n c_k H_{ik}, \quad i = 1, 2, \dots, n \\ \sum_{k=1}^n [(H_{ik} - S_{ik}W)c_k] &= 0, \quad i = 1, 2, \dots, n \end{aligned} \quad (8.53)$$

true wave functions of the system. Unfortunately, we usually need an infinite number of functions to have a complete set.

Quantum chemists may use dozens, hundreds, thousands, or even millions of terms in linear variation functions so as to get accurate results for molecules. Obviously, a computer is essential for this work. The most efficient way to solve (8.58) (which is called the **secular equation**) and the associated linear equations (8.55) is by matrix methods (Section 8.6).

To obtain an approximation to the wave function of the ground state, we take the lowest root W_1 of the secular equation and substitute it in the set of equations (8.55); we then solve this set of equations for the coefficients $c_1^{(1)}, c_2^{(1)}, \dots, c_n^{(1)}$, where the superscript (1) was added to indicate that these coefficients correspond to W_1 . [As noted in the previous section, we can determine only the ratios of the coefficients. We solve for $c_2^{(1)}, \dots, c_n^{(1)}$ in terms of $c_1^{(1)}$, and then determine $c_1^{(1)}$ by normalization.] Having found the $c_k^{(1)}$'s, we take $\phi_1 = \sum_k c_k^{(1)} f_k$ as an approximate ground-state wave function. Use of higher roots of (8.58) in (8.55) gives approximations to excited-state wave functions. These approximate wave functions can be shown to be orthogonal (Prob. 8.40).

Solution of (8.58) and (8.55) is simplified by having as many of the integrals equal to zero as possible. We can make some of the off-diagonal H_{ij} 's vanish by choosing the functions f_k as eigenfunctions of some operator $\hat{A}$ that commutes with $\hat{H}$. If f_i and f_j correspond to different eigenvalues of $\hat{A}$, then H_{ij} vanishes (Theorem 6 of Section 7.4). If the functions f_k are orthonormal, the off-diagonal S_{ij} 's vanish ($S_{ij} = \delta_{ij}$). If the initially chosen f_k 's are not orthogonal, we can use the Schmidt (or some other) procedure to find n linear combinations of these f_k 's that are orthogonal and then use the orthogonalized functions.

Equations (8.55) and (8.58) are also valid when the restriction that the variation function be real is removed (Prob. 8.39).

EXAMPLE

Add functions to the function $x(l - x)$ of the first example of Section 8.1 to form a linear variation function for the particle in a one-dimensional box of length l and find approximate energies and wave functions for the lowest four states.

In the trial function $\phi = \sum_{k=1}^n c_k f_k$, we take $f_1 = x(l - x)$. Since we want approximations to the lowest four states, n must be at least 4. There are an infinite number of possible well-behaved functions that could be used for f_2, f_3 , and f_4 . The function $x^2(l - x)^2$ obeys the boundary conditions of vanishing at $x = 0$ and $x = l$ and leads to simple integrals, so we take $f_2 = x^2(l - x)^2$.

If the origin is placed at the center of the box, the potential energy (Fig. 2.1) is an even function, and, as noted in Section 8.2, the wave functions alternate between being even and odd functions (see also Fig. 2.3). (Throughout this example, the terms even and odd will refer to having the origin at the box's center.) The functions $f_1 = x(l - x)$ and $f_2 = x^2(l - x)^2$ are both even functions (see Prob. 8.34). If we were to take $\phi = c_1 x(l - x) + c_2 x^2(l - x)^2$, we would end up with upper bounds to the energies of the lowest two states with even wave functions (the $n = 1$ and $n = 3$ states) and would get approximate wave functions for these two states. Since we also want to approximate the $n = 2$ and $n = 4$ states, we shall add in two functions that are odd. An odd function must vanish at the origin [as noted after Eq. (4.50)], so we need functions that vanish at the box midpoint $x = \frac{1}{2}l$, as well as at $x = 0$ and l . A simple function with these properties is $f_3 = x(l - x)(\frac{1}{2}l - x)$. To get f_4 , we shall multiply f_2 by $(\frac{1}{2}l - x)$. Thus we take $\phi = \sum_{k=1}^4 c_k f_k$, with

$$f_1 = x(l - x), f_2 = x^2(l - x)^2, f_3 = x(l - x)(\frac{1}{2}l - x), f_4 = x^2(l - x)^2(\frac{1}{2}l - x) \quad (8.62)$$

Note that f_1, f_2, f_3 , and f_4 are linearly independent, as assumed in (8.40).

Because f_1 and f_2 are even, while f_3 and f_4 are odd, many integrals will vanish. Thus

$$S_{13} = S_{31} = 0, \quad S_{14} = S_{41} = 0, \quad S_{23} = S_{32} = 0, \quad S_{24} = S_{42} = 0 \quad (8.63)$$

because the integrand in each of these overlap integrals is an odd function with respect to the origin at the box center. The functions f_1, f_2, f_3, f_4 are eigenfunctions of the parity operator $\hat{\Pi}$ (Section 7.5) with the even functions f_1 and f_2 having parity eigenvalue $+1$ and f_3 and f_4 having eigenvalue -1 . The operator $\hat{\Pi}$ commutes with $\hat{H}$ (since V is an even function), so by Theorem 6 of Section 7.4, H_{ij} vanishes if f_i is an odd function and f_j is even, or vice versa. Thus

$$H_{13} = H_{31} = 0, \quad H_{14} = H_{41} = 0, \quad H_{23} = H_{32} = 0, \quad H_{24} = H_{42} = 0 \quad (8.64)$$

From (8.63) and (8.64), the $n = 4$ secular equation (8.58) becomes

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W & 0 & 0 \\ H_{21} - S_{21}W & H_{22} - S_{22}W & 0 & 0 \\ 0 & 0 & H_{33} - S_{33}W & H_{34} - S_{34}W \\ 0 & 0 & H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0 \quad (8.65)$$

The secular determinant is in block-diagonal form and so is equal to the product of its blocks [Eq. (8.34)]:

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W \\ H_{21} - S_{21}W & H_{22} - S_{22}W \end{vmatrix} \times \begin{vmatrix} H_{33} - S_{33}W & H_{34} - S_{34}W \\ H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0$$

The four roots of this equation are found from the equations

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W \\ H_{21} - S_{21}W & H_{22} - S_{22}W \end{vmatrix} = 0 \quad (8.66)$$

$$\begin{vmatrix} H_{33} - S_{33}W & H_{34} - S_{34}W \\ H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0 \quad (8.67)$$

Let the roots of (8.66) (which are approximations to the $n = 1$ and $n = 3$ energies) be W_1 and W_3 and let the roots of (8.67) be W_2 and W_4 . After solving the secular equation for the W 's, we substitute them one at a time into the set of equations (8.55) to find the coefficients c_k in the variation function. From the secular equation (8.65), the set of equations (8.55) with the root W_1 is

$$\left. \begin{aligned} (H_{11} - S_{11}W_1)c_1^{(1)} + (H_{12} - S_{12}W_1)c_2^{(1)} &= 0 \\ (H_{21} - S_{21}W_1)c_1^{(1)} + (H_{22} - S_{22}W_1)c_2^{(1)} &= 0 \end{aligned} \right\} \quad (8.68a)$$

$$\left. \begin{aligned} (H_{33} - S_{33}W_1)c_3^{(1)} + (H_{34} - S_{34}W_1)c_4^{(1)} &= 0 \\ (H_{43} - S_{43}W_1)c_3^{(1)} + (H_{44} - S_{44}W_1)c_4^{(1)} &= 0 \end{aligned} \right\} \quad (8.68b)$$

Because W_1 is a root of (8.66), the set of equations (8.68a) has the determinant of its coefficients [which is the determinant in (8.66)] equal to zero. Hence (8.68a) has a nontrivial solution for $c_1^{(1)}$ and $c_2^{(1)}$. However, W_1 is not a root of (8.67), so the determinant of the coefficients of the set of equations (8.68b) is nonzero. Hence, (8.68b) has only the trivial solution $c_3^{(1)} = c_4^{(1)} = 0$. The trial function ϕ_1 corresponding to the root W_1 thus has the form $\phi_1 = \sum_{k=1}^4 c_k^{(1)} f_k = c_1^{(1)} f_1 + c_2^{(1)} f_2$. The same reasoning

shows that ϕ_3 is a linear combination of f_1 and f_2 , while ϕ_2 and ϕ_4 are each linear combinations of f_3 and f_4 :

$$\begin{aligned}\phi_1 &= c_1^{(1)}f_1 + c_2^{(1)}f_2, & \phi_3 &= c_1^{(3)}f_1 + c_2^{(3)}f_2 \\ \phi_2 &= c_3^{(2)}f_3 + c_4^{(2)}f_4, & \phi_4 &= c_3^{(4)}f_3 + c_4^{(4)}f_4\end{aligned}\quad (8.69)$$

The even wave functions ψ_1 and ψ_3 are approximated by linear combinations of the even functions f_1 and f_2 ; the odd functions ψ_2 and ψ_4 are approximated by linear combinations of the odd functions f_3 and f_4 .

When the secular equation is in block-diagonal form, it factors into two or more smaller secular equations, and the set of simultaneous equations (8.55) breaks up into two or more smaller sets of equations.

We now must evaluate the H_{ij} and S_{ij} integrals so as to solve (8.66) and (8.67) for W_1, W_2, W_3 , and W_4 . We have

$$H_{11} = \langle f_1 | \hat{H} | f_1 \rangle = \int_0^l x(l-x) \left(\frac{-\hbar^2}{2m} \right) \frac{d^2}{dx^2} [x(l-x)] dx = \frac{\hbar^2 l^3}{6m}$$

$$S_{11} = \langle f_1 | f_1 \rangle = \int_0^l x^2(l-x)^2 dx = \frac{l^5}{30}$$

where Eqs. (8.12) and (8.13) were used. Evaluation of the remaining integrals using (8.62), (8.49), and (8.52) gives (Prob. 8.35)

$$\begin{aligned}H_{12} &= H_{21} = \langle f_2 | \hat{H} | f_1 \rangle = \hbar^2 l^5 / 30m, & H_{22} &= \hbar^2 l^7 / 105m \\ H_{33} &= \hbar^2 l^5 / 40m, & H_{44} &= \hbar^2 l^9 / 1260m, & H_{34} &= H_{43} = \hbar^2 l^7 / 280m \\ S_{12} &= S_{21} = \langle f_1 | f_2 \rangle = l^7 / 140, & S_{22} &= l^9 / 630 \\ S_{33} &= l^7 / 840, & S_{44} &= l^{11} / 27720, & S_{34} &= S_{43} = l^9 / 5040\end{aligned}$$

Equation (8.66) becomes

$$\begin{vmatrix} \frac{\hbar^2 l^3}{6m} - \frac{l^5}{30}W & \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140}W \\ \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140}W & \frac{\hbar^2 l^7}{105m} - \frac{l^9}{630}W \end{vmatrix} = 0 \quad (8.70)$$

Using Theorem IV of Section 8.3, we eliminate the fractions by multiplying row one of the determinant by $420m/l^3$, row two by $1260m/l^5$, and the right side of (8.70) by both factors, to get

$$\begin{vmatrix} 70\hbar^2 - 14ml^2W & 14\hbar^2 l^2 - 3ml^4W \\ 42\hbar^2 - 9ml^2W & 12\hbar^2 l^2 - 2ml^4W \end{vmatrix} = 0$$

$$m^2 l^4 W^2 - 56ml^2 \hbar^2 W + 252\hbar^4 = 0$$

$$W = (\hbar^2/ml^2)(28 \pm \sqrt{532}) = 0.1250018\hbar^2/ml^2, \quad 1.293495\hbar^2/ml^2 \quad (8.71)$$

Substitution of the integrals into (8.67) leads to the roots (Prob. 8.36)

$$W = (\hbar^2/ml^2)(60 \pm \sqrt{1620}) = 0.5002930\hbar^2/ml^2, \quad 2.5393425\hbar^2/ml^2 \quad (8.72)$$

The approximate values $(ml^2/\hbar^2)W = 0.1250018, 0.5002930, 1.293495$, and 2.5393425 may be compared with the exact values [Eq. (2.20)] $(ml^2/\hbar^2)E = 0.125, 0.5, 1.125$, and 2 for the four lowest states. The percent errors are 0.0014% , 0.059% , 15.0% , and 27.0% for $n = 1, 2, 3$, and 4 , respectively. We did great for $n = 1$ and 2 ; lousy for $n = 3$ and 4 .

We now find the approximate wave functions corresponding to these W 's. Substitution of $W_1 = 0.1250018h^2/ml^2$ into the set of equations (8.68a) corresponding to (8.71) gives (after division by h^2)

$$\begin{aligned} 0.023095c_1^{(1)} - 0.020381c_2^{(1)}l^2 &= 0 \\ -0.061144c_1^{(1)} + 0.053960c_2^{(1)}l^2 &= 0 \end{aligned} \quad (8.73)$$

where, for example, the first coefficient is found from

$$70\hbar^2 - 14ml^2W = 70h^2/4\pi^2 - 14(0.1250018)h^2 = 0.023095h^2$$

To solve the homogeneous equations (8.73), we follow the procedure given near the end of Section 8.4. We discard the second equation of (8.73), transfer the $c_2^{(1)}$ term to the right side, and solve for the coefficient ratio; we get

$$c_1^{(1)} = k, \quad c_2^{(1)} = 1.133k/l^2$$

where k is a constant. We find k from the normalization condition:

$$\begin{aligned} \langle \phi_1 | \phi_1 \rangle &= 1 = \langle kf_1 + 1.133kf_2/l^2 | kf_1 + 1.133kf_2/l^2 \rangle \\ &= k^2(\langle f_1 | f_1 \rangle + 2.266\langle f_1 | f_2 \rangle/l^2 + 1.284\langle f_2 | f_2 \rangle/l^4) \\ &= k^2(S_{11} + 2.266S_{12}/l^2 + 1.284S_{22}/l^4) = 0.05156k^2l^5 \end{aligned}$$

where the previously found values of the overlap integrals were used. Therefore $k = 4.404/l^{5/2}$ and

$$\begin{aligned} \phi_1 &= c_1^{(1)}f_1 + c_2^{(1)}f_2 = 4.404f_1/l^{5/2} + 4.990f_2/l^{9/2} \\ \phi_1 &= l^{-1/2}[4.404(x/l)(1 - x/l) + 4.990(x/l)^2(1 - x/l)^2] \end{aligned} \quad (8.74)$$

where (8.62) was used.

Using W_2 , W_3 , and W_4 in turn in (8.55), we find the following normalized linear variation functions (Prob. 8.38), where $X \equiv x/l$:

$$\begin{aligned} \phi_2 &= l^{-1/2}[16.78X(1 - X)(\tfrac{1}{2} - X) + 71.85X^2(1 - X)^2(\tfrac{1}{2} - X)] \\ \phi_3 &= l^{-1/2}[28.65X(1 - X) - 132.7X^2(1 - X)^2] \\ \phi_4 &= l^{-1/2}[98.99X(1 - X)(\tfrac{1}{2} - X) - 572.3X^2(1 - X)^2(\tfrac{1}{2} - X)] \end{aligned} \quad (8.75)$$

8.6 Matrices, Eigenvalues, and Eigenvectors

Matrix algebra (Section 7.10) was developed during the period 1855–1858 by the British mathematician Arthur Cayley as a shorthand way of dealing with simultaneous linear equations and linear transformations from one set of variables to another. [Cayley had to support himself by working as a lawyer for 14 years until he obtained a professorship in 1860. While working as a lawyer, he published hundreds of papers in mathematics and regularly discussed mathematics with his fellow lawyer John Joseph Sylvester. Sylvester coined the term *matrix* for rectangular arrays (after the Latin word for “womb”) and obtained a mathematics professorship in 1855.] Matrix algebra remained unknown to most physicists for many years, and when Heisenberg discovered the matrix-mechanics version of quantum mechanics in 1925, he did not realize that the entities he was dealing with were matrices. When he showed his work to Born, Born, who was well-trained in mathematics, recognized that Heisenberg was using matrices. Nowadays, matrix algebra is an essential tool in quantum chemistry computations and is widely used in most branches of physics.

The set of linear inhomogeneous equations (8.35) can be written as the matrix equation

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (8.76)$$

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (8.77)$$

where $\mathbf{A}$ is the coefficient matrix and $\mathbf{x}$ and $\mathbf{b}$ are column matrices. The equivalence of (8.35) and (8.76) is readily verified using the matrix-multiplication rule (7.107).

The **determinant** of a square matrix $\mathbf{A}$ is the determinant whose elements are the same as the elements of $\mathbf{A}$. If $\det \mathbf{A} \neq 0$, the matrix $\mathbf{A}$ is said to be **nonsingular**.

The **inverse** of a square matrix $\mathbf{A}$ of order n is the square matrix whose product with $\mathbf{A}$ is the unit matrix of order n . Denoting the inverse by $\mathbf{A}^{-1}$, we have

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad (8.78)$$

where $\mathbf{I}$ is the unit matrix. One can prove that $\mathbf{A}^{-1}$ exists if and only if $\det \mathbf{A} \neq 0$. (For efficient methods of computing $\mathbf{A}^{-1}$, see *Press et al.*, Section 2.3; *Shoup*, Section 3.3; Prob. 8.51. Many spreadsheets have a built-in capability to find the inverse of a matrix.)

If $\det \mathbf{A} \neq 0$ for the coefficient matrix $\mathbf{A}$ in (8.76), then we can multiply each side of (8.77) by $\mathbf{A}^{-1}$ on the left to get $\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = \mathbf{A}^{-1}\mathbf{b}$. Since matrix multiplication is associative (Section 7.10), we have $\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$. Thus, left multiplication of (8.77) by $\mathbf{A}^{-1}$ gives $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ as the solution for the unknowns in a set of linear inhomogeneous equations.

The linear variation method is widely used to find approximate molecular wave functions, and matrix algebra gives the most computationally efficient method to solve the equations of the linear variation method. If the functions $f_1, \dots, f_n$ in the linear variation function $\phi = \sum_{k=1}^n c_k f_k$ are made to be orthonormal, then $S_{ij} = \int f_i^* f_j d\tau = \delta_{ij}$, and the homogeneous set of equations (8.55) for the coefficients c_k that minimize the variational integral becomes

$$\begin{aligned} H_{11}c_1 + H_{12}c_2 + \cdots + H_{1n}c_n &= Wc_1 \\ H_{21}c_1 + H_{22}c_2 + \cdots + H_{2n}c_n &= Wc_2 \\ \vdots & \\ H_{n1}c_1 + H_{n2}c_2 + \cdots + H_{nn}c_n &= Wc_n \end{aligned} \quad (8.79a)$$

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = W \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad (8.79b)$$

$$\mathbf{H}\mathbf{c} = W\mathbf{c} \quad (8.79c)$$

where $\mathbf{H}$ is the square matrix whose elements are $H_{ij} = \langle f_i | \hat{H} | f_j \rangle$ and $\mathbf{c}$ is the column vector of coefficients $c_1, \dots, c_n$. In (8.79c), $\mathbf{H}$ is a known matrix and $\mathbf{c}$ and W are unknowns to be solved for.

If

$$\mathbf{A}\mathbf{c} = \lambda\mathbf{c} \quad (8.80)$$

where $\mathbf{A}$ is a square matrix, $\mathbf{c}$ is a column vector with at least one nonzero element, and λ is a scalar, then $\mathbf{c}$ is said to be an **eigenvector** (or characteristic vector) of $\mathbf{A}$ and λ is an **eigenvalue** (or characteristic value) of $\mathbf{A}$.

Let us denote the n eigenvalues and the corresponding eigenvectors of $\mathbf{H}$ in the variation-method equations (8.79) by $W_1, W_2, \dots, W_n$ and $\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(n)}$, so that

$$\mathbf{H}\mathbf{c}^{(i)} = W_i \mathbf{c}^{(i)} \quad \text{for } i = 1, 2, \dots, n \quad (8.85)$$

where $\mathbf{c}^{(i)}$ is a column vector whose elements are $c_1^{(i)}, \dots, c_n^{(i)}$ and the basis functions f_i are orthonormal. Furthermore, let $\mathbf{C}$ be the square matrix whose columns are the eigenvectors of $\mathbf{H}$, and let $\mathbf{W}$ be the diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{H}$:

$$\mathbf{C} = \begin{pmatrix} c_1^{(1)} & c_1^{(2)} & \cdots & c_1^{(n)} \\ c_2^{(1)} & c_2^{(2)} & \cdots & c_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ c_n^{(1)} & c_n^{(2)} & \cdots & c_n^{(n)} \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_n \end{pmatrix} \quad (8.86)$$

The set of n eigenvalue equations (8.85) can be written as the single equation:

$$\mathbf{H}\mathbf{C} = \mathbf{C}\mathbf{W} \quad (8.87)$$

To verify the matrix equation (8.87), we show that each element $(\mathbf{H}\mathbf{C})_{ij}$ of the matrix $\mathbf{H}\mathbf{C}$ equals the corresponding element $(\mathbf{C}\mathbf{W})_{ij}$ of $\mathbf{C}\mathbf{W}$. The matrix-multiplication rule (7.107) gives $(\mathbf{H}\mathbf{C})_{ij} = \sum_k H_{ik}(\mathbf{C})_{kj} = \sum_k H_{ik}c_k^{(j)}$. Consider the eigenvalue equation $\mathbf{H}\mathbf{c}^{(j)} = W_j \mathbf{c}^{(j)}$ [Eq. (8.85)]. $\mathbf{H}\mathbf{c}^{(j)}$ and $W_j \mathbf{c}^{(j)}$ are column matrices. Using (7.107) to equate the elements in row i of each of these column matrices, we have $\sum_k H_{ik}c_k^{(j)} = W_j c_i^{(j)}$. Then

$$(\mathbf{H}\mathbf{C})_{ij} = \sum_k H_{ik}c_k^{(j)} = W_j c_i^{(j)}, \quad (\mathbf{C}\mathbf{W})_{ij} = \sum_k (\mathbf{C})_{ik}(\mathbf{W})_{kj} = \sum_k c_i^{(k)}\delta_{kj}W_k = c_i^{(j)}W_j$$

Hence $(\mathbf{H}\mathbf{C})_{ij} = (\mathbf{C}\mathbf{W})_{ij}$ and (8.87) is proved.

Provided $\mathbf{C}$ has an inverse (see below), we can multiply each side of (8.87) by $\mathbf{C}^{-1}$ on the left to get $\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{C}^{-1}(\mathbf{C}\mathbf{W})$. [Since matrix multiplication is not commutative, when we multiply each side of $\mathbf{H}\mathbf{C} = \mathbf{C}\mathbf{W}$ by $\mathbf{C}^{-1}$, we must put the factor $\mathbf{C}^{-1}$ on the left of $\mathbf{H}\mathbf{C}$ and on the left of $\mathbf{C}\mathbf{W}$ (or on the right of $\mathbf{H}\mathbf{C}$ and the right of $\mathbf{C}\mathbf{W}$).] We have $\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{C}^{-1}(\mathbf{C}\mathbf{W}) = (\mathbf{C}^{-1}\mathbf{C})\mathbf{W} = \mathbf{I}\mathbf{W} = \mathbf{W}$:

$$\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{W} \quad (8.88)$$

To simplify (8.88), we must learn more about matrices.

A square matrix $\mathbf{B}$ is a **symmetric matrix** if all its elements satisfy $b_{mn} = b_{nm}$. The elements of a symmetric matrix are symmetric about the principal diagonal; for example, $b_{12} = b_{21}$. A square matrix $\mathbf{D}$ is a **Hermitian matrix** if all its elements satisfy $d_{mn} = d_{nm}^*$. For example, if

$$\mathbf{M} = \begin{pmatrix} 2 & 5 & 0 \\ 5 & i & 2i \\ 0 & 2i & 4 \end{pmatrix}, \quad \mathbf{N} = \begin{pmatrix} 6 & 1 + 2i & 8 \\ 1 - 2i & -1 & -i \\ 8 & i & 0 \end{pmatrix} \quad (8.89)$$

then $\mathbf{M}$ is symmetric and $\mathbf{N}$ is Hermitian. (Note that the diagonal elements of a Hermitian matrix must be real; $d_{mm} = d_{mm}^*$.) A **real matrix** is one whose elements are all real numbers. A real Hermitian matrix is a symmetric matrix.

The **transpose** $\mathbf{A}^T$ (often written $\hat{\mathbf{A}}$) of the matrix $\mathbf{A}$ is the matrix formed by interchanging rows and columns of $\mathbf{A}$ so that column 1 becomes row one, column 2 becomes row two, and so on. The elements a_{mn}^T of $\mathbf{A}^T$ are related to the elements of $\mathbf{A}$ by $a_{mn}^T = a_{nm}$. For a square matrix, the transpose is found by reflecting the elements about the principal diagonal. A symmetric matrix is equal to its transpose. Thus, for the matrix $\mathbf{M}$ in (8.89), we have $\mathbf{M}^T = \mathbf{M}$.

The **complex conjugate** $\mathbf{A}^*$ of $\mathbf{A}$ is the matrix formed by taking the complex conjugate of each element of $\mathbf{A}$. The **conjugate transpose** $\mathbf{A}^\dagger$ (read as A dagger) of the matrix $\mathbf{A}$ is formed by taking the transpose of $\mathbf{A}^*$; thus $\mathbf{A}^\dagger = (\mathbf{A}^*)^T$ and

$$a_{mn}^\dagger = (a_{nm})^* \quad (8.90)$$

(Physicists call $\mathbf{A}^\dagger$ the *adjoint* of $\mathbf{A}$, a name that has been used by mathematicians to denote an entirely different matrix.) An example is

$$\mathbf{B} = \begin{pmatrix} 2 & 3+i \\ 0 & 4i \end{pmatrix}, \quad \mathbf{B}^T = \begin{pmatrix} 2 & 0 \\ 3+i & 4i \end{pmatrix}, \quad \mathbf{B}^\dagger = \begin{pmatrix} 2 & 0 \\ 3-i & -4i \end{pmatrix}$$

For a Hermitian matrix $\mathbf{A}$, Eq. (8.90) gives $a_{mn}^\dagger = (a_{nm})^* = a_{mn}$, so $(\mathbf{A}^\dagger)_{mn} = (\mathbf{A})_{mn}$ and $\mathbf{A}^\dagger = \mathbf{A}$. A Hermitian matrix is equal to its conjugate transpose. (Physicists often use the term *self-adjoint* for a Hermitian matrix.)

An **orthogonal matrix** is a square matrix whose inverse is equal to its transpose:

$$\mathbf{A}^{-1} = \mathbf{A}^T \quad \text{if } \mathbf{A} \text{ is orthogonal} \quad (8.91)$$

A **unitary matrix** is one whose inverse is equal to its conjugate transpose:

$$\mathbf{U}^{-1} = \mathbf{U}^\dagger \quad \text{if } \mathbf{U} \text{ is unitary} \quad (8.92)$$

From the definition (8.92), we have $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}$ if $\mathbf{U}$ is unitary. By equating $(\mathbf{U}^\dagger \mathbf{U})_{mn}$ to $(\mathbf{I})_{mn}$, we find (Prob. 8.43)

$$\sum_k u_{km}^* u_{kn} = \delta_{mn} \quad (8.93)$$

for columns m and n of a unitary matrix. Thus the columns of a unitary matrix (viewed as column vectors) are orthogonal and normalized (orthonormal), as defined by (8.84) and (8.83). Conversely, if (8.93) is true for all columns, then $\mathbf{U}$ is a unitary matrix. If $\mathbf{U}$ is unitary and real, then $\mathbf{U}^\dagger = \mathbf{U}^T$, and $\mathbf{U}$ is an orthogonal matrix.

One can prove that two eigenvectors of a Hermitian matrix $\mathbf{H}$ that correspond to different eigenvalues are orthogonal (see *Strang*, Section 5.5). For eigenvectors of $\mathbf{H}$ that correspond to the same eigenvalue, one can take linear combinations of them that will be orthogonal eigenvectors of $\mathbf{H}$. Moreover, the elements of an eigenvector can be multiplied by a constant to normalize the eigenvector. Hence, *the eigenvectors of a Hermitian matrix can be chosen to be orthonormal*. If the eigenvectors are chosen to be orthonormal, then the eigenvector matrix $\mathbf{C}$ in (8.86) is a unitary matrix, and $\mathbf{C}^{-1} = \mathbf{C}^\dagger$; Eq. (8.88) then becomes

$$\mathbf{C}^\dagger \mathbf{H} \mathbf{C} = \mathbf{W} \quad \text{if } \mathbf{H} \text{ is Hermitian} \quad (8.94)$$

For the common case that $\mathbf{H}$ is real as well as Hermitian (that is, $\mathbf{H}$ is real and symmetric), the c 's in (8.79a) are real (since $\mathbf{W}$ and the H_{ij} 's are real) and $\mathbf{C}$ is real as well as unitary; that is, $\mathbf{C}$ is orthogonal, with $\mathbf{C}^{-1} = \mathbf{C}^T$; Eq. (8.94) becomes

$$\mathbf{C}^T \mathbf{H} \mathbf{C} = \mathbf{W} \quad \text{if } \mathbf{H} \text{ is real and symmetric} \quad (8.95)$$

The eigenvalues of a Hermitian matrix can be proven to be real numbers (*Strang*, Section 5.5).

EXAMPLE

Find the eigenvalues and normalized eigenvectors of the Hermitian matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2i \\ -2i & 0 \end{pmatrix}$$

by solving algebraic equations. Then verify that $\mathbf{C}^\dagger \mathbf{A} \mathbf{C}$ is diagonal, where $\mathbf{C}$ is the eigenvector matrix.

The characteristic equation (8.81) for the eigenvalues λ is $\det(a_{ij} - \delta_{ij}\lambda) = 0$, which becomes

$$\begin{vmatrix} 3 - \lambda & 2i \\ -2i & -\lambda \end{vmatrix} = 0$$

$$\lambda^2 - 3\lambda - 4 = 0$$

$$\lambda_1 = 4, \quad \lambda_2 = -1$$

A useful theorem in checking eigenvalue calculations is the following (*Strang*, Exercise 5.1.9): The sum of the diagonal elements of a square matrix $\mathbf{A}$ of order n is equal to the sum of the eigenvalues λ_i of $\mathbf{A}$; that is, $\sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$. In this example, $\sum_i a_{ii} = 3 + 0$, which equals the sum $4 - 1 = 3$ of the eigenvalues.

For the root $\lambda_1 = 4$, the set of simultaneous equations (8.82) is

$$(3 - \lambda_1)c_1^{(1)} + 2ic_2^{(1)} = 0$$

$$-2ic_1^{(1)} - \lambda_1 c_2^{(1)} = 0$$

or

$$-c_1^{(1)} + 2ic_2^{(1)} = 0$$

$$-2ic_1^{(1)} - 4c_2^{(1)} = 0$$

Discarding either one of these equations, we find

$$c_1^{(1)} = 2ic_2^{(1)}$$

Normalization gives

$$1 = |c_1^{(1)}|^2 + |c_2^{(1)}|^2 = 4|c_2^{(1)}|^2 + |c_2^{(1)}|^2$$

$$|c_2^{(1)}| = 1/\sqrt{5}, \quad c_2^{(1)} = 1/\sqrt{5}$$

$$c_1^{(1)} = 2ic_2^{(1)} = 2i/\sqrt{5}$$

where the phase of $c_2^{(1)}$ was chosen to be zero.

Similarly, we find for $\lambda_2 = -1$ (Prob. 8.49)

$$c_1^{(2)} = -i/\sqrt{5}, \quad c_2^{(2)} = 2/\sqrt{5}$$

The normalized eigenvectors are then

$$\mathbf{c}^{(1)} = \begin{pmatrix} 2i/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}, \quad \mathbf{c}^{(2)} = \begin{pmatrix} -i/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

Because the eigenvalues λ_1 and λ_2 of the Hermitian matrix $\mathbf{A}$ differ, $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ are orthogonal (as you should verify). Also, $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ are normalized. Therefore, $\mathbf{C}$ is unitary and $\mathbf{C}^{-1} = \mathbf{C}^\dagger$. Forming $\mathbf{C}$ and its conjugate transpose, we have

$$\begin{aligned} \mathbf{C}^{-1} \mathbf{A} \mathbf{C} &= \mathbf{C}^\dagger \mathbf{A} \mathbf{C} = \begin{pmatrix} -2i/\sqrt{5} & 1/\sqrt{5} \\ i/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 3 & 2i \\ -2i & 0 \end{pmatrix} \begin{pmatrix} 2i/\sqrt{5} & -i/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} -2i/\sqrt{5} & 1/\sqrt{5} \\ i/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 8i/\sqrt{5} & i/\sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

which is the diagonal matrix of eigenvalues.

We have shown that if $\mathbf{H}$ is a real symmetric matrix with eigenvalues W_i and orthonormal eigenvectors $\mathbf{c}^{(i)}$ (that is, if $\mathbf{H}\mathbf{c}^{(i)} = W_i\mathbf{c}^{(i)}$ for $i = 1, 2, \dots, n$), then $\mathbf{C}^T\mathbf{H}\mathbf{C} = \mathbf{W}$ [Eq. (8.95)], where $\mathbf{C}$ is the real orthogonal matrix whose columns are the eigenvectors $\mathbf{c}^{(i)}$ and $\mathbf{W}$ is the diagonal matrix of eigenvalues W_i . The converse of this theorem is readily proved; that is, if $\mathbf{H}$ is a real symmetric matrix, $\mathbf{B}$ is a real orthogonal matrix, and $\mathbf{B}^T\mathbf{H}\mathbf{B}$ equals a diagonal matrix $\mathbf{\Lambda}$, then the columns of $\mathbf{B}$ are the eigenvectors of $\mathbf{H}$ and the diagonal elements of $\mathbf{\Lambda}$ are the eigenvalues of $\mathbf{H}$.

To find the eigenvalues and eigenvectors of a Hermitian matrix of order n , we can use either of the following procedures: (1) Solve the characteristic equation $\det(H_{ij} - \delta_{ij}W) = 0$ [Eq. (8.81)] for the eigenvalues $W_1, \dots, W_n$. Then substitute each W_k into the set of algebraic equations (8.79a) and solve for the elements $c_1^{(k)}, \dots, c_n^{(k)}$ of the k th eigenvector. (2) Search for a unitary matrix $\mathbf{C}$ such that $\mathbf{C}^\dagger\mathbf{H}\mathbf{C}$ is a diagonal matrix. The diagonal elements of $\mathbf{C}^\dagger\mathbf{H}\mathbf{C}$ are the eigenvalues of $\mathbf{H}$, and the columns of $\mathbf{C}$ are the orthonormal eigenvectors of $\mathbf{H}$. For the large matrices that occur in quantum chemistry, procedure (2) (called **matrix diagonalization**) is computationally much faster than (1).

One reason that expanding the characteristic determinant and solving the characteristic equation is not a good way to find the eigenvalues of large matrices is that, for large matrices, a very small change in a coefficient in the characteristic polynomial may produce a large change in the eigenvalues (see Prob. 8.54). Hence we might have to calculate the coefficients in the characteristic polynomial to hundreds or thousands of decimal places in order to get eigenvalues accurate to a few decimal places. Although it is true that for certain matrices a tiny change in the value of an element of that matrix might produce large changes in the eigenvalues, one can prove that for Hermitian matrices, a small change in a matrix element always produces only small changes in the eigenvalues. Hence method (2) of the preceding paragraph is the correct way to get accurate eigenvalues.

A systematic way to diagonalize a real symmetric matrix $\mathbf{H}$ is as follows. Construct an orthogonal matrix $\mathbf{O}_1$ such that the matrix $\mathbf{H}_1 \equiv \mathbf{O}_1^T\mathbf{H}\mathbf{O}_1$ has zero in place of the off-diagonal elements H_{12} and H_{21} of $\mathbf{H}$. (Because $\mathbf{H}$ is symmetric, we have $H_{12} = H_{21}$. Also, the transformed matrices $\mathbf{H}_1, \mathbf{H}_2, \dots$ are symmetric.) Then construct an orthogonal matrix $\mathbf{O}_2$ such that $\mathbf{H}_2 \equiv \mathbf{O}_2^T\mathbf{H}_1\mathbf{O}_2 = \mathbf{O}_2^T\mathbf{O}_1^T\mathbf{H}\mathbf{O}_1\mathbf{O}_2$ has zeros in place of the elements $(\mathbf{H}_1)_{13}$ and $(\mathbf{H}_1)_{31}$ of $\mathbf{H}_1$; and so on. Unfortunately, when a given pair of off-diagonal elements is made zero in a step, some off-diagonal elements made zero in a previous step are likely to become nonzero, so one has to go back and recycle through the off-diagonal elements over and over again. Generally an infinite number of steps are required to make all off-diagonal elements equal to zero. In practice, one skips a step if the absolute value of the off-diagonal elements to be zeroed in that step is less than some tiny number, and one stops the procedure when the absolute values of all off-diagonal elements are less than some tiny number. The eigenvalues are then the diagonal elements of the transformed matrix $\dots\mathbf{O}_3^T\mathbf{O}_2^T\mathbf{O}_1^T\mathbf{H}\mathbf{O}_1\mathbf{O}_2\mathbf{O}_3\dots$, and the eigenvector matrix is the product $\mathbf{O}_1\mathbf{O}_2\mathbf{O}_3\dots$. This method (the *cyclic Jacobi method*) is not very efficient for large matrices when run on a serial computer but is efficient on a parallel computer.

More-efficient approaches to diagonalize real, symmetric matrices than the Jacobi method begin by carrying out a series of orthogonal transformations to reduce the original matrix $\mathbf{H}$ to a symmetric tridiagonal matrix $\mathbf{T}$. A *tridiagonal matrix* is one whose elements are all zero except for those on the principal diagonal (elements t_{ii}) and those on the diagonals immediately above and immediately below the principal diagonal (elements $t_{i-1,i}$ and $t_{i+1,i}$, respectively). The relation between $\mathbf{T}$ and $\mathbf{H}$ is $\mathbf{T} = \mathbf{O}^T\mathbf{H}\mathbf{O}$, where $\mathbf{O}$ is a real orthogonal matrix that is the product of the orthogonal matrices used in the individual steps of going from $\mathbf{H}$ to $\mathbf{T}$. Two efficient methods of

transforming $\mathbf{H}$ to tridiagonal form are due to Givens and to Householder. An efficient method to find the eigenvalues of a symmetric tridiagonal matrix is the **QR** method. Here, $\mathbf{T}$ is expressed as the product of an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$ (one whose elements below the principal diagonal are all zero). A series of iterative steps yields matrices converging to a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{T}$, which equal the eigenvalues of $\mathbf{H}$ (Prob. 8.55). With certain refinements, the **QR** method is a very efficient way to find eigenvalues and eigenvectors (see *Strang*, Sections 5.3 and 7.3 for details).

Details of matrix diagonalization procedures and computer programs are given in *Press et al.*, Chapter 11; *Acton*, Chapters 8 and 13; *Shoup*, Chapter 4.

A major compilation of procedures and computer programs for scientific and engineering calculations is *Press et al.*; the text of older editions of this book is available free on the Internet at www.nr.com. For comments on older editions of this book, see amath.colorado.edu/computing/Fortran/numrec.html.

Programs for mathematical and scientific calculations can be found at www.netlib.org and at gams.nist.gov. Downloadable free personal-computer mathematical software and demonstration software for such commercial programs as Mathcad and Maple can be found at archives.math.utk.edu.

The procedure for using matrix algebra to solve the linear variation equations when nonorthonormal basis functions are used is outlined in Prob. 8.57.

The Excel spreadsheet can be used to find eigenvalues and eigenvectors; see Prob. 8.53.

Computer algebra systems such as Maple, Mathematica, and Mathcad and some electronic calculators have built-in commands to easily find eigenvalues and eigenvectors.

The methods for finding matrix eigenvalues and eigenvectors discussed in this section are useful for matrices of order up to 10^3 . Special methods are used to find the lowest few eigenvalues and corresponding eigenvectors of matrices of order up to 10^9 that occur in certain quantum-chemistry calculations (see Section 16.2).

As noted after Eq. (8.80), for the linear variation function $\sum_{i=1}^n c_i f_i$ with orthonormal basis functions f_i , the eigenvalues of the matrix $\mathbf{H}$ formed from the matrix elements $\langle f_i | \hat{H} | f_k \rangle$ are the roots of the secular equation, and the eigenvector corresponding to the eigenvalue W_m gives the coefficients in the variation function that corresponds to W_m . Problems 8.60 to 8.65 apply the linear variation method to problems such as the double well and the harmonic oscillator using particle-in-a-box wave functions as basis functions and using a computer algebra program such as Mathcad to find the eigenvalues and eigenvectors of the $\mathbf{H}$ matrix.

We have discussed matrix diagonalization in the context of the linear variation method. However, *finding the eigenvalues a_k and eigenfunctions g_k of any Hermitian operator ($\hat{A}g_k = a_k g_k$) can be formulated as a matrix-diagonalization problem.* If we choose a complete, orthonormal basis set $\{f_i\}$ and expand the eigenfunctions as $g_k = \sum_i c_i^{(k)} f_i$, then (Prob. 8.59) the eigenvalues of the matrix $\mathbf{A}$ whose elements are $a_{ij} = \langle f_i | \hat{A} | f_j \rangle$ are the eigenvalues of the operator $\hat{A}$, and the elements $c_i^{(k)}$ of the eigenvectors $\mathbf{c}^{(k)}$ of $\mathbf{A}$ give the coefficients in the expansions of the eigenfunctions g_k .

The material of this section further emphasizes the correspondence between linear operators and matrices and the correspondence between functions and column vectors (Section 7.10).

The PageRanks of Web pages indexed by Google are calculated by Google by constructing a certain matrix (called the Google matrix) and finding its eigenvector that corresponds to the eigenvalue 1. The order of the Google matrix equals the number of indexed pages and was between $10^{10.5}$ and 10^{11} in 2012. The components of this eigenvector are the PageRanks. For details, see www.ams.org/samplings/featurecolumn/fcarc-pagerank.

Summary

The variation theorem states that for a system with time-independent Hamiltonian $\hat{H}$ and ground-state energy E_1 , we have $\int \phi^* \hat{H} \phi d\tau / \int \phi^* \phi d\tau \geq E_1$, where ϕ is any well-behaved function that obeys the boundary conditions of the problem. The variation theorem allows us to obtain approximations for the ground-state energy and wave function.

The mathematics of determinants was reviewed in Section 8.3. A set of n linear homogeneous equations in n unknowns has a nontrivial solution if and only if the determinant of the coefficients is equal to zero.

For the linear variation function $\phi = \sum_{i=1}^n c_i f_i$, variation of the coefficients c_i to minimize the variational integral W leads to the secular equation $\det(H_{ij} - S_{ij}W) = 0$, whose roots $W_1, \dots, W_n$ are upper bounds for the n lowest bound-state energy eigenvalues; here, $H_{ij} \equiv \langle f_i | \hat{H} | f_j \rangle$ and $S_{ij} \equiv \langle f_i | f_j \rangle$. Substitution of $W_1, \dots, W_n$ one at a time into the simultaneous homogeneous equations (8.55) allows the coefficients c_i that correspond to each W to be found.

The eigenvalues λ_i and eigenvectors $\mathbf{c}_i$ of a square matrix $\mathbf{A}$ satisfy $\mathbf{A}\mathbf{c}_i = \lambda_i \mathbf{c}_i$. For a Hermitian matrix $\mathbf{H}$, the eigenvector matrix $\mathbf{C}$ (whose columns are the eigenvectors of $\mathbf{H}$) is unitary (meaning that its inverse equals its conjugate transpose $\mathbf{C}^\dagger$), and $\mathbf{C}^\dagger \mathbf{H} \mathbf{C}$ equals the diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{H}$. For a real symmetric matrix, the eigenvector matrix is orthogonal. For the linear variation function $\phi = \sum_i c_i f_i$ formed from orthonormal functions f_i , each set of optimized coefficients c_i is an eigenvector of the matrix whose elements are $\langle f_i | \hat{H} | f_j \rangle$ and the corresponding values of the variational integral are the eigenvalues of this matrix.

Problems

Sec.	8.1	8.2	8.3	8.4	8.5	8.6	general
Probs	8.1–8.16	8.17–8.18	8.19–8.22	8.23–8.28	8.29–8.40	8.41–8.56	8.57–8.66

- 8.1** Calculations on the Li atom with three well-behaved variation functions gave the following values of the variational integral: -203.2 eV, -192.0 eV, and -201.2 eV. Therefore, the true ground-state energy of Li must be ($\leq$ or $\geq$) than ? eV. (Choose one of the two inequality signs and replace the ? with a number.)
- 8.2** (a) Consider a one-particle, one-dimensional system with potential energy
- $$V = V_0 \quad \text{for } \frac{1}{4}l \leq x \leq \frac{3}{4}l, \quad V = 0 \quad \text{for } 0 \leq x \leq \frac{1}{4}l \text{ and } \frac{3}{4}l \leq x \leq l$$
- and $V = \infty$ elsewhere (where V_0 is a constant). Plot V versus x . Use the trial variation function $\phi_1 = (2/l)^{1/2} \sin(\pi x/l)$ for $0 \leq x \leq l$ to estimate the ground-state energy for $V_0 = \hbar^2/ml^2$ and compare with the true ground-state energy $E = 5.750345 \hbar^2/ml^2$. To save time in evaluating integrals, note that $\langle \phi_1 | \hat{H} | \phi_1 \rangle = \langle \phi_1 | \hat{T} | \phi_1 \rangle + \langle \phi_1 | \hat{V} | \phi_1 \rangle$ and explain why $\langle \phi_1 | \hat{T} | \phi_1 \rangle$ equals the particle-in-a-box ground-state energy $\hbar^2/8ml^2$. (b) For this system, use the variation function $\phi_2 = x(l-x)$. To save time, note that $\langle \phi_2 | \hat{T} | \phi_2 \rangle$ is given by Eq. (8.12). (Why?)
- 8.3** Verify the result for $\langle \phi | \hat{H} | \phi \rangle$ in the last example of Section 8.1.
- 8.4** If the normalized variation function $\phi = (3/l^3)^{1/2} x$ for $0 \leq x \leq l$ is applied to the particle-in-a-one-dimensional-box problem, one finds that the variational integral equals zero, which is less than the true ground-state energy. What is wrong?
- 8.5** For a particle in a three-dimensional box with sides of length a, b, c , write down the variation function that is the three-dimensional extension of the variation function $\phi = x(l-x)$ used in Section 8.1 for the particle in a one-dimensional box. Use the integrals in Eqs. (8.12) and (8.13) to evaluate the variational integral for the three-dimensional case. Find the percent error in the ground-state energy.
- 8.6** (a) A particle in a spherical box of radius b has $V = 0$ for $0 \leq r \leq b$ and $V = \infty$ for $r > b$. Use the trial function $\phi = b - r$ for $0 \leq r \leq b$ and $\phi = 0$ for $r > b$ to estimate the ground-state energy and compare with the true value $\hbar^2/8mb^2$ (Prob. 6.2). (b) Devise another

simple variation function that obeys the boundary conditions for this problem and work out the percent error in the ground-state energy given by your function.

- 8.7** Application of the variation function $\phi = e^{-cx^2}$ (where c is a variational parameter) to a problem with $V = af(x)$, where a is a positive constant and $f(x)$ is a certain function of x , gives the variational integral as $W = \hbar^2/2m + 15a/64c^3$. Find the minimum value of W (in terms of a and physical constants) for this variation function.
- 8.8** Consider a particle moving in one dimension with $V(x) = \infty$ for $x < 0$ and $V(x) = bx$ for $x \geq 0$, where b is a positive constant. (With $b = mg$, this is a particle bouncing on an impenetrable surface in a gravitational field.) (a) What must be true about ψ at $x = 0$? (b) Apply the variation function $\phi = xe^{-cx}$ for $x \geq 0$ and $\phi = 0$ for $x < 0$ to this problem. Find the optimum value of the variational parameter c and compare your energy result with the accurate ground-state energy value $1.855757081(\hbar^2/3b^{2/3}/m^{1/3})$. (See also Prob. 4.34.) Rather amazingly, quantized states of a neutron in the Earth's gravitational field have been observed: See the references in arxiv.org/abs/hep-ph/0602093.
- 8.9** Consider the variation functions $\phi_1 = af + bg$ and $\phi_2 = f + cg$, where f and g in ϕ_1 are the same functions as f and g in ϕ_2 , and a , b , and c are parameters whose values are chosen to minimize the variational integral W . Explain why ϕ_1 and ϕ_2 will give the same result for W when applied to the same problem.
- 8.10** Apply the variation function $\phi = e^{-cr}$ to the hydrogen atom. Choose the parameter c (which is real) to minimize the variational integral, and calculate the percent error in the ground-state energy.
- 8.11** A one-dimensional quartic oscillator has $V = cx^4$, where c is a constant. Devise a variation function with a parameter for this problem, and find the optimum value of the parameter to minimize the variational integral and estimate the ground-state energy in terms of c . Compare with the Numerov-method ground-state energy found in Prob. 4.32.
- 8.12** For a particle in a box of length l , use the variation function $\phi = x^k(l-x)^k$ for $0 \leq x \leq l$. You will need the integral

$$\int_0^l x^s(l-x)^t dx = l^{s+t+1} \frac{\Gamma(s+1)\Gamma(t+1)}{\Gamma(s+t+2)}$$

where the gamma function obeys the relation $\Gamma(z+1) = z\Gamma(z)$. The definition of the gamma function $\Gamma(z)$ need not concern you, since the gamma functions will ultimately cancel. (a) Show that the variational integral equals $(\hbar^2/ml^2)(4k^2+k)/(2k-1)$. (b) Find the optimum value of k and calculate the percent error in the ground-state energy for this k .

- 8.13** Consider a one-particle, one-dimensional system with $V = 0$ for $0 \leq x \leq l$ and $V = V_0$ elsewhere (Fig. 2.5). (a) Use the variation function $\phi = \sin[\pi(x+c)/(l+2c)]$ for $-c \leq x \leq l+c$ and $\phi = 0$ elsewhere, where c is a positive variational parameter. Sketch ϕ and V on the same plot. Choose c to minimize the variational integral W and find the expression for W . (Hint: Use a simple substitution to put the integrals in the form of integrals in the Appendix.) (b) Find W for $V_0 = 20\hbar^2/ml^2$ and compare with the true ground-state energy $2.814\hbar^2/ml^2$ (Prob. 4.31c).
- 8.14** Prove that, for a system with a nondegenerate ground state, $\int \phi^* \hat{H} \phi d\tau > E_1$, if ϕ is any normalized, well-behaved function that is not equal to the true ground-state wave function. *Hint:* Let b be a positive constant such that $E_1 + b < E_2$. Turn (8.4) into an inequality by replacing all E_k 's except E_1 with $E_1 + b$. (The notation $E_1, E_2, \dots$, is as in Section 8.2.)
- 8.15** A paper published in 1971 applied the normalized variation function $N \exp(-br^2/a_0^2 - cr/a_0)$ to the hydrogen atom and stated that minimization of the variational integral with respect to the parameters b and c yielded an energy 0.7% above the true ground-state energy for infinite nuclear mass. Without doing any calculations, state why this result must be in error.
- 8.16** (a) Use the triangular function (7.35) as a variation function for the ground state of the particle in a box. Note that $f''(x)$ is infinite at $x = \frac{1}{2}l$ because of the discontinuity in $f'(x)$ at this point. Therefore, in evaluating $\int f^* \hat{H} f dx$, we run into difficulty in evaluating the integral of ff'' . One way around this problem is to first show that

$$\int_0^l ff'' dx = - \int_0^l (f')^2 dx = - \int_0^{l/2} (f')^2 dx - \int_{l/2}^l (f')^2 dx \quad (8.96)$$

for any function obeying the boundary conditions. Then, using the expression on the right of (8.96), we can calculate the variational integral. Prove (8.96), and then calculate the percent error in the ground-state energy, using this triangular function. Note that there is no parameter in this trial function. [If you are ambitious, try this alternative procedure: Note that $f'(x)$ involves the Heaviside step function (Section 7.7), and therefore $f''(x)$ involves the Dirac delta function. Use the properties of the delta function to evaluate $\int_0^l f f'' dx$, and find the percent error using this triangular function.] (b) The variation function ϕ of the first example in Section 8.1 has discontinuities in ϕ' at $x = 0$ and $x = l$, so, strictly speaking, we should use one of the procedures of part (a) of this problem to evaluate $\langle \phi | \hat{H} | \phi \rangle$. Do this and show that the same value of $\langle \phi | \hat{H} | \phi \rangle$ is obtained.

- 8.17** (a) For the ground state of the hydrogen atom, use the Gaussian trial function $\phi = \exp(-cr^2/a_0^2)$. Find the optimum value of c and the percent error in the energy. (Gaussian variational functions are widely used in molecular quantum mechanics; see Section 15.4.) (b) Multiply the function in (a) by the spherical harmonic Y_2^0 , and then minimize the variational integral. This yields an upper bound to the energy of which hydrogen-atom state?
- 8.18** For the system of Prob. 8.2, explain why the variation function $\phi = (2/l)^{1/2} \sin(2\pi x/l)$ gives an upper bound to the energy E_2 of the first excited state. (*Hint:* With the origin at the center of the box, V is an even function.) Use this ϕ to evaluate the variation integral for $V_0 = \hbar^2/ml^2$ and compare with the true value $E_2 = 20.23604\hbar^2/ml^2$.

8.19 Evaluate

$$(a) \begin{vmatrix} 3 & 1 & i \\ -2 & 4 & 0 \\ 5 & 7 & \frac{1}{2} \end{vmatrix} \quad (b) \begin{vmatrix} 2 & 5 & 1 & 3 \\ 8 & 0 & 4 & -1 \\ 6 & 6 & 6 & 1 \\ 5 & -2 & -2 & 2 \end{vmatrix}$$

- 8.20** (a) Prove that the value of a determinant all of whose elements below the principal diagonal are zero is equal to the product of the diagonal elements. (Such a determinant is called *upper triangular*.) (b) How many terms are there in the expansion of an n th-order determinant?
- 8.21** Verify the block-diagonal determinant equation (8.34).
- 8.22** (a) Consider some permutation of the integers $1, 2, 3, \dots, n$. The permutation is an *even* permutation if an even number of interchanges of pairs of integers restores the permutation to the natural order $1, 2, 3, \dots, n$. An *odd permutation* requires an odd number of interchanges of pairs to reach the natural order. For example, the permutation 3124 is even, since two interchanges restore it to the natural order: $3124 \rightarrow 1324 \rightarrow 1234$. Write down and classify (even or odd) all permutations of 123. (b) Verify that the definition (8.24) of the third-order determinant is equivalent to

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum (\pm 1) a_{1i} a_{2j} a_{3k}$$

where ijk is one of the permutations of the integers 123, the sum is over the $3!$ different permutations of these integers, and the sign of each term is plus or minus, depending on whether the permutation is even or odd. (c) How would we define the n th-order determinant using this type of definition?

8.23 Use Gaussian elimination with detached coefficients to solve

$$\begin{aligned} 2x_1 - x_2 + 4x_3 + 2x_4 &= 16 \\ 3x_1 - x_3 + 4x_4 &= -5 \\ 2x_1 + x_2 + x_3 - 2x_4 &= 8 \\ -4x_1 + 6x_2 + 2x_3 + x_4 &= 3 \end{aligned}$$

8.24 When Gaussian elimination is programmed for a computer (or done on an electronic calculator), one should use exchanges of equations to avoid dividing by a coefficient that is much

smaller in absolute value than the other coefficients in the equations. Explain why division by an extremely small coefficient can lead to large errors in the results.

- 8.25** Gaussian elimination can be used to efficiently evaluate a determinant, as follows. Divide each element of row 1 of the determinant (8.20) by a_{11} and place the factor a_{11} in front of the determinant (Theorem IV of Section 8.3). Then subtract the appropriate multiples of the row 1 elements from row 2, row 3, . . . , row n to make $a_{21}, a_{31}, \dots, a_{n1}$ zero (Theorem V). Next divide the second-row elements by the current value of a_{22} and insert the factor a_{22} in front of the determinant, and so on. Ultimately, we get a determinant all of whose elements below the principal diagonal are zero. From Prob. 8.20, this determinant equals the product of its diagonal elements. Use this procedure to evaluate the determinant in Prob. 8.19b.

- 8.26** Solve each set of simultaneous equations, where $i = \sqrt{-1}$:

$$\begin{array}{ll} \text{(a)} & \begin{array}{l} 8x - 15y = 0 \\ -3x + 4y = 0 \end{array} \\ \text{(b)} & \begin{array}{l} -4x + 3iy = 0 \\ 5ix + \frac{15}{4}y = 0 \end{array} \end{array}$$

- 8.27** Solve each set of simultaneous equations using Gauss–Jordan elimination with detached coefficients or some other method:

$$\begin{array}{ll} & \begin{array}{l} x + 2y + 3z = 0 \\ 3x + y + 2z = 0 \\ 2x + 3y + z = 0 \end{array} \\ \text{(a)} & \begin{array}{l} x + 2y + 3z = 0 \\ x - y + z = 0 \\ 7x - y + 11z = 0 \end{array} \\ \text{(b)} & \end{array}$$

- 8.28** Write a computer program that uses Gaussian elimination to solve a system of n linear, simultaneous, inhomogeneous equations in n unknowns, where $n \leq 10$. Test it on a couple of examples.

- 8.29** Consider the linear variation function $\phi = \sum_i c_i f_i$. State whether each of the following statements about the basis functions f_i is true or false. (a) Each f_i must be normalized. (b) Each f_i must be well behaved. (c) The functions $\{f_i\}$ must be orthogonal to one another. (d) Each f_i must be an eigenfunction of the system's Hamiltonian operator.

- 8.30** When the variation function $\phi = c_1 f_1 + c_2 f_2$ is applied to a certain quantum-mechanical problem, one finds $\langle f_1 | \hat{H} | f_1 \rangle = 4a$, $\langle f_1 | \hat{H} | f_2 \rangle = a$, $\langle f_2 | \hat{H} | f_2 \rangle = 6a$, $\langle f_1 | f_1 \rangle = 2b$, $\langle f_2 | f_2 \rangle = 3b$, $\langle f_1 | f_2 \rangle = b$, where a and b are known positive constants. Use this ϕ to find (in terms of a and b) upper bounds to the lowest two energies, and for each W , find c_1 and c_2 for the normalized ϕ .

- 8.31** Solve the second-order secular equation (8.56) for the special case where $H_{11} = H_{22}$ and $S_{11} = S_{22}$. (Reminder: f_1 and f_2 are real functions.) Then solve for c_1/c_2 for each of the two roots W_1 and W_2 .

- 8.32** For the system of Prob. 8.2 with $V_0 = \hbar^2/ml^2$ (a particle in a box with a rectangular bump in the middle), consider the linear variation function

$$\phi = c_1 f_1 + c_2 f_2 = c_1 (2/l)^{1/2} \sin(\pi x/l) + c_2 (2/l)^{1/2} \sin(3\pi x/l)$$

for $0 \leq x \leq l$. (a) Explain why this variation function will give upper bounds to the energies E_1 and E_3 in (8.60). (b) Explain why f_1 and f_2 are orthonormal. (c) Note that $\langle f_i | \hat{H} | f_j \rangle = \langle f_i | \hat{T} | f_j \rangle + \langle f_i | V | f_j \rangle$. Explain why $\hat{T} f_j = \epsilon_j f_j$ (for $0 \leq x \leq l$), and give the $\hat{T}$ eigenvalues ϵ_1 and ϵ_2 . Show that $\langle f_i | \hat{T} | f_j \rangle = \delta_{ij} \epsilon_j$. (d) Using the results of parts (b) and (c) to help evaluate the integrals, set up and solve the secular equation. Then find the variation functions ϕ_1 and ϕ_2 that correspond to W_1 and W_2 . Compare W_1 and W_2 with the true energies $E_1 = 5.750345\hbar^2/ml^2$ and $E_3 = 44.808373\hbar^2/ml^2$. Compare W_1 with the value $W = 5.753112\hbar^2/ml^2$ found in Prob. 8.2 using $\phi = f_1$. (e) If we want to improve on the results of part (d) by using a three-term linear variation function, what would be a logical choice for f_1, f_2 , and f_3 ?

- 8.33** Apply the linear variation function

$$\phi = c_1 x^2(l-x) + c_2 x(l-x)^2, \quad 0 \leq x \leq l$$

to the particle in a one-dimensional box. (See Prob. 8.31.) Calculate the percent errors for the $n = 1$ and $n = 2$ energies. Sketch $x^2(l-x)$, $x(l-x)^2$, and the two approximate wave

functions you find. (To help sketch the functions, find the nodes and the maxima and minima of each function.)

- 8.34** Show that if the change of variable $x' \equiv x - \frac{1}{2}l$ (corresponding to shifting the origin to the center of the box) is made in the functions (8.62), then f_1 and f_2 are even functions of x' and f_3 and f_4 are odd.
- 8.35** Verify the values given for H_{12} , H_{22} , S_{12} , and S_{22} in the example in Section 8.5.
- 8.36** Derive the roots (8.72) of the odd-function secular equation in the Section 8.5 example.
- 8.37** Use a spreadsheet or a program like Mathcad to plot the percent deviation of the variation function (8.74) from the true ground-state wave function versus x/l .
- 8.38** Derive the functions ϕ_2 and ϕ_3 in Eq. (8.75) of the Section 8.5 example.
- 8.39** Let the variation function of (8.40) be complex. Then $c_j = a_j + ib_j$, where a_j and b_j are real numbers. There are $2n$ parameters to be varied, namely, the a_j 's and b_j 's. (a) Use the chain rule to show that the minimization conditions $\partial W/\partial a_i = 0$, $\partial W/\partial b_i = 0$ are equivalent to the conditions $\partial W/\partial c_i = 0$, $\partial W/\partial c_i^* = 0$. (b) Show that minimization of W leads to Eq. (8.53) and its complex conjugate, which may be discarded. Hence Eqs. (8.53) and (8.58) are valid for complex variation functions.
- 8.40** We wish to prove that the approximate wave functions obtained in the linear variation method are orthogonal and that the approximate energies obtained are upper bounds to the energies of the n lowest states. Let the approximate function ϕ_α have the value W_α for the variational integral and the coefficients $c_j^{(\alpha)}$ in (8.40). (We add α to distinguish the n different ϕ 's.) We rewrite (8.53) as

$$\sum_k [(\langle f_i | \hat{H} | f_k \rangle - \langle f_i | f_k \rangle W_\alpha) c_k^{(\alpha)}] = 0, \quad i = 1, \dots, n \quad (8.97)$$

(a) Show that $\langle f_i | \hat{H} | \phi_\alpha \rangle = 0$ by showing this integral to equal the left side of (8.97). (b) Use the result of (a) to show $\langle \phi_\beta | \hat{H} - W_\alpha | \phi_\alpha \rangle = 0$ and $\langle \phi_\alpha | \hat{H} - W_\beta | \phi_\beta \rangle^* = 0$ for all α and β . (c) Equate the two integrals in (b), and use the Hermitian property of $\hat{H}$ to show that $\langle \phi_\beta | \phi_\alpha \rangle (W_\beta - W_\alpha) = 0$. We conclude that, for $W_\alpha \neq W_\beta$, ϕ_α and ϕ_β are orthogonal. (For $W_\alpha = W_\beta$, we can form orthogonal linear combinations of ϕ_α and ϕ_β that will have the same value for the variational integral.) (d) Let $\phi_1, \phi_2, \dots, \phi_n$ be the normalized approximate wave functions found in the variation method, where the functions are listed in order of increasing value of the variational integral. Consider the function $g = \sum_{\alpha=1}^m b_\alpha \phi_\alpha$, where $m \leq n$ and the coefficients b_α are chosen so that $\langle g | \psi_1 \rangle = 0$, $\langle g | \psi_2 \rangle = 0, \dots, \langle g | \psi_{m-1} \rangle = 0$ and so that g is normalized. Here $\psi_1, \psi_2, \dots$ are the true wave functions of the lowest-energy states. Explain why $\langle g | \hat{H} | g \rangle \geq E_m$. (See Section 8.2.) (e) Use the results of parts (b) and (c) to show that $\langle \phi_\alpha | \hat{H} | \phi_\beta \rangle = 0$ for $\alpha \neq \beta$. (f) Use the result of (e) to show that $\langle g | \hat{H} | g \rangle = \sum_{\alpha=1}^m |b_\alpha|^2 W_\alpha$. (g) Use the orthonormality of the ϕ 's to show that $\sum_{\alpha=1}^m |b_\alpha|^2 = 1$. (h) Use the results of (f) and (g) to show that $\langle g | \hat{H} | g \rangle \leq W_m$. *Hint:* See Eq. (8.5). (i) Combine (h) and (d) to give the desired result: $W_m \geq E_m$ for $m = 1, 2, \dots, n$.

- 8.41** Find $\mathbf{A}^*$, $\mathbf{A}^T$, and $\mathbf{A}^\dagger$ if

$$\mathbf{A} = \begin{pmatrix} 7 & 3 & 0 \\ 2 - i & 2i & i \\ 1 + i & 4 & 2 \end{pmatrix}$$

- 8.42** (a) Which of the following matrices are real? (b) Which are symmetric? (c) Which are Hermitian?

$$\mathbf{B} = \begin{pmatrix} i & 2 + i \\ 2 - i & 3 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} i & 2 + i \\ 2 + i & -1 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} 3 & 2 - i \\ 2 + i & -1 \end{pmatrix} \quad \mathbf{F} = \begin{pmatrix} 5 & 3 \\ 3 & -1 \end{pmatrix}$$

- 8.43** Verify the orthonormality equation (8.93) for the column vectors of a unitary matrix.
- 8.44** If the functions v and w are normalized and orthogonal, and if v and w are expanded in terms of the complete, orthonormal set $\{f_i\}$ as $v = \sum_i v_i f_i$ and $w = \sum_i w_i f_i$ (where the expansion coefficients v_i and w_i are constants), show that the column vectors $\mathbf{v}$ and $\mathbf{w}$ that consist of the

expansion coefficients $v_1, v_2, \dots$ and $w_1, w_2, \dots$, respectively, are normalized and orthogonal, as defined by (8.83) and (8.84).

- 8.45** Without using a computer, find the eigenvalues and normalized eigenvectors of each of these matrices. Begin by solving the characteristic equation. Check that the sum of the eigenvalues equals the trace of the matrix.

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 3 & 2 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 2 & 0 \\ 9 & 2 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$$

- 8.46** If $\mathbf{A}$ is a diagonal square matrix of order three with unequal diagonal elements a_{11}, a_{22}, a_{33} , find the normalized eigenvectors of $\mathbf{A}$.

- 8.47** (a) Without using a computer, find the eigenvalues and normalized eigenvectors of

$$\mathbf{A} = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$$

(b) Is $\mathbf{A}$ real and symmetric? Is $\mathbf{A}$ Hermitian? (c) Is the eigenvector matrix $\mathbf{C}$ orthogonal? Is the eigenvector matrix $\mathbf{C}$ unitary? (d) Write down $\mathbf{C}^{-1}$ without doing any calculations. (e) Verify that $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ equals the diagonal matrix of eigenvalues. (f) Verify Eq. (8.80) for each eigenvector.

- 8.48** For

$$\mathbf{A} = \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix}$$

find the eigenvalues and normalized eigenvectors and answer the questions of Prob. 8.47(b)–(e).

- 8.49** Find the eigenvector $\mathbf{c}^{(2)}$ for the eigenvalue λ_2 in the example in Section 8.6.

- 8.50** Find the eigenvalues and normalized eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & -2 \\ 0 & 5 & 0 \\ -2 & 4 & 2 \end{pmatrix}$$

You can avoid solving a cubic equation if you expand the characteristic determinant in the simplest way and do not fully multiply out the characteristic polynomial.

- 8.51** An efficient way to calculate the inverse of a square matrix $\mathbf{A}$ of order n is as follows: (a) Place the n th-order unit matrix $\mathbf{I}$ at the right of the matrix $\mathbf{A}$ to form an n -row, $2n$ -column array, which we denote by $(\mathbf{A}:\mathbf{I})$. (b) Perform Gauss–Jordan elimination on the rows of $(\mathbf{A}:\mathbf{I})$ so as to reduce the $\mathbf{A}$ portion of $(\mathbf{A}:\mathbf{I})$ to the unit matrix. At the end of this process, the array will have the form $(\mathbf{I}:\mathbf{B})$. The matrix $\mathbf{B}$ is $\mathbf{A}^{-1}$. (If $\mathbf{A}^{-1}$ does not exist, it will be impossible to reduce the $\mathbf{A}$ portion of the array to $\mathbf{I}$.) Use this procedure to find the inverse of the matrix in Prob. 8.50.

- 8.52** Use a computer-algebra system to find the eigenvalues and normalized eigenvectors of the square matrix $\mathbf{B}$ of order six with elements $b_{jk} = (j^2 + k^2)/(j + k)$.

- 8.53** Excel can find eigenvalues and eigenvectors of real, symmetric matrices as follows. If the eigenvalues of the n th-order real, symmetric matrix $\mathbf{H}$ are arranged in increasing order: $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, an extension of a theorem due to Rayleigh and Ritz states that $\lambda_1 = \min(\mathbf{x}^T \mathbf{H} \mathbf{x} / \mathbf{x}^T \mathbf{x})$, where $\mathbf{x}$ is an n th-order nonzero column vector whose elements are varied to minimize the quantity in parentheses; also, $\lambda_2 = \min(\mathbf{y}^T \mathbf{H} \mathbf{y} / \mathbf{y}^T \mathbf{y})$ if $\mathbf{y}^T \mathbf{c}_1 = 0$, where $\mathbf{c}_1$ is the eigenvector corresponding to λ_1 ; $\lambda_3 = \min(\mathbf{z}^T \mathbf{H} \mathbf{z} / \mathbf{z}^T \mathbf{z})$ if $\mathbf{z}^T \mathbf{c}_1 = 0$ and $\mathbf{z}^T \mathbf{c}_2 = 0$, where $\mathbf{c}_2$ is the eigenvector corresponding to λ_2 ; and so on. (Note the close analogy with the results of Sections 8.1 and 8.2.) Use this theorem to have Excel find the eigenvalues and normalized eigenvectors of the matrix in Prob. 8.47. *Hints:* Assign names in Excel to the various matrices involved. To multiply matrices $\mathbf{A}$ and $\mathbf{B}$, select an appropriately sized rectangular array of cells where you want the matrix product to appear; then type =MMULT(A,B) and press the Control, Shift, and Enter keys simultaneously. The transpose of matrix $\mathbf{C}$ is found similarly using the formula =TRANSPOSE(C). To find λ_1 , start with a guess for $\mathbf{x}$ and use the Solver to vary $\mathbf{x}$ so as to minimize $\mathbf{x}^T \mathbf{H} \mathbf{x}$ subject to the constraint that $\mathbf{x}^T \mathbf{x} = 1$. After you find the first

eigenvalue and eigenvector, add the relevant orthogonality constraint to the Solver and find the next eigenvalue and eigenvector; and so on.

- 8.54** For a matrix of order 20 whose eigenvalues are $1, 2, 3, \dots, 20$, the characteristic equation can be written as $\prod_{m=1}^{20} (\lambda - m) = 0$, where the product notation is defined by $\prod_{k=1}^n b_k \equiv b_1 b_2 \cdots b_n$. Use a spreadsheet or computer-algebra system to graph the characteristic polynomial for λ in the range $\lambda = 0.9$ to 20.1 . Choose the scale on the vertical axis so that the points where the curve crosses the horizontal axis are clearly visible. Now add the quantity $1 \times 10^{-8} \lambda^{19}$ to the characteristic polynomial, graph the altered characteristic polynomial, and notice what has happened to the roots of the characteristic polynomial. If you see fewer than 20 roots for the altered polynomial, explain where the missing roots are. In the original characteristic polynomial, the coefficient of λ^{19} is the sum of the integers 1 to 20, which is 210, so the change made was less than one part in 10^{10} in the λ^{19} coefficient, yet it caused a major change in the eigenvalues.
- 8.55** If $\mathbf{B} = \mathbf{M}^{-1} \mathbf{A} \mathbf{M}$, prove that the square matrices $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues. Also, express the eigenvectors of $\mathbf{B}$ in terms of the eigenvectors of $\mathbf{A}$. *Hints:* Start with the eigenvalue equation for $\mathbf{A}$, replace $\mathbf{A}$ by its expression in terms of $\mathbf{B}$, and multiply the resulting equation by $\mathbf{M}^{-1}$ on the left to obtain the eigenvalue equation for $\mathbf{B}$.
- 8.56** If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $\mathbf{A}$, find the eigenvalues and eigenvectors of $\mathbf{A}^2$.
- 8.57** This problem deals with the matrix solution of the linear-variation method when the basis functions are nonorthogonal. (a) If $\{f_i\}$ in $\phi = \sum_{i=1}^n c_i f_i$ is not an orthonormal set, we take linear combinations of the functions $\{f_i\}$ to get a new set of functions $\{g_m\}$ that are orthonormal. We have $g_m = \sum_k a_{km} f_k$, $m = 1, 2, \dots, n$, where the coefficients a_{km} are constants and where $\langle g_j | g_m \rangle = \delta_{jm}$. [One procedure to choose the a_{km} coefficients is the Schmidt method (Section 7.2); another method is discussed in Prob. 8.58.] (a) In $\langle g_j | g_m \rangle = \delta_{jm}$, substitute the summation expression for each g and show that the resulting equation is equivalent to the matrix equation $\mathbf{A}^\dagger \mathbf{S} \mathbf{A} = \mathbf{I}$, where $\mathbf{I}$ is a unit matrix, $\mathbf{S}$ is the overlap matrix with elements $S_{jk} = \langle f_j | f_k \rangle$, and $\mathbf{A}$ is the matrix of coefficients a_{km} . (b) Verify that the set of equations (8.53) can be written as $\mathbf{H} \mathbf{c} = \mathbf{W} \mathbf{S} \mathbf{c}$, where $\mathbf{c}$ is the column vector of coefficients $c_1, c_2, \dots, c_n$. As we did in going from (8.79c) to (8.95), we introduce the index i to label the various eigenvalues and eigenvectors and we write $\mathbf{H} \mathbf{c} = \mathbf{W} \mathbf{S} \mathbf{c}$ as $\mathbf{H} \mathbf{c}^{(i)} = \mathbf{W} \mathbf{S} \mathbf{c}^{(i)}$ for $i = 1, 2, \dots, n$. Verify that $\mathbf{H} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{W}$, where $\mathbf{C}$ and $\mathbf{W}$ are the matrices in (8.86). (c) Since $\mathbf{A} \mathbf{A}^{-1} = \mathbf{I}$, we can write $\mathbf{H} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{W}$ as $\mathbf{H} \mathbf{A} \mathbf{A}^{-1} \mathbf{C} = \mathbf{S} \mathbf{A} \mathbf{A}^{-1} \mathbf{C} \mathbf{W}$. Multiply this last equation by $\mathbf{A}^\dagger$ on the left. Use the result of (a) to show that we get $\mathbf{H}' \mathbf{C}' = \mathbf{C}' \mathbf{W}$, where $\mathbf{C}' \equiv \mathbf{A}^{-1} \mathbf{C}$ and $\mathbf{H}' \equiv \mathbf{A}^\dagger \mathbf{H} \mathbf{A}$. Comparison with (8.87) shows that $\mathbf{H}' \mathbf{C}' = \mathbf{C}' \mathbf{W}$ is the eigenvalue equation for the $\mathbf{H}'$ matrix. The matrix procedure to solve the linear variation problem $\mathbf{H} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{W}$ with nonorthogonal basis functions is then as follows: (1) Compute the matrix elements of $\mathbf{H}$ and $\mathbf{S}$ using the nonorthogonal basis. (2) Use the overlap integrals and a procedure such as the Schmidt method to find a matrix $\mathbf{A}$ whose elements a_{km} transform the nonorthogonal functions $\{f_i\}$ to orthonormal functions $\{g_i\}$. (3) Calculate $\mathbf{H}'$ using $\mathbf{H}' \equiv \mathbf{A}^\dagger \mathbf{H} \mathbf{A}$. (4) Find the eigenvalues W_i and the eigenvectors $\mathbf{c}'^{(i)}$ of the $\mathbf{H}'$ matrix. (5) Use $\mathbf{C} = \mathbf{A} \mathbf{C}'$ to compute the coefficient matrix $\mathbf{C}$. The eigenvalues W_i found in step 4 and the coefficients found in step 5 are the desired energy estimates and variation-function coefficients.
- 8.58** The **symmetric** (or **Löwdin**) **orthogonalization** procedure is often used to orthogonalize a basis set. Given the nonorthogonal basis set $\{f_i\}$, we form the set of functions $\{g_m\}$ as the linear combinations $g_m = \sum_k a_{km} f_k$. As shown in Prob. 8.57(a), for $\{g_m\}$ to be an orthogonal set, the matrix of transformation coefficients a_{km} must satisfy $\mathbf{A}^\dagger \mathbf{S} \mathbf{A} = \mathbf{I}$, where $\mathbf{S}$ is the overlap matrix with $S_{jk} = \langle f_j | f_k \rangle$. If the square matrix $\mathbf{B}$ satisfies $\mathbf{B}^2 = \mathbf{S}$, then $\mathbf{B}$ is a *square root* of $\mathbf{S}$, which is written as $\mathbf{B} = \mathbf{S}^{1/2}$. (A matrix can have more than one square root.) Thus $\mathbf{S}^{1/2}$ satisfies $\mathbf{S}^{1/2} \mathbf{S}^{1/2} = \mathbf{S}$. The inverse of $\mathbf{S}^{1/2}$ is written as $\mathbf{S}^{-1/2}$. By the definition of the inverse, $\mathbf{S}^{-1/2} \mathbf{S}^{1/2} = \mathbf{S}^{1/2} \mathbf{S}^{-1/2} = \mathbf{I}$. (a) Since $\mathbf{S}$ is a Hermitian matrix, it can be diagonalized by the unitary matrix $\mathbf{U}$ of its orthonormal eigenvectors, and we have $\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{s}$, where $\mathbf{s}$ is the diagonal matrix of eigenvalues of $\mathbf{S}$. Show that $\mathbf{S} = \mathbf{U} \mathbf{s} \mathbf{U}^\dagger$. (b) Let s_i denote the eigenvalues of $\mathbf{S}$. Let $\mathbf{s}$ and $\mathbf{s}^{-1/2}$ denote the diagonal matrices whose diagonal elements are s_i and $s_i^{-1/2}$, respectively, where $s_i^{-1/2}$ denotes the reciprocals of the *positive* square roots of the $\mathbf{S}$

eigenvalues. [The eigenvalues of $\mathbf{S}$ can all be proven to be positive (see *Szabo and Ostlund*, p. 143), so no s_i is zero and all the numbers $s_i^{-1/2}$ exist.] Show that the matrix $\mathbf{M} \equiv \mathbf{U}\mathbf{s}^{1/2}\mathbf{U}^\dagger$ satisfies $\mathbf{M}^2 = \mathbf{U}\mathbf{S}\mathbf{U}^\dagger = \mathbf{S}$. Hence $\mathbf{M}$ is a square root of $\mathbf{S}$; $\mathbf{M} = \mathbf{S}^{1/2}$. (c) Show that the matrix $\mathbf{N} \equiv \mathbf{U}\mathbf{s}^{-1/2}\mathbf{U}^\dagger$ satisfies $\mathbf{M}\mathbf{N} = \mathbf{I}$. Since $\mathbf{M} = \mathbf{S}^{1/2}$, we have $\mathbf{S}^{1/2}\mathbf{N} = \mathbf{I}$, so $\mathbf{N} = \mathbf{S}^{-1/2}$ and $\mathbf{S}^{-1/2} = \mathbf{U}\mathbf{s}^{-1/2}\mathbf{U}^\dagger$. (d) The symmetric orthogonalization procedure takes the transformation matrix $\mathbf{A}$ to be $\mathbf{S}^{-1/2}$; that is, $\mathbf{A} = \mathbf{S}^{-1/2} = \mathbf{U}\mathbf{s}^{-1/2}\mathbf{U}^\dagger$. To show that this choice satisfies the requirement $\mathbf{A}^\dagger\mathbf{S}\mathbf{A} = \mathbf{I}$, we need to find $\mathbf{A}^\dagger$. Prove that $(\mathbf{B}\mathbf{C})^\dagger = \mathbf{C}^\dagger\mathbf{B}^\dagger$ by finding the (i, j) th elements of $(\mathbf{B}\mathbf{C})^\dagger$ and $\mathbf{C}^\dagger\mathbf{B}^\dagger$. Then set $\mathbf{C} = \mathbf{D}\mathbf{E}$ and show that $(\mathbf{B}\mathbf{D}\mathbf{E})^\dagger = \mathbf{E}^\dagger\mathbf{D}^\dagger\mathbf{B}^\dagger$. Next, show that $\mathbf{A}^\dagger = \mathbf{U}\mathbf{s}^{-1/2}\mathbf{U}^\dagger$. Then show that $\mathbf{A}^\dagger\mathbf{S}\mathbf{A} = \mathbf{I}$. Hence, to use symmetric orthogonalization, we find the eigenvalues and orthonormal eigenvectors of $\mathbf{S}$, use the eigenvalues s_i to form the matrix $\mathbf{s}^{-1/2}$, use the eigenvectors to form $\mathbf{U}$, calculate the transformation matrix $\mathbf{A} = \mathbf{U}\mathbf{s}^{-1/2}\mathbf{U}^\dagger$, and take the orthonormal functions as $g_m = \sum_k a_{km}f_k$.

- 8.59** Suppose $\hat{A}g_n = a_ng_n$. Let the eigenfunctions g_n be expanded in terms of the complete orthonormal set $\{f_i\}$ according to $g_n = \sum_k c_k^{(n)}f_k$. Substitute this expansion into $\hat{A}g_n = a_ng_n$, multiply by f_i^* , integrate over all space, and show that the set of equations $\sum_k (A_{ik} - a_n\delta_{ik})c_k^{(n)} = 0$ for $i = 1, 2, 3, \dots$ is obtained, where $A_{ik} \equiv \langle f_i | \hat{A} | f_k \rangle$. This set of equations has the same form as the set (8.53) with $S_{ik} = \delta_{ik}$. Thus, just as the W values and the coefficients c_k in (8.53) with $S_{ik} = \delta_{ik}$ can be found by finding the eigenvalues and eigenvectors of the $\mathbf{H}$ matrix, the eigenvalues a_n of $\hat{A}$ and the expansion coefficients $c_k^{(n)}$ of the eigenfunctions g_n of $\hat{A}$ can be found by finding the eigenvalues a_n and eigenvectors $\mathbf{c}^{(n)}$ of the matrix $\mathbf{A}$, whose elements are A_{ik} .
- 8.60** Consider the double-well potential with $V = \infty$ for $x < 0$, $V = 0$ for $0 \leq x \leq \frac{1}{4}l$, $V = V_0$ for $\frac{1}{4}l < x < \frac{3}{4}l$, $V = 0$ for $\frac{3}{4}l \leq x \leq l$, and $V = \infty$ for $x > l$, where l and V_0 are positive constants. Use Mathcad or some other computer-algebra program to apply the linear variation method to this problem, taking the basis functions as the m lowest particle-in-a-box (pib) wave functions (2.23). Use dimensionless variables (Section 4.4 and Prob. 4.37). Set things up so that V_0 and m can be readily changed. Use the program's ability to find eigenvalues and eigenvectors to solve for the approximate energies and wave functions. Have the program graph the lowest four approximate wave functions. Show that $T_{jk} = k^2\pi^2\hbar^2/2ml^2$ but let the program evaluate the V_{jk} integrals. *Hint:* In Mathcad, you may need to adjust the value of the TOL variable. (a) For $V_0 = 100\hbar^2/ml^2$, find the lowest four energy levels using the following numbers of pib basis functions: 4; 8; 16; 32. Compare with the following exact reduced-energy values found from Eq. (4.90): 45.802165653, 46.107222914, 113.938076461, and 143.353993916. Which basis functions contribute substantially to the ground-state wave function? Which contribute substantially to the first-excited-state wave function? (b) With four basis functions, you will find that the variational energy of the one-node function that corresponds to the first excited state lies below the variational energy of the even function that corresponds to the ground state. Is this a violation of the variation theorem inequalities (8.59) to (8.61)?
- 8.61** Revise your solution to the double-well problem 8.60 so as to treat even and odd functions separately. Do this by introducing a parameter whose value is 1 or 0, depending on whether we are treating the even or the odd wave functions. Calculate the lowest two Hamiltonian eigenvalues using the first 16 even pib functions. Repeat with the lowest 16 odd pib functions. Compare your results with those of Prob. 8.60.
- 8.62** Consider the one-particle, one-dimensional problem with $V = V_0(\hbar^2/ml^2)x$ for $0 \leq x \leq l$ and $V = \infty$ elsewhere. Modify your solution to Prob. 8.60 to estimate the lowest four energy eigenvalues for $V_0 = 200$ using the following numbers of particle-in-a-box basis functions: (a) 8; (b) 12. Note the appearance of the lowest four wave functions. Which particle-in-a-box states make substantial contributions to the ground state? to the first excited state?
- 8.63** Revise your solution to Prob. 8.60 to treat the one-dimensional harmonic oscillator using particle-in-a-box (pib) basis functions. Recall that in Section 4.4 we found that for $E_r \leq 5$, the wave function can be taken as zero outside the region $-5 \leq x_r \leq 5$, where E_r and x_r are defined by (4.75) and (4.76). Therefore, we shall take the pib basis functions to extend from -5 to 5 , with the center of the "box" at $x_r = 0$. Since the box has a length of 10 units in x_r , we have $f_j = (2/10)^{1/2} \sin[j\pi(x_r + 5)/10]$ for $|x_r| \leq 5$ and $f_j = 0$ elsewhere. You will also need to revise the kinetic- and potential-energy matrix elements. Increase the number of pib

basis functions until all energy values with $E_r < 5$ are accurate to three decimal places. Check the appearance of the lowest three variation functions. Which pib basis functions contribute most to the ground state? to the first excited state?

- 8.64** Revise your solution to Prob. 8.60 to apply the pib basis functions to the one-dimensional quartic oscillator with $V = cx^4$. See Prob. 8.63 for hints. Take the “box” to extend from $x_r = -3.5$ to 3.5 , where x_r is as found in Prob. 4.32. Increase the number of pib basis functions until the lowest three energy values remain stable to three decimal places. Compare the lowest three energies with those found by the Numerov method in Prob. 4.32. Check the appearance of the lowest three variational functions. Now repeat for the box going from $x_r = -4.5$ to 4.5 . For which box length do we get faster convergence to the true energies?
- 8.65** Apply the particle-in-a-box basis functions to the radial equation for the hydrogen atom for the $l = 0$ states. Recall that in Section 6.9, we expressed the radial factor in the H-atom wave function as $R(r) = r^{-1}F(r)$, where $F(r) = 0$ at $r = 0$. The variation function in this problem will have the form $\phi = r^{-1}F(r)Y_l^m(\theta, \phi)$; take the dimensionless function $F_r(r_r)$ to be a linear combination of 28 pib basis functions, where the box goes from $r_r = 0$ to 27 , where r_r is defined in Section 6.9. Work out the proper forms for the integrals H_{jk} and S_{jk} . Find the estimates for the lowest three $l = 0$ energies. For the ground-state variation function, how many pib basis functions appear with a coefficient greater than 0.1?
- 8.66** True or false? (a) The transpose of a column vector is a row vector. (b) The matrix product $\mathbf{bc}$ of a row vector $\mathbf{b}$ with n elements times a column vector $\mathbf{c}$ with n elements is a scalar. (c) The diagonal elements of a Hermitian matrix must be real numbers. (d) The eigenvalues of a diagonal matrix are equal to the diagonal elements. (e) For a square matrix all of whose elements below the principal diagonal are zero, the eigenvalues are equal to the diagonal elements. (See Prob. 8.20.) (f) Every nonzero column vector with n elements is an eigenvector of the unit matrix of order n . (g) Every nonzero linear combination of two eigenvectors that correspond to the same eigenvalue of a matrix is an eigenvector of that matrix. (h) Zero is not allowed as an eigenvalue of a matrix. (i) A zero column vector is not allowed as an eigenvector of a matrix. (j) Every square matrix has an inverse. (k) If $\mathbf{AB} = \mathbf{AC}$, then the matrices $\mathbf{B}$ and $\mathbf{C}$ must be equal. (l) If $\mathbf{A}$ is a square matrix of order n and $\mathbf{c}$ is a column vector with n elements, then $\mathbf{Ac}$ is a column vector with n elements. (m) Every real symmetric matrix is a Hermitian matrix. (n) The sum of the eigenvalues of a square matrix always equals the trace of the matrix.