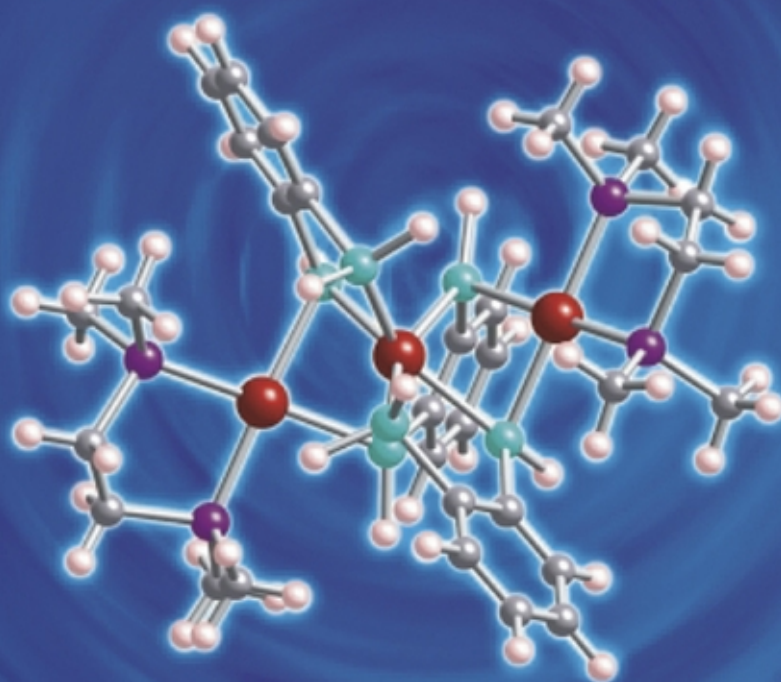


Essentials of  
**Computational  
Chemistry** Second Edition  
**Theories and Models**

Christopher J Cramer



 WILEY

# 4

## Foundations of Molecular Orbital Theory

### 4.1 Quantum Mechanics and the Wave Function

To this point, the models we have considered for representing microscopic systems have been designed based on classical, which is to say, macroscopic, analogs. We now turn our focus to contrasting models, whose foundations explicitly recognize the fundamental difference between systems of these two size extremes. Early practitioners of chemistry and physics had few, if any, suspicions that the rules governing microscopic and macroscopic systems should be different. Then, in 1900, Max Planck offered a radical proposal that blackbody radiation emitted by microscopic particles was limited to certain discrete values, i.e., it was ‘quantized’. Such quantization was essential to reconciling large differences between predictions from classical models and experiment.

As the twentieth century progressed, it became increasingly clear that quantization was not only a characteristic of light, but also of the fundamental particles from which matter is constructed. Bound electrons in atoms, in particular, are clearly limited to discrete energies (levels) as indicated by their ultraviolet and visible line spectra. This phenomenon has no classical correspondence – in a classical system, obeying Newtonian mechanics, energy can vary continuously.

In order to describe microscopic systems, then, a different mechanics was required. One promising candidate was wave mechanics, since standing waves are also a quantized phenomenon. Interestingly, as first proposed by de Broglie, matter can indeed be shown to have wavelike properties. However, it also has particle-like properties, and to properly account for this dichotomy a new mechanics, quantum mechanics, was developed. This chapter provides an overview of the fundamental features of quantum mechanics, and describes in a formal way the fundamental equations that are used in the construction of computational models. In some sense, this chapter is historical. However, in order to appreciate the differences between modern computational models, and the range over which they may be expected to be applicable, it is important to understand the foundation on which all of them are built. Following this exposition, Chapter 5 overviews the approximations inherent

in so-called semiempirical QM models, Chapter 6 focuses on *ab initio* Hartree–Fock (HF) models, and Chapter 7 describes methods for accounting for electron correlation.

We begin with a brief recapitulation of some of the key features of quantum mechanics. The fundamental postulate of quantum mechanics is that a so-called wave function,  $\Psi$ , exists for any (chemical) system, and that appropriate operators (functions) which act upon  $\Psi$  return the observable properties of the system. In mathematical notation,

$$\vartheta \Psi = e \Psi \quad (4.1)$$

where  $\vartheta$  is an operator and  $e$  is a scalar value for some property of the system. When Eq. (4.1) holds,  $\Psi$  is called an eigenfunction and  $e$  an eigenvalue, by analogy to matrix algebra where  $\Psi$  to be an  $N$ -element column vector,  $\vartheta$  to be an  $N \times N$  square matrix, and  $e$  to remain a scalar constant. Importantly, the product of the wave function  $\Psi$  with its complex conjugate (i.e.,  $|\Psi^* \Psi|$ ) has units of probability density. For ease of notation, and since we will be working almost exclusively with real, and not complex, wave functions, we will hereafter drop the complex conjugate symbol ‘\*’. Thus, the probability that a chemical system will be found within some region of multi-dimensional space is equal to the integral of  $|\Psi|^2$  over that region of space.

These postulates place certain constraints on what constitutes an acceptable wave function. For a bound particle, the normalized integral of  $|\Psi|^2$  over all space must be unity (i.e., the probability of finding it somewhere is one) which requires that  $\Psi$  be quadratically integrable. In addition,  $\Psi$  must be continuous and single-valued.

From this very formal presentation, the nature of  $\Psi$  can hardly be called anything but mysterious. Indeed, perhaps the best description of  $\Psi$  at this point is that it is an oracle – when queried with questions by an operator, it returns answers. By the end of this chapter, it will be clear the precise way in which  $\Psi$  is expressed, and we should have a more intuitive notion of what  $\Psi$  represents. However, the view that  $\Psi$  is an oracle is by no means a bad one, and will be returned to again at various points.

## 4.2 The Hamiltonian Operator

### 4.2.1 General Features

The operator in Eq. (4.1) that returns the system energy,  $E$ , as an eigenvalue is called the Hamiltonian operator,  $H$ . Thus, we write

$$H \Psi = E \Psi \quad (4.2)$$

which is the Schrödinger equation. The typical form of the Hamiltonian operator with which we will be concerned takes into account five contributions to the total energy of a system (from now on we will say molecule, which certainly includes an atom as a possibility): the kinetic energies of the electrons and nuclei, the attraction of the electrons to the nuclei, and the interelectronic and internuclear repulsions. In more complicated situations, e.g., in

the presence of an external electric field, in the presence of an external magnetic field, in the event of significant spin–orbit coupling in heavy elements, taking account of relativistic effects, etc., other terms are required in the Hamiltonian. We will consider some of these at later points in the text, but we will not find them necessary for general purposes. Casting the Hamiltonian into mathematical notation, we have

$$H = - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k} \nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} + \sum_{k < l} \frac{e^2 Z_k Z_l}{r_{kl}} \quad (4.3)$$

where  $i$  and  $j$  run over electrons,  $k$  and  $l$  run over nuclei,  $\hbar$  is Planck's constant divided by  $2\pi$ ,  $m_e$  is the mass of the electron,  $m_k$  is the mass of nucleus  $k$ ,  $\nabla^2$  is the Laplacian operator,  $e$  is the charge on the electron,  $Z$  is an atomic number, and  $r_{ab}$  is the distance between particles  $a$  and  $b$ . Note that  $\Psi$  is thus a function of  $3n$  coordinates where  $n$  is the total number of particles (nuclei and electrons), e.g., the  $x$ ,  $y$ , and  $z$  Cartesian coordinates specific to each particle. If we work in Cartesian coordinates, the Laplacian has the form

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \quad (4.4)$$

Note that the Hamiltonian operator in Eq. (4.3) is composed of kinetic energy and potential energy parts. The potential energy terms (the last three) appear exactly as they do in classical mechanics. The kinetic energy for a QM particle, however, is not expressed as  $|\mathbf{p}|^2/2m$ , but rather as the eigenvalue of the kinetic energy operator

$$T = - \frac{\hbar^2}{2m} \nabla^2 \quad (4.5)$$

Note also that, as described in Chapter 1, most of the constants appearing in Eq. (4.3) are equal to 1 when atomic units are chosen.

In general, Eq. (4.2) has *many* acceptable eigenfunctions  $\Psi$  for a given molecule, each characterized by a different associated eigenvalue  $E$ . That is, there is a complete set (perhaps infinite) of  $\Psi_i$  with eigenvalues  $E_i$ . For ease of future manipulation, we may assume without loss of generality that these wave functions are orthonormal, i.e., for a one particle system where the wave function depends on only three coordinates,

$$\iiint \Psi_i \Psi_j dx dy dz = \delta_{ij} \quad (4.6)$$

where  $\delta_{ij}$  is the Kronecker delta (equal to one if  $i = j$  and equal to zero otherwise). Orthonormal actually implies two qualities simultaneously: ‘orthogonal’ means that the integral in Eq. (4.6) is equal to zero if  $i \neq j$  and ‘normal’ means that when  $i = j$  the value of the integral is one. For ease of notation, we will henceforth replace all multiple integrals over Cartesian space with a single integral over a generalized  $3n$ -dimensional volume element  $d\mathbf{r}$ , rendering Eq. (4.6) as

$$\int \Psi_i \Psi_j d\mathbf{r} = \delta_{ij} \quad (4.7)$$

Now, consider the result of taking Eq. (4.2) for a specific  $\Psi_i$ , multiplying on the left by  $\Psi_j$ , and integrating. This process gives

$$\int \Psi_j H \Psi_i d\mathbf{r} = \int \Psi_j E_i \Psi_i d\mathbf{r} \quad (4.8)$$

Since the energy  $E$  is a scalar value, we may remove it outside the integral on the r.h.s. and use Eq. (4.7) to write

$$\int \Psi_j H \Psi_i d\mathbf{r} = E_i \delta_{ij} \quad (4.9)$$

This equation will prove useful later on, but it is worth noting at this point that it also offers a prescription for determining the molecular energy. With a wave function in hand, one simply constructs and solves the integral on the left (where  $i$  and  $j$  are identical and index the wave function of interest). Of course, we have not yet said much about the form of the wave function, so the nature of the integral in Eq. (4.8) is not obvious ... although one suspects it might be unpleasant to solve.

### 4.2.2 The Variational Principle

The power of quantum theory, as expressed in Eq. (4.1), is that if one has a molecular wave function in hand, one can calculate physical observables by application of the appropriate operator in a manner analogous to that shown for the Hamiltonian in Eq. (4.8). Regrettably, none of these equations offers us a prescription for *obtaining* the orthonormal set of molecular wave functions. Let us assume for the moment, however, that we can pick an arbitrary function,  $\Phi$ , which is indeed a function of the appropriate electronic and nuclear coordinates to be operated upon by the Hamiltonian. Since we defined the set of orthonormal wave functions  $\Psi_i$  to be complete (and perhaps infinite), the function  $\Phi$  must be some linear combination of the  $\Psi_i$ , i.e.,

$$\Phi = \sum_i c_i \Psi_i \quad (4.10)$$

where, of course, since we don't yet know the individual  $\Psi_i$ , we certainly don't know the coefficients  $c_i$  either! Note that the normality of  $\Phi$  imposes a constraint on the coefficients, however, deriving from

$$\begin{aligned} \int \Phi^2 d\mathbf{r} = 1 &= \int \sum_i c_i \Psi_i \sum_j c_j \Psi_j d\mathbf{r} \\ &= \sum_{ij} c_i c_j \int \Psi_i \Psi_j d\mathbf{r} \\ &= \sum_{ij} c_i c_j \delta_{ij} \\ &= \sum_i c_i^2 \end{aligned} \quad (4.11)$$

Now, let us consider evaluating the energy associated with wave function  $\Phi$ . Taking the approach of multiplying on the left and integrating as outlined above, we have

$$\begin{aligned}
 \int \Phi H \Phi d\mathbf{r} &= \int \left( \sum_i c_i \Psi_i \right) H \left( \sum_j c_j \Psi_j \right) d\mathbf{r} \\
 &= \sum_{ij} c_i c_j \int \Psi_i H \Psi_j d\mathbf{r} \\
 &= \sum_{ij} c_i c_j E_j \delta_{ij} \\
 &= \sum_i c_i^2 E_i
 \end{aligned} \tag{4.12}$$

where we have used Eq. (4.9) to simplify the r.h.s. Thus, the energy associated with the generic wave function  $\Phi$  is determinable from all of the coefficients  $c_i$  (that define how the orthonormal set of  $\Psi_i$  combine to form  $\Phi$ ) and their associated energies  $E_i$ . Regrettably, we still don't know the values for *any* of these quantities. However, let us take note of the following. In the set of all  $E_i$  there must be a lowest energy value (i.e., the set is bounded from below); let us call that energy, corresponding to the 'ground state',  $E_0$ . [Notice that this boundedness is a critical feature of quantum mechanics! In a classical system, one could imagine always finding a state lower in energy than another state by simply 'shrinking the orbits' of the electrons to increase nuclear-electronic attraction while keeping the kinetic energy constant.]

We may now combine the results from Eqs. (4.11) and (4.12) to write

$$\int \Phi H \Phi d\mathbf{r} - E_0 \int \Phi^2 d\mathbf{r} = \sum_i c_i^2 (E_i - E_0) \tag{4.13}$$

Assuming the coefficients to be real numbers, each term  $c_i^2$  must be greater than or equal to zero. By definition of  $E_0$ , the quantity  $(E_i - E_0)$  must also be greater than or equal to zero. Thus, we have

$$\int \Phi H \Phi d\mathbf{r} - E_0 \int \Phi^2 d\mathbf{r} \geq 0 \tag{4.14}$$

which we may rearrange to

$$\frac{\int \Phi H \Phi d\mathbf{r}}{\int \Phi^2 d\mathbf{r}} \geq E_0 \tag{4.15}$$

(note that when  $\Phi$  is normalized, the denominator on the l.h.s. is 1, but it is helpful to have Eq. (4.15) in this more general form for future use).

Equation (4.15) has extremely powerful implications. If we are looking for the best wave function to define the ground state of a system, we can judge the quality of wave functions that we arbitrarily guess by their associated energies: *the lower the better*. This result is critical because it shows us that we do not have to construct our guess wave function  $\Phi$  as a linear combination of (unknown) orthonormal wave functions  $\Psi_i$ , but we may construct it in any manner we wish. The quality of our guess will be determined by how low a value we calculate for the integral in Eq. (4.15). Moreover, since we would like to find the lowest possible energy within the constraints of how we go about constructing a wave function, we can use all of the tools that calculus makes available for locating extreme values.

### 4.2.3 The Born–Oppenheimer Approximation

Up to now, we have been discussing many-particle molecular systems entirely in the abstract. In fact, accurate wave functions for such systems are extremely difficult to express because of the correlated motions of particles. That is, the Hamiltonian in Eq. (4.3) contains pairwise attraction and repulsion terms, implying that no particle is moving independently of all of the others (the term ‘correlation’ is used to describe this interdependency). In order to simplify the problem somewhat, we may invoke the so-called Born–Oppenheimer approximation. This approximation is described with more rigor in Section 15.5, but at this point we present the conceptual aspects without delving deeply into the mathematical details.

Under typical physical conditions, the nuclei of molecular systems are moving much, much more slowly than the electrons (recall that protons and neutrons are about 1800 times more massive than electrons and note the appearance of mass in the denominator of the kinetic energy terms of the Hamiltonian in Eq. (4.3)). For practical purposes, electronic ‘relaxation’ with respect to nuclear motion is instantaneous. As such, it is convenient to decouple these two motions, and compute electronic energies for *fixed* nuclear positions. That is, the nuclear kinetic energy term is taken to be independent of the electrons, correlation in the attractive electron–nuclear potential energy term is eliminated, and the repulsive nuclear–nuclear potential energy term becomes a simply evaluated constant for a given geometry. Thus, the *electronic* Schrödinger equation is taken to be

$$(H_{\text{el}} + V_N)\Psi_{\text{el}}(\mathbf{q}_i; \mathbf{q}_k) = E_{\text{el}}\Psi_{\text{el}}(\mathbf{q}_i; \mathbf{q}_k) \quad (4.16)$$

where the subscript ‘el’ emphasizes the invocation of the Born–Oppenheimer approximation,  $H_{\text{el}}$  includes only the first, third, and fourth terms on the r.h.s. of Eq. (4.3),  $V_N$  is the nuclear–nuclear repulsion energy, and the electronic coordinates  $\mathbf{q}_i$  are independent variables but the nuclear coordinates  $\mathbf{q}_k$  are parameters (and thus appear following a semicolon rather than a comma in the variable list for  $\Psi$ ). The eigenvalue of the electronic Schrödinger equation is called the ‘electronic energy’. Note that the term  $V_N$  is a constant for a given set of fixed nuclear coordinates. Wave functions are invariant to the appearance of constant terms in the Hamiltonian, so in practice one almost always solves Eq. (4.16) without the inclusion of  $V_N$ , in which case the eigenvalue is sometimes called the ‘pure electronic energy’, and one then adds  $V_N$  to this eigenvalue to obtain  $E_{\text{el}}$ .

In general, the Born–Oppenheimer assumption is an extremely mild one, and it is entirely justified in most cases. It is worth emphasizing that this approximation has very profound consequences from a conceptual standpoint – so profound that they are rarely thought about but simply accepted as dogma. Without the Born–Oppenheimer approximation we would lack the concept of a potential energy surface: The PES is the surface defined by  $E_{\text{el}}$  over all possible nuclear coordinates. We would further lack the concepts of equilibrium and transition state geometries, since these are defined as critical points on the PES; instead we would be reduced to discussing high-probability regions of the nuclear wave functions. Of course, for some problems in chemistry, we *do* need to consider the quantum mechanical character of the nuclei, but the advantages afforded by the Born–Oppenheimer approximation should be manifest.

### 4.3 Construction of Trial Wave Functions

Equation (4.16) is simpler than Eq. (4.2) because electron–nuclear correlation has been removed. The remaining correlation, that between the individual electrons, is considerably more troubling. For the moment we will take the simplest possible approach and ignore it; we do this by considering systems with only a single electron. The electronic wave function has thus been reduced to depending only on the fixed nuclear coordinates and the three Cartesian coordinates of the single electron. The eigenfunctions of Eq. (4.16) for a molecular system may now be properly called molecular orbitals (MOs; rather unusual ones in general, since they are for a molecule having only one electron, but MOs nonetheless). To distinguish a one-electron wave function from a many-electron wave function, we will designate the former as  $\psi_{\text{el}}$  and the latter as  $\Psi_{\text{el}}$ . We will hereafter drop the subscript ‘el’ where not required for clarity; unless otherwise specified, all wave functions are electronic wave functions.

The pure electronic energy eigenvalue associated with each molecular orbital is the energy of the electron in that orbital. Experimentally, one might determine this energy by measuring the ionization potential of the electron when it occupies the orbital (fairly easy for the hydrogen atom, considerably more difficult for polynuclear molecules). To measure  $E_{\text{el}}$ , which includes the nuclear repulsion energy, one would need to determine the ‘atomization’ energy, that is, the energy required to ionize the electron *and* to remove all of the nuclei to infinite separation. In practice, atomization energies are not measured, but instead we have compilations of such thermodynamic variables as heats of formation. The relationship between these computed and thermodynamic quantities is discussed in Chapter 10.

#### 4.3.1 The LCAO Basis Set Approach

As noted earlier, we may imagine constructing wave functions in any fashion we deem reasonable, and we may judge the quality of our wave functions (in comparison to one another) by evaluation of the energy eigenvalues associated with each. The one with the lowest energy will be the most accurate and presumably the best one to use for computing other properties by the application of other operators. So, how might one go about choosing



mathematical functions with which to construct a trial wave function? This is a typical question in mathematics – how can an arbitrary function be represented by a combination of more convenient functions? The convenient functions are called a ‘basis set’. Indeed, we have already encountered this formalism – Eq. (2.10) of Chapter 2 illustrates the use of a basis set of cosine functions to approximate torsional energy functions.

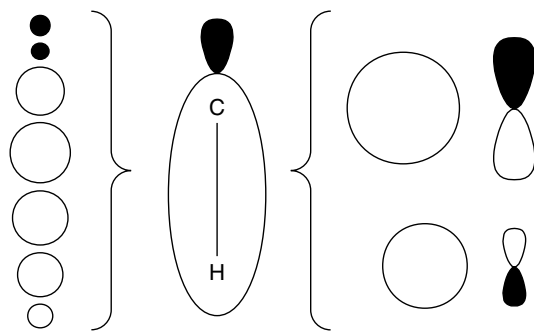
In our QM systems, we have temporarily restricted ourselves to systems of one electron. If, in addition, our system were to have only one nucleus, then we would not need to guess wave functions, but instead we could solve Eq. (4.16) *exactly*. The eigenfunctions that are determined in that instance are the familiar hydrogenic atomic orbitals, 1s, 2s, 2p, 3s, 3p, 3d, etc., whose properties and derivation are discussed in detail in standard texts on quantum mechanics. For the moment, we will not investigate the mathematical representation of these hydrogenic atomic orbitals in any detail, but we will simply posit that, as functions, they may be useful in the construction of more complicated *molecular* orbitals. In particular, just as in Eq. (4.10) we constructed a guess wave function as a linear combination of exact wave functions, so here we will construct a guess wave function  $\phi$  as a linear combination of atomic wave functions  $\varphi$ , i.e.,

$$\phi = \sum_{i=1}^N a_i \varphi_i \quad (4.17)$$

where the set of  $N$  functions  $\varphi_i$  is called the ‘basis set’ and each has associated with it some coefficient  $a_i$ . This construction is known as the linear combination of atomic orbitals (LCAO) approach.

Note that Eq. (4.17) does not specify the locations of the basis functions. Our intuition suggests that they should be centered on the atoms of the molecule, but this is certainly not a requirement. If this comment seems odd, it is worth emphasizing at this point that we should not let our chemical intuition limit our mathematical flexibility. As chemists, we choose to use atomic orbitals (AOs) because we anticipate that they will be efficient functions for the representation of MOs. However, as mathematicians, we should immediately stop thinking about our choices as orbitals, and instead consider them only to be *functions*, so that we avoid being conceptually influenced about how and where to use them.

Recall that the wave function squared has units of probability density. In essence, the electronic wave function is a road map of where the electrons are more or less likely to be found. Thus, we want our basis functions to provide us with the flexibility to allow electrons to ‘go’ where their presence at higher density lowers the energy. For instance, to describe the bonding of a hydrogen atom to a carbon, it is clearly desirable to use a p function on hydrogen, oriented along the axis of the bond, to permit electron density to be localized in the bonding region more efficiently than is possible with only a spherically symmetric s function. Does this imply that the hydrogen atom is somehow sp-hybridized? Not necessarily – the p function is simply serving the purpose of increasing the flexibility with which the *molecular* orbital may be described. If we took away the hydrogen p function and instead placed an s function *in between* the C and H atoms, we could also build up electron density in the bonding region (see Figure 4.1). Thus, the *chemical* interpretation of the coefficients in Eq. (4.17) should only be undertaken with caution, as further described in Chapter 9.



**Figure 4.1** Two different basis sets for representing a C–H  $\sigma$  bonding orbital with the size of the basis functions roughly illustrating their weight in the hybrid MO. The set on the right is the more chemically intuitive since all basis functions are centered on the atoms. Note, however, that the use of a p function to polarize the hydrogen density goes beyond a purely minimalist approach. The set on the left is composed entirely of s functions distributed along the bond. Such a basis set may seem odd in concept, but is quite capable of accurately representing the electron density in space. Indeed, the basis set on the left would have certain computational advantages, chief among them the greater simplicity of working with s functions than with p functions

One should also note that the summation in Eq. (4.17) has an upper limit  $N$ ; we cannot work with an infinite basis in any convenient way (at least not when the basis is AOs). However, the more atomic orbitals we allow into our basis, the closer our basis will come to ‘spanning’ the true molecular orbital space. Thus, the chemical idea that we would limit ourselves to, say, at most one 1s function on each hydrogen atom is needlessly confining from a mathematical standpoint. Indeed, there may be very many ‘true’ one-electron MOs that are very high in energy. Accurately describing these MOs may require some unusual basis functions, e.g., very diffuse functions to describe weakly bound electrons, like those found in Rydberg states. We will discuss these issues in much more detail in Section 6.2, but it is worth emphasizing here, at the beginning, that the distinction between orbitals and functions is a critical one in computational molecular orbital theory.

### 4.3.2 The Secular Equation

All that being said, let us now turn to evaluating the energy of our guess wave function. From Eqs. (4.15) and (4.17) we have

$$E = \frac{\int \left( \sum_i a_i \varphi_i \right) H \left( \sum_j a_j \varphi_j \right) d\mathbf{r}}{\int \left( \sum_i a_i \varphi_i \right) \left( \sum_j a_j \varphi_j \right) d\mathbf{r}}$$

$$\begin{aligned}
& \frac{\sum_{ij} a_i a_j \int \varphi_i H \varphi_j d\mathbf{r}}{\sum_{ij} a_i a_j \int \varphi_i \varphi_j d\mathbf{r}} \\
&= \frac{\sum_{ij} a_i a_j H_{ij}}{\sum_{ij} a_i a_j S_{ij}} \quad (4.18)
\end{aligned}$$

where we have introduced the shorthand notation  $H_{ij}$  and  $S_{ij}$  for the integrals in the numerator and denominator, respectively. These so-called ‘matrix elements’ are no longer as simple as they were in prior discussion, since the atomic orbital basis set, while likely to be efficient, is no longer orthonormal. These matrix elements have more common names,  $H_{ij}$  being called a ‘resonance integral’, and  $S_{ij}$  being called an ‘overlap integral’. The latter has a very clear physical meaning, namely the extent to which any two basis functions overlap in a phase-matched fashion in space. The former integral is not so easily made intuitive, but it is worth pointing out that orbitals which give rise to large overlap integrals will similarly give rise to large resonance integrals. One resonance integral which *is* intuitive is  $H_{ii}$ , which corresponds to the energy of a single electron occupying basis function  $i$ , i.e., it is essentially equivalent to the ionization potential of the AO in the environment of the surrounding molecule.

Now, it is useful to keep in mind our objective. The variational principle instructs us that as we get closer and closer to the ‘true’ one-electron ground-state wave function, we will obtain lower and lower energies from our guess. Thus, once we have selected a basis set, we would like to choose the coefficients  $a_i$  so as to *minimize* the energy for all possible linear combinations of our basis functions. From calculus, we know that a necessary condition for a function (i.e., the energy) to be at its minimum is that its derivatives with respect to all of its free variables (i.e., the coefficients  $a_i$ ) are zero. Notationally, that is

$$\frac{\partial E}{\partial a_k} = 0 \quad \forall k \quad (4.19)$$

(where we make use of the mathematical abbreviation  $\forall$  meaning ‘for all’). Performing this fairly tedious partial differentiation on Eq. (4.18) for each of the  $N$  variables  $a_k$  gives rise to  $N$  equations which must be satisfied in order for Eq. (4.19) to hold true, namely

$$\sum_{i=1}^N a_i (H_{ki} - E S_{ki}) = 0 \quad \forall k \quad (4.20)$$

This set of  $N$  equations (running over  $k$ ) involves  $N$  unknowns (the individual  $a_i$ ). From linear algebra, we know that a set of  $N$  equations in  $N$  unknowns has a non-trivial solution if and only if the determinant formed from the coefficients of the unknowns (in this case the ‘coefficients’ are the various quantities  $H_{ki} - E S_{ki}$ ) is equal to zero. Notationally again,

that is

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0 \quad (4.21)$$

Equation (4.21) is called a secular equation. In general, there will be  $N$  roots  $E$  which permit the secular equation to be true. That is, there will be  $N$  energies  $E_j$  (some of which may be equal to one another, in which case we say the roots are ‘degenerate’) where each value of  $E_j$  will give rise to a different set of coefficients,  $a_{ij}$ , which can be found by solving the set of linear Eqs. (4.20) using  $E_j$ , and these coefficients will define an optimal wave function  $\phi_j$  within the given basis set, i.e.,

$$\phi_j = \sum_{i=1}^N a_{ij} \phi_i \quad (4.22)$$

In a one-electron system, the lowest energy molecular orbital would thus define the ‘ground state’ of the system, and the higher energy orbitals would be ‘excited states’. Obviously, as these are different MOs, they have different basis function coefficients. Although we have not formally proven it, it is worth noting that the variational principle holds for the excited states as well: the calculated energy of a guess wave function for an excited state will be bounded from below by the true excited state energy (MacDonald 1933).

So, in a nutshell, to find the optimal one-electron wave functions for a molecular system, we:

1. Select a set of  $N$  basis functions.
2. For that set of basis functions, determine all  $N^2$  values of both  $H_{ij}$  and  $S_{ij}$ .
3. Form the secular determinant, and determine the  $N$  roots  $E_j$  of the secular equation.
4. For each of the  $N$  values of  $E_j$ , solve the set of linear Eqs. (4.20) in order to determine the basis set coefficients  $a_{ij}$  for that MO.

All of the MOs determined by this process are mutually orthogonal. For degenerate MOs, some minor complications arise, but those are not discussed here.

## 4.4 Hückel Theory

### 4.4.1 Fundamental Principles

To further illuminate the LCAO variational process, we will carry out the steps outlined above for a specific example. To keep things simple (and conceptual), we consider a flavor of molecular orbital theory developed in the 1930s by Erich Hückel to explain some of the unique properties of unsaturated and aromatic hydrocarbons (Hückel 1931; for historical