Section 6 More Quantitative Aspects of Electronic Structure Calculations.

Chapter 17

Electrons interact via pairwise Coulomb forces; within the "orbital picture" these interactions are modelled by less difficult to treat "averaged" potentials. The difference between the true Coulombic interactions and the averaged potential is not small, so to achieve reasonable (ca. 1 kcal/mol) chemical accuracy, high-order corrections to the orbital picture are needed.

The discipline of computational *ab initio* quantum chemistry is aimed at determining the electronic energies and wavefunctions of atoms, molecules, radicals, ions, solids, and all other chemical species. The phrase *ab initio* implies that one attempts to solve the Schrödinger equation from first principles, treating the molecule as a collection of positive nuclei and negative electrons moving under the influence of coulombic potentials, and not using any prior knowledge about this species' chemical behavior.

To make practical use of such a point of view requires that approximations be introduced; the full Schrödinger equation is too difficult to solve exactly for any but simple model problems. These approximations take the form of physical concepts (e.g., orbitals, configurations, quantum numbers, term symbols, energy surfaces, selection rules, etc.) that provide useful means of organizing and interpreting experimental data and computational methods that allow quantitative predictions to be made.

Essentially all *ab initio* quantum chemistry methods use, as a starting point from which improvements are made, a picture in which the electrons interact via a one-electron additive potential. These so-called **mean-field** potentials $V_{mf}(\mathbf{r}) = \sum_j V_{mf}(\mathbf{r_j})$ provide descriptions of atomic and molecular structure that are approximate. Their predictions must be improved to achieve reasonably accurate solutions to the true electronic Schrödinger equation. In so doing, three constructs that characterize essentially all *ab initio* quantum chemical methods are employed: **orbitals, configurations, and electron correlation.**

Since the electronic kinetic energy $T = \sum_j T_j$ operator is one-electron additive, the mean-field Hamiltonian $H^0 = T + V_{mf}$ is also of this form. The additivity of H^0 implies that the mean-field wavefunctions $\{\Phi^0_k\}$ can be formed in terms of products of functions $\{\phi_k\}$ of the coordinates of the individual electrons, and that the corresponding energies $\{E^0_k\}$ are additive. Thus, it is the ansatz that V_{mf} is separable that leads to the concept of

orbitals, which are the one-electron functions $\{\phi_j\}$. These orbitals are found by solving the one-electron Schrödinger equations:

$$(T_1 + V_{mf}(\mathbf{r}_1))\,\phi_j(\mathbf{r}_1) = \epsilon_j\,\phi_j(\mathbf{r}_1);$$

the eigenvalues $\{\epsilon_j\}$ are called **orbital energies.**

Because each of the electrons also possesses intrinsic spin, the one-electron functions $\{\phi_j\}$ used in this construction are taken to be eigenfunctions of $(T_1 + V_{mf}(\mathbf{r}_1))$ multiplied by either α or β. This set of functions is called the set of mean-field **spin-orbitals.**

Given the complete set of solutions to this one-electron equation, a complete set of N-electron mean-field wavefunctions can be written down. Each Φ^0_k is constructed by forming an antisymmetrized product of N spin-orbitals chosen from the set of $\{\phi_j\}$, allowing each spin-orbital in the list to be a function of the coordinates of one of the N electrons (e.g,

$$\Phi^0_k = |\phi_{k1}(\mathbf{r}_1)\,\phi_{k2}(\mathbf{r}_2)\,\phi_{k3}(\mathbf{r}_3)\,\cdots\,\phi_{kN-1}(\mathbf{r}_{N-1})\,\phi_{kN}(\mathbf{r}_N)|,$$

as above). The corresponding mean field energy is evaluated as the sum over those spin-orbitals that appear in Φ^0_k :

$$E^0_k = \sum_{j=1,N} \epsilon_{kj}.$$

By choosing to place N electrons into specific spin-orbitals, one has specified a **configuration**. By making other choices of which N ϕ_j to occupy, one describes other configurations. Just as the one-electron mean-field Schrödinger equation has a complete set of spin-orbital solutions $\{\phi_j$ and $\epsilon_j\}$, the N-electron mean-field Schrödinger equation has a complete set of N-electron <u>configuration state functions</u> (CSFs) Φ^0_k and energies E^0_k.

II. Electron Correlation Requires Moving Beyond a Mean-Field Model

To improve upon the mean-field picture of electronic structure, one must move beyond the single-configuration approximation. It is essential to do so to achieve higher accuracy, but it is also important to do so to achieve a <u>conceptually</u> correct view of chemical electronic structure. However, it is very disconcerting to be told that the familiar $1s^2 2s^2 2p^2$

description of the carbon atom is inadequate and that instead one must think of the ^3P ground state of this atom as a 'mixture' of $1s^22s^22p^2$, $1s^22s^23p^2$, $1s^22s^23d^2$, $2s^23s^22p^2$ (and any other configurations whose angular momenta can be coupled to produce L=1 and S=1).

Although the picture of configurations in which N electrons occupy N spin-orbitals may be very familiar and useful for systematizing electronic states of atoms and molecules, these constructs are approximations to the true states of the system. They were introduced when the mean-field approximation was made, and neither orbitals nor configurations describe the proper eigenstates $\{\psi_k, E_k\}$. The inclusion of instantaneous spatial correlations among electrons is necessary to achieve a more accurate description of atomic and molecular electronic structure. No single spin-orbital product wavefunction is capable of treating electron correlation to any extent; its product nature renders it incapable of doing so.

III. Moving from Qualitative to Quantitative Models

The preceding Chapters introduced, in a qualitative manner, many of the concepts which are used in applying quantum mechanics to electronic structures of atoms and molecules. Atomic, bonding, non-bonding, antibonding, Rydberg, hybrid, and delocalized orbitals and the configurations formed by occupying these orbitals were discussed. Spin and spatial symmetry as well as permutational symmetry were treated, and properly symmetry-adapted configuration state functions were formed. The Slater-Condon rules were shown to provide expressions for Hamiltonian matrix elements (and those involving any one- or two-electron operator) over such CSFs in terms of integrals over the orbitals occupied in the CSFs. Orbital, configuration, and state correlation diagrams were introduced to allow one to follow the evolution of electronic structures throughout a 'reaction path'.

Section 6 addresses the quantitative and computational implementation of many of the above ideas. It is not designed to address all of the state-of-the-art methods which have been, and are still being, developed to calculate orbitals and state wavefunctions. The rapid growth in computer hardware and software power and the evolution of new computer architectures makes it difficult, if not impossible, to present an up-to-date overview of the techniques that are presently at the cutting edge in computational chemistry. Nevertheless, this Section attempts to describe the essential elements of several of the more powerful and commonly used methods; it is likely that many of these elements will persist in the next

generation of computational chemistry techniques although the details of their implementation will evolve considerably. The text by Szabo and Ostlund provides excellent insights into many of the theoretical methods treated in this Section.

IV. Atomic Units

The electronic Hamiltonian is expressed, in this Section, in so-called atomic units (aus)

$$H_e = \sum_j \{ (-1/2) \nabla_j^2 - \sum_a Z_a/r_{j,a} \} + \sum_{j<k} 1/r_{j,k} .$$

These units are introduced to remove all $\hbar$, e, and m_e factors from the equations.

To effect this unit transformation, one notes that the kinetic energy operator scales as r_j^{-2} whereas the coulombic potentials scale as r_j^{-1} and as $r_{j,k}^{-1}$. So, if each of the distances appearing in the cartesian coordinates of the electrons and nuclei were expressed as a unit of length a_0 multiplied by a dimensionless length factor, the kinetic energy operator would involve terms of the form $(-\hbar^2/2(a_0)^2 m_e) \nabla_j^2$, and the coulombic potentials would appear as $Z_a e^2/(a_0)r_{j,a}$ and $e^2/(a_0)r_{j,k}$. A factor of e^2/a_0 (which has units of energy since a_0 has units of length) can then be removed from the coulombic and kinetic energies, after which the kinetic energy terms appear as $(-\hbar^2/2(e^2 a_0)m_e) \nabla_j^2$ and the potential energies appear as $Z_a/r_{j,a}$ and $1/r_{j,k}$. Then, choosing $a_0 = \hbar^2/e^2 m_e$ changes the kinetic energy terms into $-1/2 \nabla_j^2$; as a result, the entire electronic Hamiltonian takes the form given above in which no e^2, m_e, or $\hbar^2$ factors appear. The value of the so-called Bohr radius $a_0 = \hbar^2/e^2 m_e$ is 0.529 Å, and the so-called Hartree energy unit e^2/a_0, which factors out of H_e, is 27.21 eV or 627.51 kcal/mol.

Chapter 18
The single Slater determinant wavefunction (properly spin and symmetry adapted) is the starting point of the most common mean field potential. It is also the origin of the molecular orbital concept.

I. Optimization of the Energy for a Multiconfiguration Wavefunction

A. The Energy Expression

The most straightforward way to introduce the concept of optimal molecular orbitals is to consider a trial wavefunction of the form which was introduced earlier in Chapter 9.II. The expectation value of the Hamiltonian for a wavefunction of the multiconfigurational form

$$\Psi = \sum_I C_I \Phi_I,$$

where Φ_I is a space- and spin-adapted CSF which consists of determinental wavefunctions $|\phi_{I1} \phi_{I2} \phi_{I3} \cdots \phi_{IN}|$, can be written as:

$$E = \sum_{I,J = 1, M} C_I C_J < \Phi_I | H | \Phi_J > .$$

The spin- and space-symmetry of the Φ_I determine the symmetry of the state Ψ whose energy is to be optimized.

In this form, it is clear that E is a quadratic function of the CI amplitudes C_J ; it is a quartic functional of the spin-orbitals because the Slater-Condon rules express each $< \Phi_I | H | \Phi_J >$ CI matrix element in terms of one- and two-electron integrals $< \phi_i | f | \phi_j >$ and $< \phi_i \phi_j | g | \phi_k \phi_l >$ over these spin-orbitals.

B. Application of the Variational Method

The <u>variational</u> method can be used to optimize the above expectation value expression for the electronic energy (i.e., to make the functional stationary) as a function of the CI coefficients C_J and the LCAO-MO coefficients $\{C_{\nu,i}\}$ that characterize the spin-orbitals. However, in doing so the set of $\{C_{\nu,i}\}$ can not be treated as entirely independent variables. The fact that the spin-orbitals $\{\phi_i\}$ are assumed to be orthonormal imposes a set of constraints on the $\{C_{\nu,i}\}$:

$$< \phi_i | \phi_j > = \delta_{i,j} = \sum_{\mu,\nu} C^*_{\mu,i} < \chi_\mu | \chi_\nu > C_{\nu,j}.$$

These constraints can be enforced within the variational optimization of the energy function mentioned above by introducing a set of Lagrange multipliers $\{\varepsilon_{i,j}\}$, one for each constraint condition, and subsequently differentiating

$$E - \sum_{i,j} \varepsilon_{i,j} [\delta_{i,j} - \sum_{\mu,\nu} C^*_{\mu,i} < \chi_\mu | \chi_\nu > C_{\nu,j}]$$

with respect to each of the $C_{J,i}$ variables.

C. The Fock and Secular Equations

Upon doing so, the following set of equations is obtained (early references to the derivation of such equations include A. C. Wahl, J. Chem. Phys. <u>41</u>, 2600 (1964) and F. Grein and T. C. Chang, Chem. Phys. Lett. <u>12</u>, 44 (1971); a more recent overview is presented in R. Shepard, p 63, in Adv. in Chem. Phys. LXIX, K. P. Lawley, Ed., Wiley-Interscience, New York (1987); the subject is also treated in the textbook <u>Second Quantization Based Methods in Quantum Chemistry</u>, P. Jørgensen and J. Simons, Academic Press, New York (1981))) :

$$\sum_{J=1, M} H_{I,J} \ C_J \ = E \ C_I, \quad I = 1, 2, ... \ M, \text{ and}$$

$$F \ \phi_i = \sum_j \epsilon_{i,j} \ \phi_j,$$

where the $\epsilon_{i,j}$ are Lagrange multipliers.

The first set of equations govern the $\{C_J\}$ amplitudes and are called the CI- secular equations. The second set determine the LCAO-MO coefficients of the spin-orbitals $\{\phi_j\}$ and are called the Fock equations. The Fock operator F is given in terms of the one- and two-electron operators in H itself as well as the so-called one- and two-electron density matrices $\gamma_{i,j}$ and $\Gamma_{i,j,k,l}$ which are defined below. These density matrices reflect the averaged occupancies of the various spin orbitals in the CSFs of Ψ. The resultant expression for F is:

$$F \ \phi_i = \sum_j \gamma_{i,j} \ h \ \phi_j + \sum_{j,k,l} \Gamma_{i,j,k,l} \ J_{j,l} \ \phi_k,$$

where h is the one-electron component of the Hamiltonian (i.e., the kinetic energy operator and the sum of coulombic attractions to the nuclei). The operator $J_{j,l}$ is defined by:

$$J_{j,l} \ \phi_k(r) = \int \phi^*_j(r') \ \phi_l(r') 1/|r-r'| \ d \ '\ \phi_k(r),$$

where the integration denoted d ' is over the spatial and spin coordinates. The so-called spin integration simply means that the α or β spin function associated with ϕ_l must be the

same as the α or β spin function associated with ϕ_j or the integral will vanish. This is a consequence of the orthonormality conditions $<\alpha|\alpha> = <\beta|\beta> = 1, <\alpha|\beta> = <\beta|\alpha> = 0$.

D. One- and Two- Electron Density Matrices

The density matrices introduced above can most straightforwardly be expressed in terms of the CI amplitudes and the nature of the orbital occupancies in the CSFs of Ψ as follows:

1. $\gamma_{i,i}$ is the sum over all CSFs, in which ϕ_i is occupied, of the square of the C_I coefficient of that CSF:

$$\gamma_{i,i} = \sum_I (\text{with } \phi_i \text{ occupied}) \, C^2_I .$$

2. $\gamma_{i,j}$ is the sum over pairs of CSFs which differ by a single spin-orbital occupancy (i.e., one having ϕ_i occupied where the other has ϕ_j occupied after the two are placed into maximal coincidence-the sign factor (sign) arising from bringing the two to maximal coincidence is attached to the final density matrix element):

$$\gamma_{i,j} = \sum_{I,J} (\text{sign})(\text{with } \phi_i \text{ occupied in I where } \phi_j \text{ is in J}) \, C_I C_J .$$

The two-electron density matrix elements are given in similar fashion:

3. $\quad \Gamma_{i,j,i,j} = \sum_I (\text{with both } \phi_i \text{ and } \phi_j \text{ occupied}) \, C_I C_I ;$

4. $\quad \Gamma_{i,j,j,i} = -\sum_I (\text{with both } \phi_i \text{ and } \phi_j \text{ occupied}) \, C_I C_I = -\Gamma_{i,j,i,j}$

(it can be shown, in general that $\Gamma_{i,j,k,l}$ is odd under exchange of i and j, odd under exchange of k and l and even under (i,j)<=>(k,l) exchange; this implies that $\Gamma_{i,j,k,l}$ vanishes if i = j or k = l.) ;

5. $\quad \Gamma_{i,j,k,j} = \sum_{I,J} (\text{sign})(\text{with } \phi_j \text{ in both I and J}$
$\qquad\qquad$ and ϕ_i in I where ϕ_k is in J) $C_I C_J$

$$= \Gamma_{j,i,j,k} = -\Gamma_{i,j,j,k} = -\Gamma_{j,i,k,j};$$

6. $\gamma_{i,j,k,l} = \sum_{I,J} (\text{sign})(\text{with } \phi_i \text{ in } I \text{ where } \phi_k \text{ is in } J \text{ and } \phi_j \text{ in } I \text{ where } \phi_l \text{ is in } J)\, C_I C_J$

$$= \gamma_{j,i,l,k} = -\gamma_{j,i,k,l} = -\gamma_{i,j,l,k} = \gamma_{j,i,l,k} \cdot$$

These density matrices are themselves quadratic functions of the CI coefficients and they reflect all of the permutational symmetry of the determinental functions used in constructing Ψ; they are a compact representation of all of the Slater-Condon rules as applied to the particular CSFs which appear in Ψ. They contain all information about the spin-orbital occupancy of the CSFs in Ψ. The one- and two- electron integrals $< \phi_i | f | \phi_j >$ and $< \phi_i \phi_j | g | \phi_k \phi_l >$ contain all of the information about the magnitudes of the kinetic and Coulombic interaction energies.

II. The Single-Determinant Wavefunction

The simplest trial function of the form given above is the single Slater determinant function:

$$\Psi = |\phi_1 \phi_2 \phi_3 \cdots \phi_N|.$$

For such a function, the CI part of the energy minimization is absent (the classic papers in which the SCF equations for closed- and open-shell systems are treated are C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951); 32, 179 (1960)) and the density matrices simplify greatly because only one spin-orbital occupancy is operative. In this case, the orbital optimization conditions reduce to:

$$F \phi_i = \sum_j \epsilon_{i,j} \phi_j \, ,$$

where the so-called Fock operator F is given by

$$F \phi_i = h \phi_i + \sum_{j(\text{occupied})} [J_j - K_j] \phi_i \, .$$

The coulomb (J_j) and exchange (K_j) operators are defined by the relations:

$$J_j \phi_i = \int \phi^*_j(r') \phi_j(r') 1/|r-r'| d' \phi_i(r) \, , \text{ and}$$

$$K_j \phi_i = \int \phi_j^*(r') \phi_i(r') 1/|r-r'| \, dr' \, \phi_j(r).$$

Again, the integration implies integration over the spin variables associated with the ϕ_j (and, for the exchange operator, ϕ_i), as a result of which the exchange integral vanishes unless the spin function of ϕ_j is the same as that of ϕ_i; the coulomb integral is non-vanishing no matter what the spin functions of ϕ_j and ϕ_i.

The sum over coulomb and exchange interactions in the Fock operator runs only over those spin-orbitals that are occupied in the trial Ψ. Because a unitary transformation among the orbitals that appear in Ψ leaves the determinant unchanged (this is a property of determinants- det (UA) = det (U) det (A) = 1 det (A), if U is a unitary matrix), it is possible to choose such a unitary transformation to make the $\epsilon_{i,j}$ matrix diagonal. Upon so doing, one is left with the so-called <u>canonical Hartree-Fock equations</u>:

$$F \phi_i = \epsilon_i \phi_j,$$

where ϵ_i is the diagonal value of the $\epsilon_{i,j}$ matrix after the unitary transformation has been applied; that is, ϵ_i is an eigenvalue of the $\epsilon_{i,j}$ matrix. These equations are of the eigenvalue-eigenfunction form with the Fock operator playing the role of an effective one-electron Hamiltonian and the ϕ_i playing the role of the one-electron eigenfunctions.

It should be noted that the Hartree-Fock equations $F \phi_i = \epsilon_i \phi_j$ possess solutions for the spin-orbitals which appear in Ψ (the so-called <u>occupied</u> spin-orbitals) as well as for orbitals which are not occupied in Ψ (the so-called <u>virtual</u> spin-orbitals). In fact, the F operator is hermitian, so it possesses a complete set of orthonormal eigenfunctions; only those which appear in Ψ appear in the coulomb and exchange potentials of the Fock operator. The physical meaning of the occupied and virtual orbitals will be clarified later in this Chapter (Section VII.A)

III. The Unrestricted Hartree-Fock Spin Impurity Problem

As formulated above in terms of spin-orbitals, the Hartree-Fock (HF) equations yield orbitals that do not guarantee that Ψ possesses proper spin symmetry. To illustrate the point, consider the form of the equations for an open-shell system such as the Lithium atom Li. If $1s\alpha$, $1s\beta$, and $2s\alpha$ spin-orbitals are chosen to appear in the trial function Ψ, then the Fock operator will contain the following terms:

$$F = h + J_{1s} + J_{1s} + J_{2s} - [K_{1s} + K_{1s} + K_{2s}].$$

Acting on an α spin-orbital ϕ_k with F and carrying out the spin integrations, one obtains

$$F \phi_k = h \phi_k + (2J_{1s} + J_{2s}) \phi_k - (K_{1s} + K_{2s}) \phi_k.$$

In contrast, when acting on a β spin-orbital, one obtains

$$F \phi_k = h \phi_k + (2J_{1s} + J_{2s}) \phi_k - (K_{1s}) \phi_k.$$

Spin-orbitals of α and β type do not experience the same exchange potential in this model, which is clearly due to the fact that ϕ contains two α spin-orbitals and only one β spin-orbital.

One consequence of the spin-polarized nature of the effective potential in F is that the optimal $1s\alpha$ and $1s\beta$ spin-orbitals, which are themselves solutions of $F \phi_i = \varepsilon_i \phi_i$, do not have identical orbital energies (i.e., $\varepsilon_{1s\alpha} \neq \varepsilon_{1s\beta}$) and are not spatially identical to one another (i.e., $\phi_{1s\alpha}$ and $\phi_{1s\beta}$ do not have identical LCAO-MO expansion coefficients). This resultant spin polarization of the orbitals in ϕ gives rise to spin impurities in ϕ. That is, the determinant $| 1s\alpha \; 1s'\beta \; 2s\alpha |$ is not a pure doublet spin eigenfunction although it is an S_z eigenfunction with $M_s = 1/2$; it contains both $S = 1/2$ and $S = 3/2$ components. If the $1s\alpha$ and $1s'\beta$ spin-orbitals were spatially identical, then $| 1s\alpha \; 1s'\beta \; 2s\alpha |$ would be a pure spin eigenfunction with $S = 1/2$.

The above single-determinant wavefunction is commonly referred to as being of the **unrestricted Hartree-Fock** (UHF) type because no restrictions are placed on the spatial nature of the orbitals which appear in ϕ. In general, UHF wavefunctions are not of pure spin symmetry for any open-shell system. Such a UHF treatment forms the starting point of early versions of the widely used and highly successful Gaussian 70 through Gaussian-8X series of electronic structure computer codes which derive from J. A. Pople and co-workers (see, for example, M. J. Frisch, J. S. Binkley, H. B. Schlegel, K Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohling, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whitehead, D. J. Fox, E. M. Fleuder, and J. A. Pople, Gaussian 86 , Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA (1984)).

The inherent spin-impurity problem is sometimes 'fixed' by using the orbitals which are obtained in the UHF calculation to subsequently form a properly spin-adapted wavefunction. For the above Li atom example, this amounts to forming a new

wavefunction (after the orbitals are obtained via the UHF process) using the techniques detailed in **Section 3** and **Appendix G**:

$$\Psi = 1/\sqrt{2}\ [\ |1s\ \ 1s'\ \ 2s\ | - |\ 1s\ \ 1s'\ \ 2s\ |\ ].$$

This wavefunction is a pure S = 1/2 state. This prescription for avoiding spin contamination (i.e., carrying out the UHF calculation and then forming a new spin-pure Ψ) is referred to as <u>spin-projection</u>.

It is, of course, possible to first form the above spin-pure Ψ as a trial wavefunction and to then determine the orbitals 1s 1s' and 2s which minimize its energy; in so doing, one is dealing with a spin-pure function from the start. The problem with carrying out this process, which is referred to as a <u>spin-adapted</u> Hartree-Fock calculation, is that the resultant 1s and 1s' orbitals still do not have identical spatial attributes. Having a set of orbitals (1s, 1s', 2s, and the virtual orbitals) that form a non-orthogonal set (1s and 1s' are neither identical nor orthogonal) makes it difficult to progress beyond the single-configuration wavefunction as one often wishes to do. That is, it is difficult to use a spin-adapted wavefunction as a starting point for a correlated-level treatment of electronic motions.

Before addressing head-on the problem of how to best treat orbital optimization for open-shell species, it is useful to examine how the HF equations are solved in practice in terms of the LCAO-MO process.

IV. The LCAO-MO Expansion

The HF equations $F\phi_i = \varepsilon_i \phi_i$ comprise a set of integro-differential equations; their differential nature arises from the kinetic energy operator in h, and the coulomb and exchange operators provide their integral nature. The solutions of these equations must be achieved iteratively because the J_i and K_i operators in F depend on the orbitals ϕ_i which are to be solved for. Typical iterative schemes begin with a 'guess' for those ϕ_i which appear in Ψ, which then allows F to be formed. Solutions to $F\phi_i = \varepsilon_i \phi_i$ are then found, and those ϕ_i which possess the space and spin symmetry of the occupied orbitals of Ψ and which have the proper energies and nodal character are used to generate a new F operator (i.e., new J_i and K_i operators). The new F operator then gives new ϕ_i and ε_i via solution of the new $F\phi_i = \varepsilon_i \phi_i$ equations. This iterative process is continued until the ϕ_i and ε_i do not vary significantly from one iteration to the next, at which time one says that the process has converged. This iterative procedure is referred to as the Hartree-Fock <u>self-consistent field</u>

(SCF) procedure because iteration eventually leads to coulomb and exchange potential fields that are consistent from iteration to iteration.

In practice, solution of $F \phi_i = \epsilon_i \phi_i$ as an integro-differential equation can be carried out only for atoms (C. Froese-Fischer, Comp. Phys. Commun. <u>1</u>, 152 (1970)) and linear molecules (P. A. Christiansen and E. A. McCullough, J. Chem. Phys. <u>67</u>, 1877 (1977)) for which the angular parts of the ϕ_i can be exactly separated from the radial because of the axial- or full- rotation group symmetry (e.g., $\phi_i = Y_{l,m} R_{n,l}(r)$ for an atom and $\phi_i = \exp(im\phi) R_{n,l,m}(r, \theta)$ for a linear molecule). In such special cases, $F \phi_i = \epsilon_i \phi_i$ gives rise to a set of coupled equations for the $R_{n,l}(r)$ or $R_{n,l,m}(r, \theta)$ which can and have been solved. However, for non-linear molecules, the HF equations have not yet been solved in such a manner because of the three-dimensional nature of the ϕ_i and of the potential terms in F.

In the most commonly employed procedures used to solve the HF equations for non-linear molecules, the ϕ_i are expanded in a basis of functions χ_μ according to the LCAO-MO procedure:

$$\phi_i = \sum_\mu C_{\mu,i} \chi_\mu .$$

Doing so then reduces $F \phi_i = \epsilon_i \phi_i$ to a matrix eigenvalue-type equation of the form:

$$\sum_\nu F_{\mu,\nu} C_{\nu,i} = \epsilon_i \sum_\nu S_{\mu,\nu} C_{\nu,i} ,$$

where $S_{\mu,\nu} = \langle \chi_\mu | \chi_\nu \rangle$ is the overlap matrix among the atomic orbitals (aos) and

$$F_{\mu,\nu} = \langle \chi_\mu | h | \chi_\nu \rangle + \sum_{\kappa,\lambda} [\gamma_{\kappa,\lambda} \langle \chi_\mu \chi_\kappa | g | \chi_\nu \chi_\lambda \rangle - \gamma_{\kappa,\lambda}^{ex} \langle \chi_\mu \chi_\kappa | g | \chi_\lambda \chi_\nu \rangle]$$

is the matrix representation of the Fock operator in the ao basis. The coulomb and exchange- density matrix elements in the ao basis are:

$$\gamma_{\kappa,\lambda} = \sum_{i(occupied)} C_{\kappa,i} C_{\lambda,i}, \text{ and}$$

$$\gamma_{\kappa,\lambda}^{ex} = \sum_{i(occ., \text{ and same spin})} C_{\kappa,i} C_{\lambda,i},$$

where the sum in $\gamma_{\kappa,\lambda}^{ex}$ runs over those occupied spin-orbitals whose m_s value is equal to that for which the Fock matrix is being formed (for a closed-shell species, $\gamma_{\kappa,\lambda}^{ex} = 1/2 \gamma_{\kappa,\lambda}$).

It should be noted that by moving to a matrix problem, one does not remove the need for an iterative solution; the $F_{\mu,\nu}$ matrix elements depend on the $C_{\nu,i}$ LCAO-MO coefficients which are, in turn, solutions of the so-called Roothaan matrix Hartree-Fock equations- $\sum_\nu F_{\mu,\nu} C_{\nu,i} = \varepsilon_i \sum_\nu S_{\mu,\nu} C_{\nu,i}$. One should also note that, just as $F \phi_i = \varepsilon_i \phi_j$ possesses a complete set of eigenfunctions, the matrix $F_{\mu,\nu}$, whose dimension M is equal to the number of atomic basis orbitals used in the LCAO-MO expansion, has M eigenvalues ε_i and M eigenvectors whose elements are the $C_{\nu,i}$. Thus, there are occupied and virtual molecular orbitals (mos) each of which is described in the LCAO-MO form with $C_{\nu,i}$ coefficients obtained via solution of

$$\sum_\nu F_{\mu,\nu} C_{\nu,i} = \varepsilon_i \sum_\nu S_{\mu,\nu} C_{\nu,i} .$$

V. Atomic Orbital Basis Sets

A. STOs and GTOs

The basis orbitals commonly used in the LCAO-MO-SCF process fall into two classes:

1. Slater-type orbitals

$$\phi_{n,l,m} (r, \theta, \phi) = N_{n,l,m,\zeta} Y_{l,m} (\theta, \phi) r^{n-1} e^{-\zeta r},$$

which are characterized by quantum numbers n, l, and m and exponents (which characterize the 'size' of the basis function) ζ. The symbol $N_{n,l,m,\zeta}$ denotes the normalization constant.

2. Cartesian Gaussian-type orbitals

$$\phi_{a,b,c} (r, \theta, \phi) = N'_{a,b,c,\alpha} x^a y^b z^c \exp(-\alpha r^2),$$

characterized by quantum numbers a, b, and c which detail the angular shape and direction of the orbital and exponents α which govern the radial 'size' of the basis function. For example, orbitals with a, b, and c values of 1,0,0 or 0,1,0 or 0,0,1 are p_x , p_y , and p_z orbitals; those with a,b,c values of 2,0,0 or 0,2,0 or 0,0,2 and

1,1,0 or 0,1,1 or 1,0,1 span the space of five d orbitals and one s orbital (the sum of the 2,0,0 and 0,2,0 and 0,0,2 orbitals is an s orbital because $x^2 + y^2 + z^2 = r^2$ is independent of θ and ϕ).

For both types of orbitals, the coordinates r, θ, and ϕ refer to the position of the electron relative to a set of axes attached to the center on which the basis orbital is located. Although Slater-type orbitals (STOs) are preferred on fundamental grounds (e.g., as demonstrated in Appendices A and B, the hydrogen atom orbitals are of this form and the exact solution of the many-electron Schrödinger equation can be shown to be of this form (in each of its coordinates) near the nuclear centers), STOs are used primarily for atomic and linear-molecule calculations because the multi-center integrals $< \chi_a \chi_b | g | \chi_c \chi_d >$ (each basis orbital can be on a separate atomic center) which arise in polyatomic-molecule calculations can not efficiently be performed when STOs are employed. In contrast, such integrals can routinely be done when Gaussian-type orbitals (GTOs) are used. This fundamental advantage of GTOs has lead to the dominance of these functions in molecular quantum chemistry.

To understand why integrals over GTOs can be carried out when analogous STO-based integrals are much more difficult, one must only consider the orbital products ($\chi_a \chi_c$ (r_1) and $\chi_b \chi_d$ (r_2)) which arise in such integrals. For orbitals of the GTO form, such products involve $\exp(-\alpha_a (\mathbf{r}-\mathbf{R}_a)^2) \exp(-\alpha_c (\mathbf{r}-\mathbf{R}_c)^2)$. By completing the square in the exponent, this product can be rewritten as follows:

$$\exp(-\alpha_a (\mathbf{r}-\mathbf{R}_a)^2) \exp(-\alpha_c (\mathbf{r}-\mathbf{R}_c)^2)$$

$$= \exp(-(\alpha_a + \alpha_c)(\mathbf{r}-\mathbf{R}')^2) \exp(-\alpha'(\mathbf{R}_a-\mathbf{R}_c)^2),$$

where

$$\mathbf{R}' = [\alpha_a \mathbf{R}_a + \alpha_c \mathbf{R}_c]/(\alpha_a + \alpha_c) \text{ and}$$

$$\alpha' = \alpha_a \alpha_c/(\alpha_a + \alpha_c).$$

Thus, the product of two GTOs on different centers is equal to a single other GTO at a center $\mathbf{R}'$ between the two original centers. As a result, even a four-center two-electron integral over GTOs can be written as, at most, a two-center two-electron integral; it turns out that this reduction in centers is enough to allow all such integrals to be carried out. A

similar reduction does not arise for STOs because the product of two STOs can not be rewritten as a new STO at a new center.

To overcome the primary weakness of GTO functions, that they have incorrect behavior near the nuclear centers (i.e., their radial derivatives vanish at the nucleus whereas the derivatives of STOs are non-zero), it is common to combine two, three, or more GTOs, with combination coefficients which are fixed and <u>not</u> treated as LCAO-MO parameters, into new functions called contracted GTOs or CGTOs. Typically, a series of tight, medium, and loose GTOs (i.e., GTOs with large, medium, and small   values, respectively) are multiplied by so-called contraction coefficients and summed to produce a CGTO which appears to possess the proper 'cusp' (i.e., non-zero slope) at the nuclear center (although even such a combination can not because each GTO has zero slope at the nucleus).

B. Basis Set Libraries

Much effort has been devoted to developing sets of STO or GTO basis orbitals for main-group elements and the lighter transition metals. This ongoing effort is aimed at providing standard basis set libraries which:
1. Yield reasonable chemical accuracy in the resultant wavefunctions and energies.
2. Are cost effective in that their use in practical calculations is feasible.
3. Are relatively transferrable in the sense that the basis for a given atom is flexible enough to be used for that atom in a variety of bonding environments (where the atom's hybridization and local polarity may vary).

C. The Fundamental Core and Valence Basis

In constructing an atomic orbital basis to use in a particular calculation, one must choose from among several classes of functions. First, the size and nature of the primary core and valence basis must be specified. Within this category, the following choices are common:
1. A <u>minimal basis</u> in which the number of STO or CGTO orbitals is equal to the number of core and valence atomic orbitals in the atom.
2. A <u>double-zeta</u> (DZ) basis in which twice as many STOs or CGTOs are used as there are core and valence atomic orbitals. The use of more basis functions is motivated by a desire to provide additional variational flexibility to the LCAO-MO process. This flexibility allows the LCAO-MO process to generate molecular orbitals of variable diffuseness as the

local electronegativity of the atom varies. Typically, double-zeta bases include pairs of functions with one member of each pair having a smaller exponent (ζ or α value) than in the minimal basis and the other member having a larger exponent.

3. A triple-zeta (TZ) basis in which three times as many STOs or CGTOs are used as the number of core and valence atomic orbitals.

4. Dunning has developed CGTO bases which range from approximately DZ to substantially beyond TZ quality (T. H. Dunning, J. Chem. Phys. 53, 2823 (1970); T. H. Dunning and P. J. Hay in Methods of Electronic Structure Theory, H. F. Schaefer, III Ed., Plenum Press, New York (1977))). These bases involve contractions of primitive GTO bases which Huzinaga had earlier optimized (S. Huzinaga, J. Chem. Phys. 42, 1293 (1965)) for use as uncontracted functions (i.e., for which Huzinaga varied the α values to minimize the energies of several electronic states of the corresponding atom). These Dunning bases are commonly denoted, for example, as follows for first-row atoms: (10s,6p/5s,4p), which means that 10 s-type primitive GTOs have been contracted to produce 5 separate s-type CGTOs and that 6 primitive p-type GTOs were contracted to generate 4 separate p-type CGTOs. More recent basis sets from the Dunning group are given in T. Dunning, J. Chem. Phys. 90, 1007 (1990).

5. Even-tempered basis sets (M. W. Schmidt and K. Ruedenberg, J. Chem. Phys. 71, 3961 (1979)) consist of GTOs in which the orbital exponents α_k belonging to series of orbitals consist of geometrical progressions: $\alpha_k = a \, \beta^k$, where a and β characterize the particular set of GTOs.

6. STO-3G bases were employed some years ago (W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys. 51, 2657 (1969)) but are less popular recently. These bases are constructed by least squares fitting GTOs to STOs which have been optimized for various electronic states of the atom. When three GTOs are employed to fit each STO, a STO-3G basis is formed.

7. 4-31G, 5-31G, and 6-31G bases (R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. 54, 724 (1971); W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. 56, 2257 (1972); P. C. Hariharan and J. A. Pople, Theoret. Chim. Acta. (Berl.) 28, 213 (1973); R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. 72, 650 (1980)) employ a single CGTO of contraction length 4, 5, or 6 to describe the core orbital. The valence space is described at the DZ level with the first CGTO constructed from 3 primitive GTOs and the second CGTO built from a single primitive GTO.

The values of the orbital exponents (ζs or αs) and the GTO-to-CGTO contraction coefficients needed to implement a particular basis of the kind described above have been tabulated in several journal articles and in computer data bases (in particular, in the data

base contained in the book <u>Handbook of Gaussian Basis Sets: A. Compendium for Ab initio Molecular Orbital Calculations</u>, R. Poirer, R. Kari, and I. G. Csizmadia, Elsevier Science Publishing Co., Inc., New York, New York (1985)).

Several other sources of basis sets for particular atoms are listed in the Table shown below (here JCP and JACS are abbreviations for the Journal of Chemical Physics and the Journal of The American Chemical Society, respectively).

<u>Literature Reference</u>	<u>Basis Type</u>	<u>Atoms</u>
Hehre, W.J.; Stewart, R.F.; Pople, J.A. JCP <u>51</u>, 2657 (1969). Hehre, W.J.; Ditchfield, R.; Stewart, R.F.; Pople, J.A. JCP <u>52</u>, 2769 (1970).	STO-3G	H-Ar
Binkley, J.S.; Pople, J.A.; Hehre, W.J. JACS <u>102</u>, 939 (1980).	3-21G	H-Ne
Gordon, M.S.; Binkley, J.S.; Pople, J.A.; Pietro, W.J.; Hehre, W.J. JACS <u>104</u>, 2797 (1982).	3-21G	Na-Ar
Dobbs, K.D.; Hehre, W.J. J. Comput. Chem. <u>7</u>, 359 (1986).	3-21G	K,Ca,Ga
Dobbs, K.D.; Hehre, W.J. J. Comput. Chem. <u>8,</u> 880 (1987).	3-21G	Sc-Zn
Ditchfield, R.; Hehre, W.J.; Pople, J.A. JCP <u>54</u>, 724 (1971).	6-31G	H
Dill, J.D.; Pople, J.A. JCP <u>62</u>, 2921 (1975).	6-31G	Li,B
Binkley, J.S.; Pople, J.A. JCP <u>66</u>, 879 (1977).	6-31G	Be

Hehre, W.J.; Ditchfield, R.; Pople, J.A. JCP <u>56</u>, 2257 (1972).	6-31G	C-F
Francl, M.M.; Pietro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFrees, D.J.; Pople, J.A. JCP <u>77</u>, 3654 (1982).	6-31G	Na-Ar
Dunning, T. JCP <u>53</u>, 2823 (1970).	(4s/2s)	H
	(4s/3s)	H
	(9s5p/3s2p)	B-F
	(9s5p/4s2p)	B-F
	(9s5p/5s3p)	B-F
Dunning, T. JCP <u>55</u>, 716 (1971).	(5s/3s)	H
	(10s/4s)	Li
	(10s/5s)	Be
	(10s6p/5s3p)	B-Ne
	(10s6p/5s4p)	B-Ne
Krishnan, R.; Binkley, J.S.; Seeger, R.; Pople, J.A. JCP <u>72</u>, 650 (1980).	6-311G	H-Ne
Dunning, unpublished VDZ.	(4s/2s)	H
	(9s5p/3s2)	Li,Be,C-Ne
	(12s8p/4s3p)	Na-Ar
Dunning, unpublished VTZ.	(5s/3s)	H
	(6s/3s)	H
	(12s6p/4s3p)	Li,Be,C-Ne
	(17s10p/5s4p)	Mg-Ar
Dunning, unpublished VQZ.	(7s/4s)	H
	(8s/4s)	H
	(16s7p/5s4p)	B-Ne

Dunning, T. JCP <u>90</u>, 1007 (1989).   (4s1p/2s1p)   H
(pVDZ,pVTZ,pVQZ correlation-consistent)   (5s2p1d/3s2p1d)   H
                                  (6s3p1d1f/4s3p2d1f)   H
                                  (9s4p1d/3s2p1d)   B-Ne
                                  (10s5p2d1f/4s3p2d1f) B-Ne
                                  (12s6p3d2f1g/5s4p3d2f1g) B-Ne

Huzinaga, S.; Klobukowski, M.; Tatewaki, H.   (14s/2s)   Li,Be
Can. J. Chem. <u>63</u>, 1812 (1985).   (14s9p/2s1p)   B-Ne
                                    (16s9p/3s1p)   Na-Mg
                                  (16s11p/3s2p) Al-Ar

Huzinaga, S.; Klobukowski, M.   (14s10p/2s1p)   B-Ne
THEOCHEM. <u>44</u>, 1 (1988).   (17s10p/3s1p)   Na-Mg
                                  (17s13p/3s2p)   Al-Ar
                                  (20s13p/4s2p)   K-Ca
                                (20s13p10d/4s2p1d) Sc-Zn
                                  (20s14p9d/4s3d1d)   Ga

McLean, A.D.; Chandler, G.S.   (12s8p/4s2p) Na-Ar, P$^-$,S$^-$,Cl$^-$
JCP <u>72</u>, 5639 (1980).   (12s8p/5s3p) Na-Ar, P$^-$,S$^-$,Cl$^-$
                                  (12s8p/6s4p) Na-Ar, P$^-$,S$^-$,Cl$^-$
                                  (12s9p/6s4p) Na-Ar, P$^-$,S$^-$,Cl$^-$
                                  (12s9p/6s5p) Na-Ar, P$^-$,S$^-$,Cl$^-$
Dunning, T.H.Jr.; Hay, P.J. Chapter 1 in   (11s7p/6s4p)   Al-Cl
'Methods of Electronic Structure Theory',
Schaefer, H.F.III, Ed., Plenum Press,
N.Y., 1977.

Hood, D.M.; Pitzer, R.M.; Schaefer, H.F.III (14s11p6d/10s8p3d) Sc-Zn
JCP <u>71</u>, 705 (1979).

Schmidt, M.W.; Ruedenberg, K.   ([N]s), N=3-10   H
JCP <u>71</u>, 3951 (1979).   ([2N]s), N=3-10   He

(regular even-tempered)   ([2N]s), N=3-14   Li,Be
                                           ([2N]s[N]p),N=3-11   B,N-Ne
                                       ([2N]s[N]p),N=3-13   C
                                       ([2N]s[N]p),N=4-12   Na,Mg
                                       ([2N-6]s[N]p),N=7-15   Al-Ar

D. Polarization Functions

          In addition to the fundamental core and valence basis described above, one usually
adds a set of so-called <u>polarization functions</u> to the basis. Polarization functions are
functions of one higher angular momentum than appears in the atom's valence orbital space
(e.g, d-functions for C, N , and O and p-functions for H). These polarization functions
have exponents (or   ) which cause their radial sizes to be similar to the sizes of the
primary valence orbitals
(i.e., the polarization p orbitals of the H atom are similar in size to the 1s orbital). Thus,
they are <u>not</u> orbitals which provide a description of the atom's valence orbital with one
higher l-value; such higher-l valence orbitals would be radially more diffuse and would
therefore require the use of STOs or GTOs with smaller exponents.
          The primary purpose of polarization functions is to give additional angular
flexibility to the LCAO-MO process in forming the valence molecular orbitals. This is
illustrated below where polarization d   orbitals are seen to contribute to formation of the
bonding   orbital of a carbonyl group by allowing polarization of the Carbon atom's p
orbital toward the right and of the Oxygen atom's p   orbital toward the left.
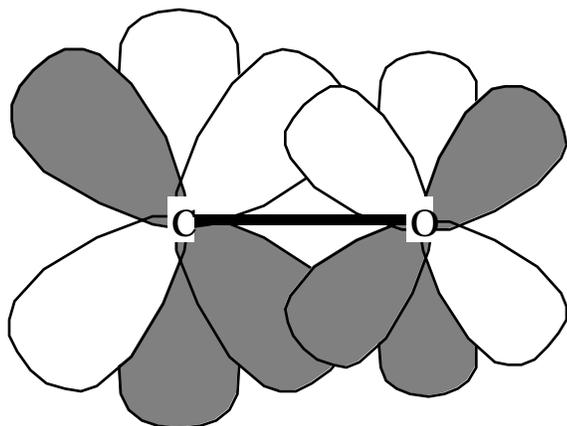
Polarization functions are essential in strained ring compounds because they provide the angular flexibility needed to direct the electron density into regions between bonded atoms.

Functions with higher l-values and with 'sizes' more in line with those of the lower-l orbitals are also used to introduce additional angular correlation into the calculation by permitting polarized orbital pairs (see Chapter 10) involving higher angular correlations to be formed. Optimal polarization functions for first and second row atoms have been tabulated (B. Roos and P. Siegbahn, Theoret. Chim. Acta (Berl.) 17, 199 (1970); M. J. Frisch, J. A. Pople, and J. S. Binkley, J. Chem. Phys. 80 , 3265 (1984)).

E. Diffuse Functions

When dealing with anions or Rydberg states, one must augment the above basis sets by adding so-called diffuse basis orbitals. The conventional valence and polarization functions described above do not provide enough radial flexibility to adequately describe either of these cases. Energy-optimized diffuse functions appropriate to anions of most lighter main group elements have been tabulated in the literature (an excellent source of Gaussian basis set information is provided in Handbook of Gaussian Basis Sets, R. Poirier, R. Kari, and I. G. Csizmadia, Elsevier, Amsterdam (1985)) and in data bases. Rydberg diffuse basis sets are usually created by adding to conventional valence-plus-polarization bases sequences of primitive GTOs whose exponents are smaller than that (call it α_{diff}) of the most diffuse GTO which contributes strongly to the valence CGTOs. As a 'rule of thumb', one can generate a series of such diffuse orbitals which are liniarly independent yet span considerably different regions of radial space by introducing primitive GTOs whose exponents are $\alpha_{diff}/3$, $\alpha_{diff}/9$, $\alpha_{diff}/27$, etc.

Once one has specified an atomic orbital basis for each atom in the molecule, the LCAO-MO procedure can be used to determine the $C_{\nu,i}$ coefficients that describe the occupied and virtual orbitals in terms of the chosen basis set. It is important to keep in mind that the basis orbitals are not themselves the true orbitals of the isolated atoms; even the proper atomic orbitals are combinations (with atomic values for the $C_{\nu,i}$ coefficients) of the basis functions. For example, in a minimal-basis-level treatment of the Carbon atom, the 2s atomic orbital is formed by combining, with opposite sign to achieve the radial node, the two CGTOs (or STOs); the more diffuse s-type basis function will have a larger $C_{i,}$ coefficient in the 2s atomic orbital. The 1s atomic orbital is formed by combining the same two CGTOs but with the same sign and with the less diffuse basis function having a larger

$C_{,i}$ coefficient. The LCAO-MO-SCF process itself determines the magnitudes and signs of the $C_{,i}$.

VI. The Roothaan Matrix SCF Process

The matrix SCF equations introduced earlier

$$F_{\mu,} \ C_{,i} = \ _i \ \ S_{\mu,} \ C_{,i}$$

must be solved both for the occupied and virtual orbitals' energies $_i$ and $C_{,i}$ values. Only the occupied orbitals' $C_{,i}$ coefficients enter into the Fock operator

$$F_{\mu,} \ = < \ \mu \,|\,h\,| \ \ > + \ \ , \ \ [\ , < \ \mu \ \ |\,g\,| \ \ >$$

$$- \ \ , \ ^{ex}< \ \mu \ \ |\,g\,| \ \ >],$$

but both the occupied and virtual orbitals are solutions of the SCF equations. Once atomic basis sets have been chosen for each atom, the one- and two-electron integrals appearing in $F_{\mu,}$ must be evaluated. Doing so is a time consuming process, but there are presently several highly efficient computer codes which allow such integrals to be computed for s, p, d, f, and even g, h, and i basis functions. After executing one of these 'integral packages' for a basis with a total of N functions, one has available (usually on the computer's hard disk) of the order of $\mathbf{N^2/2}$ one-electron and $\mathbf{N^4/8}$ two-electron integrals over these atomic basis orbitals (the factors of 1/2 and 1/8 arise from permutational symmetries of the integrals). When treating extremely large atomic orbital basis sets (e.g., 200 or more basis functions), modern computer programs calculate the requisite integrals but never store them on the disk. Instead, their contributions to $F_{\mu,}$ are accumulated 'on the fly' after which the integrals are discarded.

To begin the SCF process, one must input to the computer routine which computes $F_{\mu,}$ initial 'guesses' for the $C_{,i}$ values corresponding to the occupied orbitals. These initial guesses are typically made in one of the following ways:

1. If one has available $C_{,i}$ values for the system from an SCF calculation performed earlier at a nearby molecular geometry, one can use these $C_{,i}$ values to begin the SCF process.

2. If one has $C_{\mu,i}$ values appropriate to fragments of the system (e.g., for C and O atoms if the CO molecule is under study or for CH_2 and O if H_2CO is being studied), one can use these.

3. If one has no other information available, one can carry out one iteration of the SCF process in which the two-electron contributions to $F_{\mu,\nu}$ are ignored (i.e., take $F_{\mu,\nu} = <\mu|h|\nu>$) and use the resultant solutions to $\sum F_{\mu,\nu} C_{\nu,i} = \epsilon_i \sum S_{\mu,\nu} C_{\nu,i}$ as initial guesses for the $C_{\mu,i}$. Using only the one-electron part of the Hamiltonian to determine initial values for the LCAO-MO coefficients may seem like a rather severe step; it is, and the resultant $C_{\mu,i}$ values are usually far from the converged values which the SCF process eventually produces. However, the initial $C_{\mu,i}$ obtained in this manner have proper symmetries and nodal patterns because the one-electron part of the Hamiltonian has the same symmetry as the full Hamiltonian.

Once initial guesses are made for the $C_{\mu,i}$ of the occupied orbitals, the full $F_{\mu,\nu}$ matrix is formed and new ϵ_i and $C_{\mu,i}$ values are obtained by solving $\sum F_{\mu,\nu} C_{\nu,i} = \epsilon_i \sum S_{\mu,\nu} C_{\nu,i}$. These new orbitals are then used to form a new $F_{\mu,\nu}$ matrix from which new ϵ_i and $C_{\mu,i}$ are obtained. This iterative process is carried on until the ϵ_i and $C_{\mu,i}$ do not vary (within specified tolerances) from iteration to iteration, at which time one says that the SCF process has converged and reached self-consistency.

As presented, the Roothaan SCF process is carried out in a fully <u>ab initio</u> manner in that all one- and two-electron integrals are computed in terms of the specified basis set; no experimental data or other input is employed. As described in Appendix F, it is possible to introduce approximations to the coulomb and exchange integrals entering into the Fock matrix elements that permit many of the requisite $F_{\mu,\nu}$ elements to be evaluated in terms of experimental data or in terms of a small set of 'fundamental' orbital-level coulomb interaction integrals that can be computed in an <u>ab initio</u> manner. This approach forms the basis of so-called 'semi-empirical' methods. Appendix F provides the reader with a brief introduction to such approaches to the electronic structure problem and deals in some detail with the well known Hückel and CNDO- level approximations.

VII. Observations on Orbitals and Orbital Energies

A. The Meaning of Orbital Energies

The physical content of the Hartree-Fock orbital energies can be seen by observing that $F \phi_i = \epsilon_i \phi_i$ implies that ϵ_i can be written as:

$$\varepsilon_i = <\phi_i | F | \phi_i> = <\phi_i | h | \phi_i> + \sum_{j(occupied)} <\phi_i | J_j - K_j | \phi_i>$$

$$= <\phi_i | h | \phi_i> + \sum_{j(occupied)} [J_{i,j} - K_{i,j}].$$

In this form, it is clear that ε_i is equal to the average value of the kinetic energy plus coulombic attraction to the nuclei for an electron in ϕ_i plus the sum over all of the spin-orbitals occupied in ψ of coulomb minus exchange interactions between ϕ_i and these occupied spin-orbitals. If ϕ_i itself is an occupied spin-orbital, the term $[J_{i,i} - K_{i,i}]$ disappears and the latter sum represents the coulomb minus exchange interaction of ϕ_i with all of the N-1 <u>other</u> occupied spin-orbitals. If ϕ_i is a virtual spin-orbital, this cancellation does not occur, and one obtains the coulomb minus exchange interaction of ϕ_i with all N of the occupied spin-orbitals.

In this sense, the orbital energies for occupied orbitals pertain to interactions which are appropriate to a total of N electrons, while the orbital energies of virtual orbitals pertain to a system with N+1 electrons. It is this fact that makes SCF virtual orbitals not optimal (in fact, not usually very good) for use in subsequent correlation calculations where, for instance, they are used, in combination with the occupied orbitals, to form polarized orbital pairs as discussed in Chapter 12. To correlate a pair of electrons that occupy a valence orbital requires double excitations into a virtual orbital that is not too dislike in size. Although the virtual SCF orbitals themselves suffer these drawbacks, the space they span can indeed be used for treating electron correlation. To do so, it is useful to recombine (in a unitary manner to preserve orthonormality) the virtual orbitals to 'focus' the correlating power into as few orbitals as possible so that the multiconfigurational wavefunction can be formed with as few CSFs as possible. Techniques for effecting such reoptimization or improvement of the virtual orbitals are treated later in this text.

B.. Koopmans' Theorem

Further insight into the meaning of the energies of occupied and virtual orbitals can be gained by considering the following model of the vertical (i.e., at fixed molecular geometry) detachment or attachment of an electron to the original N-electron molecule:
1. In this model, <u>both</u> the parent molecule and the species generated by adding or removing an electron are treated at the single-determinant level.
2. In this model, the Hartree-Fock orbitals of the parent molecule are used to describe both the parent and the species generated by electron addition or removal. It is said that such a model neglects '<u>orbital relaxation</u>' which would accompany the electron addition or

removal (i.e., the reoptimization of the spin-orbitals to allow them to become appropriate to the daughter species).

Within this simplified model, the energy difference between the daughter and the parent species can be written as follows (ϕ_k represents the particular spin-orbital that is added or removed):

1. For electron detachment:

$$E^{N-1} - E^N = \langle |\phi_1 \phi_2 \cdots \phi_{k-1} \cdots \phi_N| H |\phi_1 \phi_2 \cdots \phi_{k-1} \cdots \phi_N| \rangle -$$

$$\langle |\phi_1 \phi_2 \cdots \phi_{k-1} \phi_k \cdots \phi_N| H || \phi_1 \phi_2 \cdots \phi_{k-1} \phi_k \cdots \phi_N| \rangle$$

$$= - \langle \phi_k | h | \phi_k \rangle - \sum_{j=(1,k-1,k+1,N)} [J_{k,j} - K_{k,j}] = -\epsilon_k \, ;$$

2. For electron attachment:

$$E^N - E^{N+1} = \langle |\phi_1 \phi_2 \cdots \phi_N| H |\phi_1 \phi_2 \cdots \phi_N| \rangle -$$

$$\langle |\phi_1 \phi_2 \cdots \phi_N \phi_k| H || \phi_1 \phi_2 \cdots \phi_N \phi_k| \rangle$$

$$= - \langle \phi_k | h | \phi_k \rangle - \sum_{j=(1,N)} [J_{k,j} - K_{k,j}] = -\epsilon_k \, .$$

So, within the limitations of the single-determinant, frozen-orbital model set forth, the ionization potentials (IPs) and electron affinities (EAs) are given as the negative of the occupied and virtual spin-orbital energies, respectively. This statement is referred to as Koopmans' theorem (T. Koopmans, Physica $\underline{1}$, 104 (1933)); it is used extensively in quantum chemical calculations as a means for estimating IPs and EAs and often yields results that are at least qualitatively correct (i.e., ± 0.5 eV).

C. Orbital Energies and the Total Energy

For the N-electron species whose Hartree-Fock orbitals and orbital energies have been determined, the total SCF electronic energy can be written, by using the Slater-Condon rules, as:

$$E = \sum_{i(occupied)} \langle \phi_i | h | \phi_i \rangle + \sum_{i>j(occupied)} [J_{i,j} - K_{i,j}].$$

For this same system, the sum of the orbital energies of the occupied spin-orbitals is given by:

$$\sum_{i(occupied)} \epsilon_i = \sum_{i(occupied)} < \phi_i \,|\, h \,|\, \phi_i >$$

$$+ \sum_{i,j(occupied)} [J_{i,j} - K_{i,j}].$$

These two seemingly very similar expressions differ in a very important way; the sum of occupied orbital energies, when compared to the total energy, double counts the coulomb minus exchange interaction energies. Thus, within the Hartree-Fock approximation, the sum of the occupied orbital energies is <u>not</u> equal to the total energy. The total SCF energy can be computed in terms of the sum of occupied orbital energies by taking one-half of $\sum_{i(occupied)} \epsilon_i$ and then adding to this one-half of $\sum_{i(occupied)} < \phi_i \,|\, h \,|\, \phi_i >$:

$$E = 1/2 [\sum_{i(occupied)} < \phi_i \,|\, h \,|\, \phi_i > + \sum_{i(occupied)} \epsilon_i].$$

The fact that the sum of orbital energies is not the total SCF energy also means that as one attempts to develop a qualitative picture of the energies of CSFs along a reaction path, as when orbital and configuration correlation diagrams are constructed, one must be careful not to equate the sum of orbital energies with the total configurational energy; the former is higher than the latter by an amount equal to the sum of the coulomb minus exchange interactions.

D. The Brillouin Theorem

The condition that the SCF energy $<\phi_1... \phi_N | H | \phi_1... \phi_N>$ be stationary with respect to variations $\delta\phi_i$ in the occupied spin-orbitals (that preserve orthonormality) can be written

$$<\delta\phi_1... \delta\phi_i... \delta\phi_N | H | \phi_1... \phi_i... \phi_N> = 0.$$

The infinitesimal variation of ϕ_i can be expressed in terms of its (small) components along the other occupied ϕ_j and along the virtual ϕ_m as follows:

$$\delta\phi_i = \sum_{j=occ} U_{ij} \phi_j + \sum_m U_{im} \phi_m.$$

When substituted into $|\phi_1\ldots\,\phi_i\ldots\,|$, the terms $\sum_{j'=occ}|\phi_1\ldots\,\phi_j\ldots\,\phi_N|U_{ij}$ vanish because ϕ_j already appears in the original Slater determinant $|\phi_1\ldots\,\phi_N|$, so $|\phi_1\ldots\,\phi_j\ldots\,|$ contains ϕ_j twice. Only the sum over virtual orbitals remains, and the stationary property written above becomes

$$\sum_m U_{im}\langle|\phi_1\ldots\,\phi_m\ldots\,\phi_N|\,H\,|\phi_1\ldots\,\phi_i\ldots\,\phi_N|\rangle = 0.$$

The Slater-Condon rules allow one to express the Hamiltonian matrix elements appearing here as

$$\langle|\phi_1\ldots\,\phi_m\ldots\,\phi_N|\,H\,|\phi_1\ldots\,\phi_i\ldots\,\phi_N|\rangle = \langle\phi_m|h|\phi_i\rangle + \sum_{j=occ,\ \neq i}\langle\phi_m|[J_j-K_j]|\phi_i\rangle,$$

which (because the term with j=i can be included since it vanishes) is equal to the following element of the Fock operator: $\langle\phi_m|F|\phi_i\rangle = \varepsilon_i\,\delta_{im} = 0$. This result proves that Hamiltonian matrix elements between the SCF determinant and those that are singly excited relative to the SCF determinant vanish because they reduce to Fock-operator integrals connecting the pair of orbitals involved in the 'excitation'. This stability property of the SCF energy is known as the Brillouin theorem (i.e., that $|\phi_1\,\phi_i\,\phi_N|$ and $|\phi_1\ldots\,\phi_m\ldots\,\phi_N|$ have zero Hamiltonian matrix elements <u>if</u> the ϕs are SCF orbitals). It is exploited in quantum chemical calculations in two manners:

(i) When multiconfiguration wavefunctions are formed from SCF spin-orbitals, it allows one to neglect Hamiltonian matrix elements between the SCF configuration and those that are 'singly excited' in constructing the secular matrix.

(ii) A so-called generalized Brillouin theorem (GBT) arises when one deals with energy optimization for a multiconfigurational variational trial wavefunction for which the orbitals and C_I mixing coefficients are simultaneously optimized. This GBT causes certain Hamiltonian matrix elements to vanish, which, in turn, simplifies the treatment of electron correlation for such wavefunctions. This matter is treated in more detail later in this text.

Chapter 19

Corrections to the mean-field model are needed to describe the instantaneous Coulombic interactions among the electrons. This is achieved by including more than one Slater determinant in the wavefunction.

Much of the development of the previous chapter pertains to the use of a single Slater determinant trial wavefunction. As presented, it relates to what has been called the

unrestricted Hartree-Fock (UHF) theory in which each spin-orbital ϕ_i has its own orbital energy ϵ_i and LCAO-MO coefficients $C_{\nu,i}$; there may be different $C_{\nu,i}$ for α spin-orbitals than for β spin-orbitals. Such a wavefunction suffers from the spin contamination difficulty detailed earlier.

To allow for a properly spin- and space- symmetry adapted trial wavefunction and to permit Ψ to contain more than a single CSF, methods which are more flexible than the single-determinant HF procedure are needed. In particular, it may be necessary to use a combination of determinants to describe such a proper symmetry function. Moreover, as emphasized earlier, whenever two or more CSFs have similar energies (i.e., Hamiltonian expectation values) and can couple strongly through the Hamiltonian (e.g., at avoided crossings in configuration correlation diagrams), the wavefunction must be described in a multiconfigurational manner to permit the wavefunction to evolve smoothly from reactants to products. Also, whenever dynamical electron correlation effects are to be treated, a multiconfigurational Ψ must be used; in this case, CSFs that are <u>doubly excited</u> relative to one or more of the essential CSFs (i.e., the dominant CSFs that are included in the so-called <u>reference wavefunction</u>) are included to permit polarized-orbital-pair formation.

Multiconfigurational functions are needed not only to account for electron correlation but also to permit orbital readjustments to occur. For example, if a set of SCF orbitals is employed in forming a multi-CSF wavefunction, the variational condition that the energy is stationary with respect to variations in the LCAO-MO coefficients is no longer obeyed (i.e., the SCF energy functional is stationary when SCF orbitals are employed, but the MC-energy functional is generally not stationary if SCF orbitals are employed). For such reasons, it is important to include CSFs that are <u>singly excited</u> relative to the dominant CSFs in the reference wavefunction.

That singly excited CSFs allow for orbital relaxation can be seen as follows. Consider a wavefunction consisting of one CSF $|\phi_1\ldots\phi_i\ldots\phi_N|$ to which singly excited CSFs of the form $|\phi_1\ldots\phi_m\ldots\phi_N|$ have been added with coefficients $C_{i,m}$:

$$\Psi = \sum_m C_{i,m} |\phi_1\cdots\phi_m\cdots\phi_N| + |\phi_1\cdots\phi_i\cdots\phi_N|.$$

All of these determinants have all of their columns equal except the ith column; therefore, they can be combined into a single new determinant:

$$\Psi = |\phi_1\cdots\phi_i'\cdots\phi_N|,$$

where the relaxed orbital ϕ_i' is given by

$$\phi_i' = \phi_i + \sum_m C_{i,m}\phi_m.$$

The sum of CSFs that are singly excited in the ith spin-orbital with respect to $|\phi_1 \ldots \phi_i \ldots \phi_N|$ is therefore seen to allow the spin-orbital ϕ_i to relax into the new spin-orbital ϕ_i'. It is in this sense that singly excited CSFs allow for orbital reoptimization.

In summary, doubly excited CSFs are often employed to permit polarized orbital pair formation and hence to allow for electron correlations. Singly excited CSFs are included to permit orbital relaxation (i.e., orbital reoptimization) to occur.

I. Different Methods

There are numerous procedures currently in use for determining the 'best' wavefunction of the form:

$$\Psi = \sum_I C_I \Phi_I,$$

where Φ_I is a spin-and space- symmetry adapted CSF consisting of determinants of the form $|\phi_{I1} \phi_{I2} \phi_{I3} \ldots \phi_{IN}|$. Excellent overviews of many of these methods are included in <u>Modern Theoretical Chemistry</u> Vols. 3 and 4, H. F. Schaefer, III Ed., Plenum Press, New York (1977) and in <u>Advances in Chemical Physics</u>, Vols. LXVII and LXIX, K. P. Lawley, Ed., Wiley-Interscience, New York (1987). Within the present Chapter, these two key references will be denoted MTC, Vols. 3 and 4, and ACP, Vols. 67 and 69, respectively.

In all such trial wavefunctions, there are two fundamentally different kinds of parameters that need to be determined- the CI coefficients C_I and the LCAO-MO coefficients describing the ϕ_{Ik}. The most commonly employed methods used to determine these parameters include:

1. The **multiconfigurational self-consistent field** (MCSCF) method in which the expectation value $<\Psi|H|\Psi>/<\Psi|\Psi>$ is treated variationally and simultaneously made stationary with respect to variations in the C_I and $C_{\nu,i}$ coefficients subject to the constraints that the spin-orbitals and the full N-electron wavefunction remain normalized:

$$<\phi_i|\phi_j> = \delta_{i,j} = \sum_{\nu,\mu} C_{\nu,i} S_{\nu,\mu} C_{\mu,i} \text{, and}$$

$$\sum_I C_I^2 = 1.$$

The articles by H.-J. Werner and by R. Shepard in ACP Vol. 69 provide up to date reviews of the status of this approach. The article by A. C. Wahl and G. Das in MTC Vol. 3 covers the 'earlier' history on this topic. F. W. Bobrowicz and W. A. Goddard, III provide, in MTC Vol. 3, an overview of the GVB approach, which, as discussed in Chapter 12, can be viewed as a specific kind of MCSCF calculation.

2. The **configuration interaction** (CI) method in which the LCAO-MO coefficients are determined first (and independently) via either a single-configuration SCF calculation or an MCSCF calculation using a small number of CSFs. The CI coefficients are subsequently determined by making the expectation value $< \Psi | H | \Psi > / < \Psi | \Psi >$ stationary with respect to variations in the C_I only. In this process, the optimizations of the orbitals and of the CSF amplitudes are done in separate steps. The articles by I. Shavitt and by B. O. Ross and P. E. M. Siegbahn in MTC, Vol. 3 give excellent early overviews of the CI method.

3. The **Møller-Plesset perturbation method** (MPPT) uses the single-configuration SCF process (usually the UHF implementation) to first determine a set of LCAO-MO coefficients and, hence, a set of orbitals that obey $F \phi_i = \epsilon_i \phi_i$. Then, using an unperturbed Hamiltonian equal to the sum of these Fock operators for each of the N electrons $H^0 = \sum_{i=1,N} F(i)$, perturbation theory (see Appendix D for an introduction to time-independent perturbation theory) is used to determine the C_I amplitudes for the CSFs. The MPPT procedure is also referred to as the many-body perturbation theory (MBPT) method. The two names arose because two different schools of physics and chemistry developed them for somewhat different applications. Later, workers realized that they were identical in their working equations when the UHF H^0 is employed as the unperturbed Hamiltonian. In this text, we will therefore refer to this approach as MPPT/MBPT.

The amplitude for the so-called <u>reference</u> CSF used in the SCF process is taken as unity and the other CSFs' amplitudes are determined, relative to this one, by Rayleigh-Schrödinger perturbation theory using the full N-electron Hamiltonian minus the sum of Fock operators $H-H^0$ as the perturbation. The Slater-Condon rules are used for evaluating matrix elements of $(H-H^0)$ among these CSFs. The essential features of the MPPT/MBPT approach are described in the following articles: J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. <u>14</u>, 545 (1978); R. J. Bartlett and D. M. Silver,

J. Chem. Phys. <u>62</u>, 3258 (1975); R. Krishnan and J. A. Pople, Int. J. Quantum Chem. <u>14</u>, 91 (1978).

4. The **Coupled-Cluster method** expresses the CI part of the wavefunction in a somewhat different manner (the early work in chemistry on this method is described in J. Cizek, J. Chem. Phys. <u>45</u>, 4256 (1966); J. Paldus, J. Cizek, and I. Shavitt, Phys. Rev. <u>A5</u>, 50 (1972); R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. <u>14</u>, 561 (1978); G. D. Purvis and R. J. Bartlett, J. Chem. Phys. <u>76</u>, 1910 (1982)):

$$\Psi = \exp(T)\,\Phi \quad,$$

where Φ is a single CSF (usually the UHF single determinant) which has been used to independently determine a set of spin-orbitals and LCAO-MO coefficients via the SCF process. The operator T generates, when acting on Φ, single, double, etc. 'excitations' (i.e., CSFs in which one, two, etc. of the occupied spin-orbitals in Φ have been replaced by virtual spin-orbitals). T is commonly expressed in terms of operators that effect such spin-orbital removals and additions as follows:

$$T = \sum_{i,m} t_i^m\, m^+\, i + \sum_{i,j,m,n} t_{i,j}^{m,n}\, m^+\, n^+\, j\, i + ...,$$

where the operator m^+ is used to denote <u>creation</u> of an electron in virtual spin-orbital ϕ_m and the operator j is used to denote <u>removal</u> of an electron from occupied spin-orbital ϕ_j.

The t_i^m, $t_{i,j}^{m,n}$, etc. amplitudes, which play the role of the CI coefficients in CC theory, are determined through the set of equations generated by projecting the Schrödinger equation in the form

$$\exp(-T)\, H\, \exp(T)\, \Phi = E\, \Phi$$

against CSFs which are single, double, etc. excitations relative to Φ. For example, for double excitations $\Phi_{i,j}^{m,n}$ the equations read:

$$< \Phi_{i,j}^{m,n} \mid \exp(-T)\, H\, \exp(T) \mid \Phi > = E < \Phi_{i,j}^{m,n} \mid \Phi > = 0;$$

zero is obtained on the right hand side because the excited CSFs

$|\Phi_{i,j}^{m,n}\rangle$ are orthogonal to the reference function $|\Phi\rangle$. The elements on the left hand side of the CC equations can be expressed, as described below, in terms of one- and two-electron integrals over the spin-orbitals used in forming the reference and excited CSFs.

A. Integral Transformations

All of the above methods require the evaluation of one- and two-electron integrals over the N atomic <u>orbital basis:</u> $\langle \chi_a |f| \chi_b \rangle$ and $\langle \chi_a \chi_b |g| \chi_c \chi_d \rangle$. Eventually, all of these methods provide their working equations and energy expressions in terms of one- and two-electron integrals over the N final <u>molecular orbitals:</u> $\langle \phi_i |f| \phi_j \rangle$ and $\langle \phi_i \phi_j |g| \phi_k \phi_l \rangle$. The mo-based integrals can only be evaluated by <u>transforming</u> the AO-based integrals as follows:

$$\langle \phi_i \phi_j |g| \phi_k \phi_l \rangle = \sum_{a,b,c,d} C_{a,i} C_{b,j} C_{c,k} C_{d,l} \langle \chi_a \chi_b |g| \chi_c \chi_d \rangle,$$

and

$$\langle \phi_i |f| \phi_j \rangle = \sum_{a,b} C_{a,i} C_{b,j} \langle \chi_a |f| \chi_b \rangle.$$

It would seem that the process of evaluating all N^4 of the $\langle \phi_i \phi_j |g| \phi_k \phi_l \rangle$, each of which requires N^4 additions and multiplications, would require computer time proportional to N^8. However, it is possible to perform the full transformation of the two-electron integral list in a time that scales as $\mathbf{N^5}$. This is done by first performing a transformation of the $\langle \chi_a \chi_b |g| \chi_c \chi_d \rangle$ to an intermediate array labeled $\langle \chi_a \chi_b |g| \chi_c \phi_l \rangle$ as follows:

$$\langle \chi_a \chi_b |g| \chi_c \phi_l \rangle = \sum_d C_{d,l} \langle \chi_a \chi_b |g| \chi_c \chi_d \rangle.$$

This partial transformation requires N^5 multiplications and additions. The list $\langle \chi_a \chi_b |g| \chi_c \phi_l \rangle$ is then transformed to a second-level transformed array $\langle \chi_a \chi_b |g| \phi_k \phi_l \rangle$:

$$\langle \chi_a \chi_b |g| \phi_k \phi_l \rangle = \sum_c C_{c,k} \langle \chi_a \chi_b |g| \chi_c \phi_l \rangle,$$

which requires another N^5 operations. This sequential, one-index-at-a-time transformation is repeated four times until the final $\langle \phi_i \phi_j |g| \phi_k \phi_l \rangle$ array is in hand. The entire transformation done this way requires $\mathbf{4N^5}$ multiplications and additions.

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. These transformations consume a large fraction of the computer time used in most such calculations, and represent a severe bottleneck to progress in applying ab initio electronic structure methods to larger systems.

B. Configuration List Choices

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. Each of these methods has its own approach to describing the configurations $\{\Phi_J\}$ included in the calculation and how the $\{C_J\}$ amplitudes and the total energy E is to be determined.

The number of configurations (N_C) varies greatly among the methods and is an important factor to keep in mind when planning to carry out an *ab initio* calculation. Under certain circumstances (e.g., when studying Woodward-Hoffmann forbidden reactions where an avoided crossing of two configurations produces an activation barrier), it may be essential to use more than one electronic configuration. Sometimes, one configuration (e.g., the SCF model) is adequate to capture the qualitative essence of the electronic structure. In all cases, many configurations will be needed if highly accurate treatment of electron-electron correlations are desired.

The value of N_C determines how much computer time and memory is needed to solve the N_C-dimensional $\sum_J H_{I,J} C_J = E C_I$ secular problem in the CI and MCSCF methods. Solution of these matrix eigenvalue equations requires computer time that scales as N_C^2 (if few eigenvalues are computed) to N_C^3 (if most eigenvalues are obtained).

So-called complete-active-space (CAS) methods form all CSFs that can be created by distributing N valence electrons among P valence orbitals. For example, the eight non-core electrons of H_2O might be distributed, in a manner that gives $M_S = 0$, among six valence orbitals (e.g., two lone-pair orbitals, two OH bonding orbitals, and two OH* antibonding orbitals). The number of configurations thereby created is 225 . If the same eight electrons were distributed among ten valence orbitals 44,100 configurations results; for twenty and thirty valence orbitals, 23,474,025 and 751,034,025 configurations arise, respectively. Clearly, practical considerations dictate that CAS-based approaches be limited to situations in which a few electrons are to be correlated using a few valence orbitals. The primary advantage of CAS configurations is discussed below in Sec. II. C.

II. Strengths and Weaknesses of Various Methods

A. Variational Methods Such as MCSCF, SCF, and CI Produce Energies that are Upper Bounds, but These Energies are not Size-Extensive

Methods that are based on making the energy functional $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$ stationary (i.e., variational methods) yield upper bounds to the lowest energy of the symmetry which characterizes the CSFs which comprise Ψ. These methods also can provide approximate excited-state energies and wavefunctions (e. g., in the form of other solutions of the secular equation $\sum_J H_{I,J} C_J = E C_I$ that arises in the CI and MCSCF methods). Excited-state energies obtained in this manner can be shown to 'bracket' the true energies of the given symmetry in that between any two approximate energies obtained in the variational calculation, there exists at least one true eigenvalue. This characteristic is commonly referred to as the 'bracketing theorem' (E. A. Hylleraas and B. Undheim, Z. Phys. 65, 759 (1930); J. K. L. MacDonald, Phys. Rev. 43, 830 (1933)). These are strong attributes of the variational methods, as is the long and rich history of developments of analytical and computational tools for efficiently implementing such methods (see the discussions of the CI and MCSCF methods in MTC and ACP).

However, all variational techniques suffer from at least one serious drawback; they are not **size-extensive** (J. A. Pople, pg. 51 in Energy, Structure, and Reactivity, D. W. Smith and W. B. McRae, Eds., Wiley, New York (1973)). This means that the energy computed using these tools can not be trusted to scale with the size of the system. For example, a calculation performed on two CH_3 species at large separation may not yield an energy equal to twice the energy obtained by performing the same kind of calculation on a single CH_3 species. Lack of size-extensivity precludes these methods from use in extended systems (e.g., solids) where errors due to improper scaling of the energy with the number of molecules produce nonsensical results.

By carefully adjusting the kind of variational wavefunction used, it is possible to circumvent size-extensivity problems for selected species. For example, a CI calculation on Be_2 using all $^1\Sigma_g$ CSFs that can be formed by placing the four valence electrons into the orbitals $2\sigma_g, 2\sigma_u, 3\sigma_g, 3\sigma_u, 1\pi_u$, and $1\pi_g$ can yield an energy equal to twice that of the Be atom described by CSFs in which the two valence electrons of the Be atom are placed into the 2s and 2p orbitals in all ways consistent with a ^1S symmetry. Such special choices of configurations give rise to what are called complete-active-space (CAS) MCSCF or CI calculations (see the article by B. O. Roos in ACP for an overview of this approach).

Let us consider an example to understand why the CAS choice of configurations works. The ^1S ground state of the Be atom is known to form a wavefunction that is a strong mixture of CSFs that arise from the $2s^2$ and $2p^2$ configurations:

$$\Psi_{Be} = C_1\,|1s^2\,2s^2\,| + C_2\,|\,1s^2\,2p^2\,|,$$

where the latter CSF is a short-hand representation for the proper spin- and space-symmetry adapted CSF

$$|\,1s^2\,2p^2\,| = 1/\sqrt{3}\,[\,|1s\;\;1s\;\;2p_0\;\;2p_0\;\,| - |1s\;\;1s\;\;2p_1\;\;2p_{-1}\;\,|$$

$$-\,|1s\;\;1s\;\;2p_{-1}\;\;2p_1\;\,|\,].$$

The reason the CAS process works is that the Be_2 CAS wavefunction has the flexibility to dissociate into the product of two CAS Be wavefunctions:

$$\Psi = \Psi_{Be_a}\,\Psi_{Be_b}$$

$$= \{C_1\,|1s^2\,2s^2\,| + C_2\,|\,1s^2\,2p^2\,|\}_a\{C_1\,|1s^2\,2s^2\,| + C_2\,|\,1s^2\,2p^2\,|\}_b,$$

where the subscripts a and b label the two Be atoms, because the four electron CAS function distributes the four electrons in all ways among the $2s_a$, $2s_b$, $2p_a$, and $2p_b$ orbitals. In contrast, if the Be_2 calculation had been carried out using only the following CSFs : $|\,1\sigma_g^2\,1\sigma_u^2\,2\sigma_g^2\,2\sigma_u^2\,|$ and all single and double excitations relative to this (dominant) CSF, which is a very common type of CI procedure to follow, the Be_2 wavefunction would not have contained the particular CSFs $|\,1s^2\,2p^2\,|_a\,|\,1s^2\,2p^2\,|_b$ because these CSFs are four-fold excited relative to the $|\,1\sigma_g^2\,1\sigma_u^2\,2\sigma_g^2\,2\sigma_u^2\,|$ 'reference' CSF.

In general, one finds that if the 'monomer' uses CSFs that are K-fold excited relative to its dominant CSF to achieve an accurate description of its electron correlation, a size-extensive variational calculation on the 'dimer' will require the inclusion of CSFs that are 2K-fold excited relative to the dimer's dominant CSF. To perform a size-extensive variational calculation on a species containing M monomers therefore requires the inclusion of CSFs that are MxK-fold excited relative to the M-mer's dominant CSF.

B. Non-Variational Methods Such as MPPT/MBPT and CC do not Produce Upper Bounds, but Yield Size-Extensive Energies

In contrast to variational methods, perturbation theory and coupled-cluster methods achieve their energies from a 'transition formula' $< \Phi | H | \Psi >$ rather than from an expectation value
$< \Psi | H | \Psi >$. It can be shown (H. P. Kelly, Phys. Rev. <u>131</u>, 684 (1963)) that this difference allows non-variational techniques to yield size-extensive energies. This can be seen in the MPPT/MBPT case by considering the energy of two non-interacting Be atoms. The reference CSF is $\Phi = | 1s_a^2\, 2s_a^2\, 1s_b^2\, 2s_b^2 |$; the Slater-Condon rules limit the CSFs in Ψ which can contribute to

$$E = < \Phi | H | \Psi > = < \Phi | H | \sum_J C_J \Phi_J >,$$

to be Φ itself and those CSFs that are singly or doubly excited relative to Φ. These 'excitations' can involve atom a, atom b, or both atoms. However, any CSFs that involve excitations on both atoms
(e.g., $| 1s_a^2\, 2s_a\, 2p_a\, 1s_b^2\, 2s_b\, 2p_b |$) give rise, via the SC rules, to one- and two- electron integrals over orbitals on both atoms; these integrals (e.g., $< 2s_a\, 2p_a | g | 2s_b\, 2p_b >$) vanish if the atoms are far apart, as a result of which the contributions due to such CSFs vanish in our consideration of size-extensivity. Thus, only CSFs that are excited on one or the other atom contribute to the energy:

$$E = < \Phi_a\, \Phi_b | H | \sum_{Ja} C_{Ja}\, \Phi^*_{Ja}\, \Phi_b + \sum_{Jb} C_{Jb}\, \Phi_a\, \Phi^*_{Jb} >,$$

where Φ_a and Φ_b as well as Φ^*_{Ja} and Φ^*_{Jb} are used to denote the a and b parts of the reference and excited CSFs, respectively.

This expression, once the SC rules are used to reduce it to one- and two- electron integrals, is of the additive form required of any size-extensive method:

$$E = < \Phi_a | H | \sum_{Ja} C_{Ja}\, \Phi^*_{Ja} > + < \Phi_b | H | \sum_{Jb} C_{Jb}\, \Phi^*_{Jb} >,$$

and will yield a size-extensive energy <u>if</u> the equations used to determine the C_{Ja} and C_{Jb} amplitudes are themselves separable. In MPPT/MBPT, these amplitudes are expressed, in first order, as:

$$C_{Ja} = < \Phi_a\, \Phi_b | H | \Phi^*_{Ja}\, \Phi_b > / [E^0_a + E^0_b - E^*_{Ja} - E^0_b]$$

(and analogously for C_{Jb}). Again using the SC rules, this expression reduces to one that involves only atom a:

$$C_{Ja} = < \Phi_a | H | \Phi^*_{Ja} >/[E^0_a - E^*_{Ja}].$$

The additivity of E and the separability of the equations determining the C_J coefficients make the MPPT/MBPT energy size-extensive. This property can also be demonstrated for the Coupled-Cluster energy (see the references given above in Chapter 19. I.4). However, size-extensive methods have at least one serious weakness; their energies do <u>not</u> provide upper bounds to the true energies of the system (because their energy functional is not of the expectation-value form for which the upper bound property has been proven).

C. Which Method is Best?

At this time, it may not possible to say which method is preferred for applications where all are practical. Nor is it possible to assess, in a way that is applicable to most chemical species, the accuracies with which various methods predict bond lengths and energies or other properties. However, there are reasons to recommend some methods over others in specific cases.   For example, certain applications require a size-extensive energy (e.g., extended systems that consist of a large or macroscopic number of units or studies of weak intermolecular interactions), so MBPT/MPPT or CC or CAS-based MCSCF are preferred. Moreover, certain chemical reactions (e.g., Woodward-Hoffmann forbidden reactions) and certain bond-breaking events require two or more 'essential' electronic configurations. For them, single-configuration-based methods such as conventional CC and MBTP/MPPT should not be used; MCSCF or CI calculations would be better. Very large molecules, in which thousands of atomic orbital basis functions are required, may be impossible to treat by methods whose effort scales as N^4 or higher; density functional methods would be better to use then.

For all calculations, the choice of atomic orbital basis set must be made carefully, keeping in mind the N^4 scaling of the one- and two-electron integral evaluation step and the N^5 scaling of the two-electron integral transformation step. Of course, basis functions that describe the essence of the states to be studied are essential (e.g., Rydberg or anion states require diffuse functions, and strained rings require polarization functions).

As larger atomic basis sets are employed, the size of the CSF list used to treat dynamic correlation increases rapidly. For example, most of the above methods use singly and doubly excited CSFs for this purpose. For large basis sets, the number of such CSFs,

N_C, scales as the number of electrons squared, n_e^2, times the number of basis functions squared, N^2. Since the effort needed to solve the CI secular problem varies as N_C^2 or N_C^3, a dependence as strong as N^4 to N^6 can result. To handle such large CSF spaces, all of the multiconfigurational techniques mentioned in this paper have been developed to the extent that calculations involving of the order of 100 to 5,000 CSFs are routinely performed and calculations using 10,000, 100,000, and even several million CSFs are practical.

Other methods, most of which can be viewed as derivatives of the techniques introduced above, have been and are still being developed. This ongoing process has been, in large part, stimulated by the explosive growth in computer power and change in computer architecture that has been realized in recent years. All indications are that this growth pattern will continue, so *ab initio* quantum chemistry will likely have an even larger impact on future chemistry research and education (through new insights and concepts).

III. Further Details on Implementing Multiconfigurational Methods

A. The MCSCF Method

The simultaneous optimization of the LCAO-MO and CI coefficients performed within an MCSCF calculation is a quite formidable task. The variational energy functional is a quadratic function of the CI coefficients, and so one can express the stationary conditions for these variables in the secular form:

$$\sum_J H_{I,J} \, C_J \; = E \, C_I \, .$$

However, E is a quartic function of the $C_{\nu,i}$ coefficients because each matrix element $< \Phi_I | H | \Phi_J >$ involves one- and two-electron integrals over the mos ϕ_i, and the two-electron integrals depend quartically on the $C_{\nu,i}$ coefficients. The stationary conditions with respect to these $C_{\nu,i}$ parameters must be solved iteratively because of this quartic dependence.

It is well known that minimization of a function (E) of several non-linear parameters (the $C_{\nu,i}$) is a difficult task that can suffer from poor convergence and may locate local rather than global minima. In an MCSCF wavefunction containing many CSFs, the energy is only weakly dependent on the orbitals that are weakly occupied (i.e., those that appear in CSFs with small C_I values); in contrast, E is strongly dependent on the $C_{\nu,i}$ coefficients of those orbitals that appear in the CSFs with larger C_I values. One is therefore faced with minimizing a function of many variables (there may be as many $C_{\nu,i}$ as the square of the

number of orbital basis functions) that depends strongly on several of the variables and weakly on many others. This is a very difficult job.

For these reasons, in the MCSCF method, the number of CSFs is usually kept to a small to moderate number (e.g., a few to several hundred) chosen to describe essential correlations (i.e., configuration crossings, proper dissociation) and important dynamical correlations (those electron-pair correlations of angular, radial, left-right, etc. nature that arise when low-lying 'virtual' orbitals are present). In such a compact wavefunction, only spin-orbitals with reasonably large occupations (e.g., as characterized by the diagonal elements of the one-particle density matrix $\gamma_{i,j}$) appear. As a result, the energy functional is expressed in terms of variables on which it is strongly dependent, in which case the non-linear optimization process is less likely to be pathological.

Such a compact MCSCF wavefunction is designed to provide a good description of the set of strongly occupied spin-orbitals and of the CI amplitudes for CSFs in which only these spin-orbitals appear. It, of course, provides no information about the spin-orbitals that are not used to form the CSFs on which the MCSCF calculation is based. As a result, the MCSCF energy is invariant to a unitary transformation among these 'virtual' orbitals.

In addition to the references mentioned earlier in ACP and MTC, the following papers describe several of the advances that have been made in the MCSCF method, especially with respect to enhancing its rate and range of convergence: E. Dalgaard and P. Jørgensen, J. Chem. Phys. <u>69</u>, 3833 (1978); H. J. Aa. Jensen, P. Jørgensen, and H. Ågren, J. Chem. Phys. <u>87</u>, 457 (1987); B. H. Lengsfield, III and B. Liu, J. Chem. Phys. <u>75</u>, 478 (1981).

B. The Configuration Interaction Method

In the CI method, one usually attempts to realize a high-level treatment of electron correlation. A set of orthonormal molecular orbitals are first obtained from an SCF or MCSCF calculation (usually involving a small to moderate list of CSFs). The LCAO-MO coefficients of these orbitals are <u>no longer considered</u> as variational parameters in the subsequent CI calculation; only the C_I coefficients are to be further optimized.

The CI wavefunction

$$\Psi = \sum_J C_J \Phi_J$$

is most commonly constructed from CSFs Φ_J that include:

1. All of the CSFs in the SCF (in which case only a single CSF is included) or MCSCF wavefunction that was used to generate the molecular orbitals ϕ_i . This set of CSFs are referred to as spanning the 'reference space' of the subsequent CI calculation, and the particular combination of these CSFs used in this orbital optimization (i.e., the SCF or MCSCF wavefunction) is called the reference function.

2. CSFs that are generated by carrying out single, double, triple, etc. level 'excitations' (i.e., orbital replacements) relative to reference CSFs. CI wavefunctions limited to include contributions through various levels of excitation (e.g., single, double, etc.) are denoted S (singly excited), D (doubly), SD (singly and doubly), SDT (singly, doubly, and triply), and so on.

The orbitals from which electrons are removed and those into which electrons are excited can be restricted to focus attention on correlations among certain orbitals. For example, if excitations out of core electrons are excluded, one computes a total energy that contains no correlation corrections for these core orbitals. Often it is possible to so limit the nature of the orbital excitations to focus on the energetic quantities of interest (e.g., the CC bond breaking in ethane requires correlation of the σ_{CC} orbital but the 1s Carbon core orbitals and the CH bond orbitals may be treated in a non-correlated manner).

Clearly, the number of CSFs included in the CI calculation can be far in excess of the number considered in typical MCSCF calculations; CI wavefunctions including 5,000 to 50,000 CSFs are routinely used, and functions with one to several million CSFs are within the realm of practicality (see, for example, J. Olsen, B. Roos, Poul Jørgensen, and H. J. Aa. Jensen, J. Chem. Phys. 89, 2185 (1988) and J. Olsen, P. Jørgensen, and J. Simons, Chem. Phys. Letters 169, 463 (1990)).

The need for such large CSF expansions should not come as a surprise once one considers that (i) each electron pair requires at least two CSFs (let us say it requires P of them, on average, a dominant one and P-1 others which are doubly excited) to form polarized orbital pairs, (ii) there are of the order of N(N-1)/2 = X electron pairs in an atom or molecule containing N electrons, and (iii) that the number of terms in the CI wavefunction scales as P^X. So, for an H_2O molecule containing ten electrons, there would be P^{55} terms in the CI expansion. This is 3.6×10^{16} terms if P=2 and 1.7×10^{26} terms if P=3. Undoubtedly, this is an over estimate of the number of CSFs needed to describe electron correlation in H_2O, but it demonstrates how rapidly the number of CSFs can grow with the number of electrons in the system.

The $H_{I,J}$ matrices that arise in CI calculations are evaluated in terms of one- and two- electron integrals over the molecular orbitals using the equivalent of the Slater-Condon rules. For large CI calculations, the full $H_{I,J}$ matrix is not actually evaluated and stored in the computer's memory (or on its disk); rather, so-called 'direct CI' methods (see the article by Roos and Siegbahn in MTC) are used to compute and immediately sum contributions to the sum $\sum_J H_{I,J} C_J$ in terms of integrals, density matrix elements, and approximate values of the C_J amplitudes. Iterative methods (see, for example, E. R. Davidson, J. Comput. Phys. 17, 87 (1975)), in which approximate values for the C_J coefficients and energy E are refined through sequential application of $\sum_J H_{I,J}$ to the preceding estimate of the C_J vector, are employed to solve these large CI matrix eigenvalue problems.

C. The MPPT/MBPT Method

In the MPPT/MBPT method, once the reference CSF is chosen and the SCF orbitals belonging to this CSF are determined, the wavefunction Ψ and energy E are determined in an order-by-order manner. This is one of the primary strengths of the MPPT/MBPT technique; it does not require one to make further (potentially arbitrary) choices once the basis set and dominant (SCF) configuration are specified. In contrast to the MCSCF and CI treatments, one need not make choices of CSFs to include in or exclude from Ψ. The MPPT/MBPT perturbation equations determine what CSFs must be included through any particular order.

For example, the first-order wavefunction correction Ψ^1 (i.e., $\Psi = \Phi + \Psi^1$ through first order) is given by:

$$\Psi^1 = -\sum_{i<j,m<n} <\Phi_{i,j}^{m,n} | H - H^0 | \Phi > [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]^{-1} | \Phi_{i,j}^{m,n} >$$

$$= -\sum_{i<j,m<n} [< i,j |g| m,n >- < i,j |g| n,m >][\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]^{-1} | \Phi_{i,j}^{m,n} >$$

where the SCF orbital energies are denoted ϵ_k and $\Phi_{i,j}^{m,n}$ represents a CSF that is <u>doubly excited</u> relative to Φ. Thus, only doubly excited CSFs contribute to the <u>first-order wavefunction</u>; as a result, the energy E is given through second order as:

$$E = < \Phi | H^0 | \Phi > + < \Phi | H - H^0 | \Phi > + < \Phi | H - H^0 | \Psi^1 >$$

$$= < \Phi | H | \Phi > - \sum_{i<j,m<n} |< \Phi_{i,j}^{m,n} | H - H^0 | \Phi >|^2 / [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$$

$$= E_{SCF} - \sum_{i<j,m<n} |<i,j|g|m,n> - <i,j|g|n,m>|^2/[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$$

$$= E^0 + E^1 + E^2.$$

These contributions have been expressed, using the SC rules, in terms of the two-electron integrals $<i,j|g|m,n>$ coupling the excited spin-orbitals to the spin-orbitals from which electrons were excited as well as the orbital energy differences $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$ accompanying such excitations. In this form, it becomes clear that major contributions to the correlation energy of the pair of occupied orbitals ϕ_i ϕ_j are made by double excitations into virtual orbitals ϕ_m ϕ_n that have large coupling (i..e., large $<i,j|g|m,n>$ integrals) and small orbital energy gaps, $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$.

In higher order corrections to the wavefunction and to the energy, contributions from CSFs that are singly, triply, etc. excited relative to Φ appear, and additional contributions from the doubly excited CSFs also enter. It is relatively common to carry MPPT/MBPT calculations (see the references given above in Chapter 19.I.3 where the contributions of the Pople and Bartlett groups to the development of MPPT/MBPT are documented) through to third order in the energy (whose evaluation can be shown to require only Ψ^0 and Ψ^1). The entire GAUSSIAN-8X series of programs, which have been used in thousands of important chemical studies, calculate E through third order in this manner.

In addition to being size-extensive and not requiring one to specify input beyond the basis set and the dominant CSF, the MPPT/MBPT approach is able to include the effect of <u>all</u> CSFs (that contribute to any given order) without having to find any eigenvalues of a matrix. This is an important advantage because matrix eigenvalue determination, which is necessary in MCSCF and CI calculations, requires computer time in proportion to the third power of the dimension of the $H_{I,J}$ matrix. Despite all of these advantages, it is important to remember the primary disadvantages of the MPPT/MBPT approach; its energy is not an upper bound to the true energy and it may not be able to treat cases for which two or more CSFs have equal or nearly equal amplitudes because it obtains the amplitudes of all but the dominant CSF from perturbation theory formulas that assume the perturbation is 'small'.

D. The Coupled-Cluster Method

The implementation of the CC method begins much as in the MPPT/MBPT case; one selects a reference CSF that is used in the SCF process to generate a set of spin-orbitals to be used in the subsequent correlated calculation. The set of working equations of the CC

technique given above in Chapter 19.I.4 can be written explicitly by introducing the form of the so-called cluster operator T,

$$T = \sum_{i,m} t_i^m \, m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} \, m^+ n^+ j i + ...,$$

where the combination of operators $m^+ i$ denotes <u>creation</u> of an electron in virtual spin-orbital ϕ_m and <u>removal</u> of an electron from occupied spin-orbital ϕ_i to generate a single excitation. The operation $m^+ n^+ j i$ therefore represents a double excitation from $\phi_i \phi_j$ to $\phi_m \phi_n$. Expressing the cluster operator T in terms of the amplitudes t_i^m, $t_{i,j}^{m,n}$, etc. for singly, doubly, etc. excited CSFs, and expanding the exponential operators in exp(-T) H exp(T) one obtains:

$$< \Phi_i^m | H + [H,T] + 1/2 \, [[H,T],T] + 1/6 \, [[[H,T],T],T]$$
$$+ 1/24 \, [[[[H,T],T],T],T] | \Phi > = 0;$$

$$< \Phi_{i,j}^{m,n} | H + [H,T] + 1/2 \, [[H,T],T] + 1/6 \, [[[H,T],T],T]$$
$$+ 1/24 \, [[[[H,T],T],T],T] | \Phi > = 0;$$

$$< \Phi_{i,j,k}^{m,n,p} | H + [H,T] + 1/2 \, [[H,T],T] + 1/6 \, [[[H,T],T],T]$$
$$+ 1/24 \, [[[[H,T],T],T],T] | \Phi > = 0,$$

and so on for higher order excited CSFs. It can be shown, because of the one- and two-electron operator nature of H, that the expansion of the exponential operators truncates exactly at the fourth power; that is terms such as [[[[[H,T],T],T],T],T] and higher commutators vanish identically (this is demonstrated in Chapter 4 of <u>Second Quantization Based Methods in Quantum Chemistry</u>, P. Jørgensen and J. Simons, Academic Press, New York (1981).

As a result, the exact CC equations are <u>quartic equations</u> for the t_i^m, $t_{i,j}^{m,n}$, etc. amplitudes. Although it is a rather formidable task to evaluate all of the commutator matrix elements appearing in the above CC equations, it can be and has been done (the references given above to Purvis and Bartlett are especially relevant in this context). The result is to express each such matrix element, via the Slater-Condon rules, in terms of one- and two-electron integrals over the spin-orbitals used in determining Φ, including those in Φ itself and the 'virtual' orbitals not in Φ.

In general, these quartic equations must then be solved in an iterative manner and are susceptible to convergence difficulties that are similar to those that arise in MCSCF-type

calculations. In any such iterative process, it is important to start with an approximation (to the t amplitudes, in this case) which is reasonably close to the final converged result. Such an approximation is often achieved, for example, by neglecting all of the terms that are non-linear in the t amplitudes (because these amplitudes are assumed to be less than unity in magnitude). This leads, for the CC working equations obtained by projecting onto the doubly excited CSFs, to:

$$< i,j \mid g \mid m,n >' + [\ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \, t_{i,j}{}^{m,n} +$$

$$\sum_{i',j',m',n'} < \Phi_{i,j}{}^{m,n} \mid H - H^0 \mid \Phi_{i',j'}{}^{m',n'} > t_{i',j'}{}^{m',n'} = 0 \, ,$$

where the notation $< i,j \mid g \mid m,n >'$ is used to denote the two-electron integral difference $< i,j \mid g \mid m,n > - < i,j \mid g \mid n,m >$. If, in addition, the factors that couple different doubly excited CSFs are ignored (i.e., the sum over i',j',m',n') , the equations for the t amplitudes reduce to the equations for the CSF amplitudes of the first-order MPPT/MBPT wavefunction:

$$t_{i,j}{}^{m,n} = - < i,j \mid g \mid m,n >' / [\ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \, .$$

As Bartlett and Pople have both demonstrated, there is, in fact, close relationship between the MPPT/MBPT and CC methods when the CC equations are solved iteratively starting with such an MPPT/MBPT-like initial 'guess' for these double-excitation amplitudes.

The CC method, as presented here, suffers from the same drawbacks as the MPPT/MBPT approach; its energy is not an upper bound and it may not be able to accurately describe wavefunctions which have two or more CSFs with approximately equal amplitude. Moreover, solution of the non-linear CC equations may be difficult and slowly (if at all) convergent. It has the same advantages as the MPPT/MBPT method; its energy is size-extensive, it requires no large matrix eigenvalue solution, and its energy and wavefunction are determined once one specifies the basis and the dominant CSF.

E. Density Functional Methods

These approaches provide alternatives to the conventional tools of quantum chemistry. The CI, MCSCF, MPPT/MBPT, and CC methods move beyond the single-configuration picture by adding to the wave function more configurations whose

amplitudes they each determine in their own way. This can lead to a very large number of CSFs in the correlated wave function, and, as a result, a need for extraordinary computer resources.

The density functional approaches are different. Here one solves a set of orbital-level equations

$$[- \hbar^2/2m_e \nabla^2 - \sum_A Z_A e^2/|r-R_A| + \int \rho(r') e^2/|r-r'| dr'$$

$$+ U(r)] \phi_i = \epsilon_i \phi_i$$

in which the orbitals $\{\phi_i\}$ 'feel' potentials due to the nuclear centers (having charges Z_A), Coulombic interaction with the total electron density $\rho(r')$, and a so-called exchange-correlation potential denoted $U(r')$. The particular electronic state for which the calculation is being performed is specified by forming a corresponding density $\rho(r')$. Before going further in describing how DFT calculations are carried out, let us examine the origins underlying this theory.

The so-called Hohenberg-Kohn theorem states that the ground-state electron density $\rho(r)$ describing an N-electron system uniquely determines the potential $V(r)$ in the Hamiltonian

$$H = \sum_j \{-\hbar^2/2m_e \nabla_j^2 + V(r_j) + e^2/2 \sum_{k \neq j} 1/r_{j,k} \},$$

and, because H determines the ground-state energy and wave function of the system, the ground-state density $\rho(\mathbf{r})$ determines the ground-state properties of the system. The proof of this theorem proceeds as follows:

a. $\rho(\mathbf{r})$ determines N because $\int \rho(\mathbf{r}) \, d^3r = N$.

b. Assume that there are two distinct potentials (aside from an additive constant that simply shifts the zero of total energy) $V(\mathbf{r})$ and $V'(\mathbf{r})$ which, when used in H and H', respectively, to solve for a ground state produce E_0, $\Psi(\mathbf{r})$ and E_0', $\Psi'(\mathbf{r})$ that have the same one-electron density: $\int |\Psi|^2 \, dr_2 \, dr_3 \ldots dr_N = \rho(\mathbf{r}) = \int |\Psi'|^2 \, dr_2 \, dr_3 \ldots dr_N$.

c. If we think of Ψ' as trial variational wave function for the Hamiltonian H, we know that

$$E_0 < \langle \Psi'|H|\Psi'\rangle = \langle \Psi'|H'|\Psi'\rangle + \int \rho(\mathbf{r}) \, [V(\mathbf{r}) - V'(\mathbf{r})] \, d^3r = E_0' + \int \rho(\mathbf{r}) \, [V(\mathbf{r}) - V'(\mathbf{r})] \, d^3r.$$

d. Similarly, taking Ψ as a trial function for the H' Hamiltonian, one finds that

$$E_0' < E_0 + \int \rho(\mathbf{r}) \, [V'(\mathbf{r}) - V(\mathbf{r})] \, d^3r.$$

e. Adding the equations in c and d gives

$$E_0 + E_0' < E_0 + E_0',$$

a clear contradiction.

Hence, there cannot be two distinct potentials V and V' that give the same ground-state $\rho(\mathbf{r})$. So, the ground-state density $\rho(\mathbf{r})$ uniquely determines N and V, and thus H, and therefore Ψ and E_0. Furthermore, because Ψ determines all properties of the ground state, then $\rho(\mathbf{r})$, in principle, determines all such properties. This means that even the kinetic energy and the electron-electron interaction energy of the ground-state are determined by

ρ(**r**). It is easy to see that $\int \rho$(**r**) V(**r**) d^{3}r = V[ρ] gives the average value of the electron-nuclear (plus any additional one-electron additive potential) interaction in terms of the ground-state density ρ(**r**), but how are the kinetic energy T[ρ] and the electron-electron interaction V_{ee}[ρ] energy expressed in terms of ρ?

The main difficulty with DFT is that the Hohenberg-Kohn theorem shows that the ground-state values of T, V_{ee}, V, etc. are all unique functionals of the ground-state ρ (i.e., that they can, in principle, be determined once ρ is given), but it does not tell us what these functional relations are.

To see how it might make sense that a property such as the kinetic energy, whose operator $-\hbar^2/2m_e \nabla^2$ involves derivatives, can be related to the electron density, consider a simple system of N non-interacting electrons moving in a three-dimensional cubic "box" potential. The energy states of such electrons are known to be

$$E = (h^2/2m_e L^2) (n_x^2 + n_y^2 + n_z^2),$$

where L is the length of the box along the three axes, and n_x, n_y, and n_z are the quantum numbers describing the state. We can view $n_x^2 + n_y^2 + n_z^2 = R^2$ as defining the squared radius of a sphere in three dimensions, and we realize that the density of quantum states in this space is one state per unit volume in the n_x, n_y, n_z space. Because n_x, n_y, and n_z must be positive integers, the volume covering all states with energy less than or equal to a specified energy $E = (h^2/2m_e L^2) R^2$ is 1/8 the volume of the sphere of radius R:

$$\Phi(E) = 1/8 (4\pi/3) R^3 = (\pi/6) (8m_e L^2 E/h^2)^{3/2}.$$

Since there is one state per unit of such volume, $\Phi(E)$ is also the number of states with energy less than or equal to E, and is called the <u>integrated density of states</u>. The number of states $g(E)\,dE$ with energy between E and E+dE, the <u>density of states</u>, is the derivative of Φ:

$$g(E) = d\Phi/dE = (\pi/4)\,(8m_e L^2/h^2)^{3/2}\,E^{1/2}\ .$$

If we calculate the total energy for N electrons, with the states having energies up to the so-called <u>Fermi energy</u> (i.e., the energy of the highest occupied molecular orbital HOMO) doubly occupied, we obtain the ground-state energy:

$$E_0 = 2\int_0^{E_F} g(E)E\,dE = (8\pi/5)\,(2m_e/h^2)^{3/2}\,L^3\,E_F^{5/2}.$$

The total number of electrons N can be expressed as

$$N = 2\int_0^{E_F} g(E)\,dE = (8\pi/3)\,(2m_e/h^2)^{3/2}\,L^3\,E_F^{3/2},$$

which can be solved for E_F in terms of N to then express E_0 in terms of N instead of E_F:

$$E_0 = (3h^2/10m_e)\,(3/8\pi)^{2/3}\,L^3\,(N/L^3)^{5/3}\ .$$

This gives the total energy, which is also the kinetic energy in this case because the potential energy is zero within the "box", in terms of the electron density $\rho(x,y,z) = (N/L^3)$. It therefore may be plausible to express kinetic energies in terms of electron

densities $\rho(\mathbf{r})$, but it is by no means clear how to do so for "real" atoms and molecules with electron-nuclear and electron-electron interactions operative.

In one of the earliest DFT models, the Thomas-Fermi theory, the kinetic energy of an atom or molecule is approximated using the above kind of treatment on a "local" level. That is, for each volume element in $\mathbf{r}$ space, one assumes the expression given above to be valid, and then one integrates over all $\mathbf{r}$ to compute the total kinetic energy:

$$T_{TF}[\rho] = (3h^2/10m_e)(3/8\pi)^{2/3} \int [\rho(\mathbf{r})]^{5/3} d^3r = C_F \int [\rho(\mathbf{r})]^{5/3} d^3r ,$$

where the last equality simply defines the C_F constant (which is 2.8712 in atomic units). Ignoring the correlation and exchange contributions to the total energy, this T is combined with the electron-nuclear V and Coulombic electron-electron potential energies to give the Thomas-Fermi total energy:

$$E_{0,TF}[\rho] = C_F \int [\rho(\mathbf{r})]^{5/3} d^3r + \int V(\mathbf{r})\rho(\mathbf{r}) d^3r + e^2/2 \int \rho(\mathbf{r})\rho(\mathbf{r'})/|\mathbf{r}\text{-}\mathbf{r'}| \, d^3r \, d^3r',$$

This expression is an example of how E_0 is given as a local density functional approximation (LDA). The term local means that the energy is given as a functional (i.e., a function of ρ) which depends only on $\rho(\mathbf{r})$ at points in space but not on $\rho(\mathbf{r})$ at more than one point in space.

Unfortunately, the Thomas-Fermi energy functional does not produce results that are of sufficiently high accuracy to be of great use in chemistry. What is missing in this theory are a. the exchange energy and b. the correlation energy; moreover, the kinetic energy is treated only in the approximate manner described.

In the book by Parr and Yang, it is shown how Dirac was able to address the exchange energy for the 'uniform electron gas' (N Coulomb <u>interacting</u> electrons moving in a uniform positive background charge whose magnitude balances the charge of the N electrons). If the exact expression for the exchange energy of the uniform electron gas is applied on a local level, one obtains the commonly used Dirac <u>local density approximation to the exchange energy</u>:

$$E_{ex,Dirac}[\ ] = -C_x \int [\ (\mathbf{r})]^{4/3} \, d^3r,$$

with $C_x = (3/4)(3/\ )^{1/3} = 0.7386$ in atomic units. Adding this exchange energy to the Thomas-Fermi total energy $E_{0,TF}[\ ]$ gives the so-called Thomas-Fermi-Dirac (TFD) energy functional.

Because electron densities vary rather strongly spatially near the nuclei, corrections to the above approximations to $T[\ ]$ and $E_{ex.Dirac}$ are needed. One of the more commonly used so-called <u>gradient-corrected</u> approximations is that invented by Becke, and referred to as the Becke88 exchange functional:

$$E_{ex}(Becke88) = E_{ex,Dirac}[\ ] - \int x^2 \ ^{4/3} (1+6 \ x \sinh^{-1}(x))^{-1} \, \mathbf{dr},$$

where $x = \ ^{-4/3}|\ |$, and is a parameter chosen so that the above exchange energy can best reproduce the known exchange energies of specific electronic states of the inert gas atoms (Becke finds to equal 0.0042). A common gradient correction to the earlier $T[\ ]$ is called the Weizsacker correction and is given by

$$T_{Weizsacker} = (1/72)(\hbar/m_e) \int |\nabla\rho(\mathbf{r})|^2/\rho(\mathbf{r}) \, d\mathbf{r}.$$

Although the above discussion suggests how one might compute the ground-state energy once the ground-state density $\rho(\mathbf{r})$ is given, one still needs to know how to obtain ρ. Kohn and Sham (KS) introduced a set of so-called KS orbitals obeying the following equation:

$$\{-1/2 \nabla^2 + V(\mathbf{r}) + e^2/2 \int \rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| \, d\mathbf{r}' + U_{xc}(\mathbf{r})\} \phi_j = \varepsilon_j \phi_j,$$

where the so-called exchange-correlation potential $U_{xc}(\mathbf{r}) = \delta E_{xc}[\rho]/\delta\rho(\mathbf{r})$ could be obtained by functional differentiation if the exchange-correlation energy functional $E_{xc}[\rho]$ were known. KS also showed that the KS orbitals $\{\phi_j\}$ could be used to compute the density by simply adding up the orbital densities multiplied by orbital occupancies n_j:

$$\rho(\mathbf{r}) = \sum_j n_j |\phi_j(\mathbf{r})|^2.$$

(here $n_j = 0,1$, or 2 is the occupation number of the orbital ϕ_j in the state being studied) and that the kinetic energy should be calculated as

$$T = \sum_j n_j \langle \phi_j(\mathbf{r})|-1/2 \nabla^2 |\phi_j(\mathbf{r})\rangle.$$

The same investigations of the idealized 'uniform electron gas' that identified the Dirac exchange functional, found that the correlation energy (per electron) could also be written exactly as a <u>function</u> of the electron density ρ of the system, but only in two

limiting cases- the high-density limit (large ρ) and the low-density limit. There still exists

no exact expression for the correlation energy even for the uniform electron gas that is valid

at arbitrary values of ρ. Therefore, much work has been devoted to creating efficient and

accurate interpolation formulas connecting the low- and high- density uniform electron gas

expressions. One such expression is

$$E_C[\rho] = \int \rho(\mathbf{r}) \; \varepsilon_c(\rho) \; d\mathbf{r},$$

where

$$\varepsilon_c(\rho) = A/2\{\ln(x/X) + 2b/Q \; \tan^{-1}(Q/(2x+b)) - bx_0/X_0 \; [\ln((x-x_0)^2/X)$$

$$+2(b+2x_0)/Q \; \tan^{-1}(Q/(2x+b))]$$

is the correlation energy per electron. Here $x = r_s^{1/2}$, $X = x^2 + bx + c$, $X_0 = x_0^2 + bx_0 + c$ and

$Q = (4c - b^2)^{1/2}$, $A = 0.0621814$, $x_0 = -0.409286$, $b = 13.0720$, and $c = 42.7198$. The

parameter r_s is how the density ρ enters since $4/3 \; \pi r_s^3$ is equal to $1/\rho$; that is, r_s is the radius

of a sphere whose volume is the effective volume occupied by one electron. A reasonable

approximation to the full $E_{xc}[\rho]$ would contain the Dirac (and perhaps gradient corrected)

exchange functional plus the above $E_C[\rho]$, but there are many alternative approximations to

the exchange-correlation energy functional. Currently, many workers are doing their best to

"cook up" functionals for the correlation and exchange energies, but no one has yet

invented functionals that are so reliable that most workers agree to use them.

To summarize, in implementing any DFT, one usually proceeds as follows:

1. An atomic orbital basis is chosen in terms of which the KS orbitals are to be expanded.

2. Some initial guess is made for the LCAO-KS expansion coefficients $C_{j,a}$: $\phi_j = \sum_a C_{j,a} \chi_a$.

3. The density is computed as $\rho(\mathbf{r}) = \sum_j n_j |\phi_j(\mathbf{r})|^2$. Often, $\rho(\mathbf{r})$ is expanded in an atomic orbital basis, which need not be the same as the basis used for the ϕ_j, and the expansion coefficients of ρ are computed in terms of those of the ϕ_j. It is also common to use an atomic orbital basis to expand $\rho^{1/3}(\mathbf{r})$ which, together with ρ, is needed to evaluate the exchange-correlation functional's contribution to E_0.

4. The current iteration's density is used in the KS equations to determine the Hamiltonian $\{-1/2 \nabla^2 + V(\mathbf{r}) + e^2/2 \int \rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| \, d\mathbf{r}' + U_{xc}(\mathbf{r}) \}$ whose "new" eigenfunctions $\{\phi_j\}$ and eigenvalues $\{\varepsilon_j\}$ are found by solving the KS equations.

5. These new ϕ_j are used to compute a new density, which, in turn, is used to solve a new set of KS equations. This process is continued until convergence is reached (i.e., until the ϕ_j used to determine the current iteration's ρ are the same ϕ_j that arise as solutions on the next iteration.

6. Once the converged $\rho(\mathbf{r})$ is determined, the energy can be computed using the earlier expression

$$E[\rho] = \sum_j n_j \langle \phi_j(\mathbf{r})|-1/2 \nabla^2 |\phi_j(\mathbf{r})\rangle + \int V(\mathbf{r}) \rho(\mathbf{r}) \, d\mathbf{r} + e^2/2 \int \rho(\mathbf{r}) \rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| \, d\mathbf{r} \, d\mathbf{r}' + E_{xc}[\rho].$$

In closing this section, it should once again be emphasized that this area is currently undergoing explosive growth and much scrutiny. As a result, it is nearly certain that many of the specific functionals discussed above will be replaced in the near future by improved and more rigorously justified versions. It is also likely that extensions of DFT to excited states (many workers are actively pursuing this) will be placed on more solid ground and made applicable to molecular systems. Because the computational effort involved in these approaches scales much less strongly with basis set size than for conventional (SCF, MCSCF, CI, etc.) methods, density functional methods offer great promise and are likely to contribute much to quantum chemistry in the next decade.

Chapter 20

Many physical properties of a molecule can be calculated as expectation values of a corresponding quantum mechanical operator. The evaluation of other properties can be formulated in terms of the "response" (i.e., derivative) of the electronic energy with respect to the application of an external field perturbation.

I. Calculations of Properties Other Than the Energy

There are, of course, properties other than the energy that are of interest to the practicing chemist. Dipole moments, polarizabilities, transition probabilities among states, and vibrational frequencies all come to mind. Other properties that are of importance involve operators whose quantum numbers or symmetry indices label the state of interest. Angular momentum and point group symmetries are examples of the latter properties; for these quantities the properties are precisely specified once the quantum number or symmetry label is given (e.g., for a ^3P state, the average value of L^2 is $<^3P|L^2|^3P> = \hbar^2 1(1+1) = 2\hbar^2$).

Although it may be straightforward to specify what property is to be evaluated, often computational difficulties arise in carrying out the calculation. For some <u>ab initio</u> methods, these difficulties are less severe than for others. For example, to compute the electric dipole transition matrix element $< \Psi_2 | \mathbf{r} | \Psi_1>$ between two states Ψ_1 and Ψ_2, one must evaluate the integral involving the one-electron dipole operator $\mathbf{r} = \sum_j e \mathbf{r}_j - \sum_a e Z_a \mathbf{R}_a$; here the first sum runs over the N electrons and the second sum runs over the nuclei whose charges are denoted Z_a. To evaluate such transition matrix elements in terms of the Slater-Condon rules is relatively straightforward as long as Ψ_1 and Ψ_2 are expressed in terms of Slater determinants involving a single set of orthonormal spin-orbitals. If Ψ_1 and Ψ_2, have been obtained, for example, by carrying out separate MCSCF calculations on the two states in question, the energy optimized spin-orbitals for one state will not be the same as the optimal spin-orbitals for the second state. As a result, the determinants in Ψ_1 and those in Ψ_2 will involve spin-orbitals that are not orthonormal to one another. Thus, the SC rules can not immediately be applied. Instead, a transformation of the spin-orbitals of Ψ_1 and Ψ_2 to a single set of orthonormal functions must be carried out. This then expresses Ψ_1 and Ψ_2 in terms of new Slater determinants over this new set of orthonormal spin-orbitals, after which the SC rules can be exploited.

In contrast, if Ψ_1 and Ψ_2 are obtained by carrying out a CI calculation using a single set of orthonormal spin-orbitals (e.g., with Ψ_1 and Ψ_2 formed from two different

eigenvectors of the resulting secular matrix), the SC rules can immediately be used to evaluate the transition dipole integral.

A. Formulation of Property Calculations as Responses

Essentially all experimentally measured properties can be thought of as arising through the <u>response</u> of the system to some externally applied perturbation or disturbance. In turn, the calculation of such properties can be formulated in terms of the response of the energy E or wavefunction to a perturbation. For example, molecular dipole moments μ are measured, via electric-field deflection, in terms of the change in energy

$$E = \mu \cdot \mathbf{E} + 1/2 \, \mathbf{E} \cdot \quad \cdot \mathbf{E} + 1/6 \, \mathbf{E} \cdot \mathbf{E} \cdot \mathbf{E} \cdot \quad + \ldots$$

caused by the application of an external electric field **E** which is spatially inhomogeneous, and thus exerts a force

$$\mathbf{F} = - \quad E$$

on the molecule proportional to the dipole moment (good treatments of response properties for a wide variety of wavefunction types (i.e., SCF, MCSCF, MPPT/MBPT, etc.) are given in <u>Second Quantization Based Methods in Quantum Chemistry</u> , P. Jørgensen and J. Simons, Academic Press, New York (1981) and in <u>Geometrical Derivatives of Energy Surfaces and Molecular Properties</u>, P. Jørgensen and J. Simons, Eds., NATO ASI Series, Vol. 166, D. Reidel, Dordrecht (1985)).

To obtain expressions that permit properties other than the energy to be evaluated in terms of the state wavefunction , the following strategy is used:

1. The perturbation $V = H - H^0$ appropriate to the particular property is identified. For dipole moments (μ), polarizabilities ( ), and hyperpolarizabilities ( ), V is the interaction of the nuclei and electrons with the external electric field

$$V = \quad_a Z_a e \, \mathbf{R}_a \cdot \mathbf{E} - \quad_j e \, \mathbf{r}_j \cdot \mathbf{E}.$$

For vibrational frequencies, one needs the derivatives of the energy E with respect to deformation of the bond lengths and angles of the molecule, so V is the sum of all changes in the electronic Hamiltonian that arise from displacements $\mathbf{R}_a$ of the atomic centers

$$V = \sum_a (\nabla_{\mathbf{R}_a} H) \cdot \nabla_{\mathbf{R}_a} .$$

2. A power series expansion of the state energy E, computed in a manner consistent with how Ψ is determined (i.e., as an expectation value for SCF, MCSCF, and CI wavefunctions or as $<\Psi|H|\Phi>$ for MPPT/MBPT or as $<\Phi|\exp(-T)H\exp(T)|\Phi>$ for CC wavefunctions), is carried out in powers of the perturbation V:

$$E = E^0 + E^{(1)} + E^{(2)} + E^{(3)} + \ldots$$

In evaluating the terms in this expansion, the dependence of $H = H^0+V$ <u>and</u> of Ψ (which is expressed as a solution of the SCF, MCSCF, ..., or CC equations for H <u>not</u> for H^0) must be included.

3. The desired physical property must be extracted from the power series expansion of E in powers of V.

B. The MCSCF Response Case

1. The Dipole Moment

  To illustrate how the above developments are carried out and to demonstrate how the results express the desired quantities in terms of the original wavefunction, let us consider, for an MCSCF wavefunction, the response to an external electric field. In this case, the Hamiltonian is given as the conventional one- and two-electron operators H^0 to which the above one-electron electric dipole perturbation V is added. The MCSCF wavefunction Ψ and energy E are assumed to have been obtained via the MCSCF procedure with $H=H^0+\lambda V$, where λ can be thought of as a measure of the strength of the applied electric field.
  The terms in the expansion of E(λ) in powers of λ:

$$E = E(\lambda=0) + \lambda (dE/d\lambda)_0 + 1/2 \lambda^2 (d^2E/d\lambda^2)_0 + \ldots$$

are obtained by writing the total derivatives of the MCSCF energy functional with respect to λ and evaluating these derivatives at $\lambda=0$

(which is indicated by the subscript $(..)_0$ on the above derivatives):

$$E(\lambda=0) = <\Psi(\lambda=0)|H^0|\Psi(\lambda=0)> = E^0,$$

$$(dE/d\lambda)_0 = <\Psi(\lambda=0)|V|\Psi(\lambda=0)> + 2\sum_J(\partial C_J/\partial\lambda)_0 <\Psi/\partial C_J|H^0|\Psi(\lambda=0)>$$

$$+ 2\sum_{i,a}(\partial C_{a,i}/\partial\lambda)_0 <\partial\Psi/\partial C_{a,i}|H^0|\Psi(\lambda=0)>$$

$$+ 2\sum_\chi(\partial\chi/\partial\lambda)_0 <\partial\Psi/\partial\chi|H^0|\Psi(\lambda=0)>,$$

and so on for higher order terms. The factors of 2 in the last three terms come through using the hermiticity of H^0 to combine terms in which derivatives of Ψ occur.

The first-order correction can be thought of as arising from the response of the wavefunction (as contained in its LCAO-MO and CI amplitudes and basis functions χ) plus the response of the Hamiltonian to the external field. Because the MCSCF energy functional has been made stationary with respect to variations in the C_J and $C_{i,a}$ amplitudes, the second and third terms above vanish:

$$\partial E/\partial C_J = 2<\partial\Psi/\partial C_J|H^0|\Psi(\lambda=0)> = 0,$$

$$\partial E/\partial C_{a,i} = 2<\partial\Psi/\partial C_{a,i}|H^0|\Psi(\lambda=0)> = 0.$$

If, as is common, the atomic orbital bases used to carry out the MCSCF energy optimization are not explicitly dependent on the external field, the third term also vanishes because $(\partial\chi/\partial\lambda)_0 = 0$. Thus for the MCSCF case, the first-order response is given as the average value of the perturbation over the wavefunction with $\lambda=0$:

$$(dE/d\lambda)_0 = <\Psi(\lambda=0)|V|\Psi(\lambda=0)>.$$

For the external electric field case at hand, this result says that the field-dependence of the state energy will have a linear term equal to

$$<\Psi(\lambda=0)|V|\Psi(\lambda=0)> = <\Psi|\sum_a Z_a e \, \mathbf{R_a}\cdot\mathbf{e} - \sum_j e \, \mathbf{r_j}\cdot\mathbf{e}|\Psi>,$$

where $\mathbf{e}$ is a unit vector in the direction of the applied electric field (the magnitude of the field λ having already been removed in the power series expansion). Since the dipole

moment is determined experimentally as the energy's slope with respect to field strength, this means that the dipole moment is given as:

$$\mu = < \Psi | \sum_a Z_a e \, \mathbf{R}_a - \sum_j e \, \mathbf{r}_j | \Psi >.$$

2. The Geometrical Force

These same techniques can be used to determine the response of the energy to displacements $\delta \mathbf{R}_a$ of the atomic centers. In such a case, the perturbation is

$$V = \sum_a \delta \mathbf{R}_a \cdot \nabla_{\mathbf{R}_a} (- \sum_i Z_a e^2 / |\mathbf{r}_i - \mathbf{R}_a|)$$

$$= - \sum_a Z_a e^2 \, \delta \mathbf{R}_a \cdot \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3.$$

Here, the one-electron operator $\sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3$ is referred to as 'the Hellmann-Feynman' force operator; it is the derivative of the Hamiltonian with respect to displacement of center-a in the x, y, or z direction.

The expressions given above for $E(\lambda=0)$ and $(dE/d\lambda)_0$ can once again be used, but with the Hellmann-Feynman form for V. Once again, for the MCSCF wavefunction, the variational optimization of the energy gives

$$< \partial \Psi / \partial C_J | H^0 | \Psi(\lambda=0)> = < \partial \Psi / \partial C_{a,i} | H^0 | \Psi(\lambda=0)> = 0.$$

However, because the atomic basis orbitals are attached to the centers, and because these centers are displaced in forming V, it is no longer true that $(\partial \Psi / \partial \lambda)_0 = 0$; the variation in the wavefunction caused by movement of the basis functions now contributes to the first-order energy response. As a result, one obtains

$$(dE/d\lambda)_0 = - \sum_a Z_a e^2 \, \delta \mathbf{R}_a \cdot < \Psi | \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3 | \Psi >$$

$$+ 2 \sum_a \delta \mathbf{R}_a \cdot (\partial \Psi / \partial \mathbf{R}_a)_0 < \partial \Psi / \partial \lambda | H^0 | \Psi(\lambda=0)>.$$

The first contribution to the force

$$\mathbf{F}_a = - Z_a e^2 < \Psi | \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3 | \Psi >$$

$$+ 2 \sum (\partial_{\mathbf{R}_a} \Psi)_0 < \quad / \quad |H^0| \ (\ =0)>$$

along the x, y, and z directions for center-a involves the expectation value, with respect to the MCSCF wavefunction with =0, of the Hellmann-Feynman force operator. The second contribution gives the forces due to infinitesimal displacements of the basis functions on center-a.

The evaluation of the latter contributions can be carried out by first realizing that

$$\Psi = \sum_J C_J \ | \phi_{J1} \ \phi_{J2} \ \phi_{J3} \cdots \ \phi_{Jn} \cdots \ \phi_{JN}|$$

with

$$\phi_j = \sum_\mu C_{\mu,j} \ \chi_\mu$$

involves the basis orbitals through the LCAO-MO expansion of the ϕ_js. So the derivatives of the basis orbitals contribute as follows:

$$(\partial_{\mathbf{R}_a} \Psi) < \quad / \quad | = \sum_J \sum_j, C_J \ C_{,j} <| \phi_{J1} \ \phi_{J2} \ \phi_{J3} \cdots \ \partial_{\mathbf{R}_a} \ \cdots \ \phi_{JN}|.$$

Each of these factors can be viewed as combinations of CSFs with the same C_J and $C_{,j}$ coefficients as in but with the jth spin-orbital involving basis functions that have been differentiated with respect to displacement of center-a. It turns out that such derivatives of Gaussian basis orbitals can be carried out analytically (giving rise to new Gaussians with one higher and one lower l-quantum number).

When substituted into $(\partial_{\mathbf{R}_a} \Psi)_0 < \quad / \quad |H^0| \ (\ =0)>$, these basis derivative terms yield

$$(\partial_{\mathbf{R}_a} \Psi)_0 < \quad / \quad |H^0| \ (\ =0)>$$

$$= \sum_J \sum_j, C_J \ C_{,j} <| \phi_{J1} \ \phi_{J2} \ \phi_{J3} \cdots \ \partial_{\mathbf{R}_a} \ \cdots \ \phi_{JN}|H^0| \ >,$$

whose evaluation via the Slater-Condon rules is straightforward. It is simply the expectation value of H^0 with respect to (with the same density matrix elements that arise

in the evaluation of 's energy) <u>but</u> with the one- and two-electron integrals over the atomic basis orbitals involving one of these differentiated functions:

$$< \mu |g| > \quad \mathbf{R}_a < \mu |g| >= < \mathbf{R}_a \mu |g| >$$

$$+< \mu \mathbf{R}_a |g| > +< \mu |g| \mathbf{R}_a > +< \mu |g| \mathbf{R}_a >.$$

In summary, the force $\mathbf{F}_a$ felt by the nuclear framework due to a displacement of center-a along the x, y, or z axis is given as

$$\mathbf{F}_a = - Z_a e^2 < |_i (\mathbf{r}_i - \mathbf{R}_a)/|\mathbf{r}_i - \mathbf{R}_a|^3 | > + (\mathbf{R}_a < |H^0| >),$$

where the second term is the energy of but with all atomic integrals replaced by integral derivatives: $< \mu |g| >$
$\mathbf{R}_a < \mu |g| >.$

C. Responses for Other Types of Wavefunctions

It should be stressed that the MCSCF wavefunction yields especially compact expressions for responses of E with respect to an external perturbation because of the variational conditions

$$< / C_J|H^0| (=0)> = < / C_{a,i}|H^0| (=0)> =0$$

that apply. The SCF case, which can be viewed as a special case of the MCSCF situation, also admits these simplifications. However, the CI, CC, and MPPT/MBPT cases involve additional factors that arise because the above variational conditions do not apply (in the CI case, $< / C_J|H^0| (=0)> = 0$ still applies, but the orbital condition $< / C_{a,i}|H^0| (=0)> =0$ does not because the orbitals are not varied to make the CI energy functional stationary).

Within the CC, CI, and MPPT/MBPT methods, one must evaluate the so-called responses of the C_I and $C_{a,i}$ coefficients $(C_J/)_0$ and $(C_{a,i}/)_0$ that appear in the full energy response as (see above)
$2 _J (C_J/)_0 < / C_J|H^0| (=0)> +2 _{i,a}(C_{a,i}/)_0 < / C_{a,i}|H^0| (=0)>.$ To do so requires solving a set of response equations that are obtained by differentiating whatever

equations govern the C_I and $C_{a,i}$ coefficients in the particular method (e.g., CI, CC, or MPPT/MBPT) with respect to the external perturbation. In the geometrical derivative case, this amounts to differentiating with respect to x, y, and z displacements of the atomic centers. These response equations are discussed in <u>Geometrical Derivatives of Energy Surfaces and Molecular Properties</u>, P. Jørgensen and J. Simons, Eds., NATO ASI Series, Vol. 166, D. Reidel, Dordrecht (1985). Their treatment is somewhat beyond the scope of this text, so they will not be dealt with further here.

D. The Use of Geometrical Energy Derivatives

1. Gradients as Newtonian Forces

The first energy derivative is called the gradient **g** and is the negative of the force **F** (with components along the ath center denoted $\mathbf{F_a}$) experienced by the atomic centers $\mathbf{F} = -\mathbf{g}$. These forces, as discussed in Chapter 16, can be used to carry out classical trajectory simulations of molecular collisions or other motions of large organic and biological molecules for which a quantum treatment of the nuclear motion is prohibitive.

The second energy derivatives with respect to the x, y, and z directions of centers a and b (for example, the x, y component for centers a and b is $H_{ax,by} = (\partial^2 E/\partial x_a \partial y_b)_0$) form the Hessian matrix **H**. The elements of **H** give the local curvatures of the energy surface along the 3N cartesian directions.

The gradient and Hessian can be used to systematically locate local minima (i.e., stable geometries) and transition states that connect one local minimum to another. At each of these stationary points, all forces and thus all elements of the gradient **g** vanish. At a local minimum, the **H** matrix has 5 or 6 zero eigenvalues corresponding to translational and rotational displacements of the molecule (5 for linear molecules; 6 for non-linear species) and 3N-5 or 3N-6 positive eigenvalues. At a transition state, **H** has one negative eigenvalue, 5 or 6 zero eigenvalues, and 3N-6 or 3N-7 positive eigenvalues.

2. Transition State Rate Coefficients

The transition state theory of Eyring or its extensions due to Truhlar and co-workers (see, for example, D. G. Truhlar and B. C. Garrett, Ann. Rev. Phys. Chem. <u>35</u>, 159 (1984)) allow knowledge of the Hessian matrix at a transition state to be used to compute a rate coefficient k_{rate} appropriate to the chemical reaction for which the transition state applies.

More specifically, the geometry of the molecule at the transition state is used to compute a rotational partition function $Q^{\dagger}_{rot}$ in which the principal moments of inertia I_a, I_b, and I_c (see Chapter 13) are those of the transition state (the † symbol is, by convention, used to label the transition state):

$$Q^{\dagger}_{rot} = \prod_{n=a,b,c} \sqrt{\frac{8\pi^2 I_n kT}{h^2}} \, ,$$

where k is the Boltzmann constant and T is the temperature in °K.

The eigenvalues {ω} of the mass weighted Hessian matrix (see below) are used to compute, for each of the 3N-7 vibrations with real and positive ω values, a vibrational partition function that is combined to produce a transition-state vibrational partition function:

$$Q^{\dagger}_{vib} = \prod_{\nu=1,3-7} \frac{\exp(-\hbar\omega_{\nu}/2kT)}{1-\exp(-\hbar\omega_{\nu}/kT)} \, .$$

The electronic partition function of the transition state is expressed in terms of the activation energy (the energy of the transition state relative to the electronic energy of the reactants) $E^{\dagger}$ as:

$$Q^{\dagger}_{electronic} = \omega^{\dagger} \exp(-E^{\dagger}/kT)$$

where $\omega^{\dagger}$ is the degeneracy of the electronic state at the transition state geometry.

In the original Eyring version of transition state theory (TST), the rate coefficient k_{rate} is then given by:

$$k_{rate} = \frac{kT}{h} \, \omega^{\dagger} \exp(-E^{\dagger}/kT) \frac{Q^{\dagger}_{rot}Q^{\dagger}_{vib}}{Q_{reactants}} \, ,$$

where $Q_{reactants}$ is the conventional partition function for the reactant materials.

For example, in a bimolecular reaction such as:

$$F + H_2 \rightarrow FH + H,$$

the reactant partition function

$$Q_{reactants} = Q_F \; Q_{H_2}$$

is written in terms of the translational and electronic (the degeneracy of the ^2P state produces the 2 (3) overall degeneracy factor) partition functions of the F atom

$$Q_F = \left(\frac{2 \; m_F kT}{h^2}\right)^{3/2} 2 \; (3)$$

and the translational, electronic, rotational, and vibrational partition functions of the H_2 molecule

$$Q_{H_2} = \left(\frac{2 \; m_{H_2} kT}{h^2}\right)^{3/2} \frac{8 \; ^2 I_{H_2} kT}{2h^2} \frac{\exp(-\hbar \; \omega_{H_2}/2kT)}{1-\exp(-\hbar \; \omega_{H_2}/kT)} \; .$$

The factor of 2 in the denominator of the H_2 molecule's rotational partition function is the "symmetry number" that must be inserted because of the identity of the two H nuclei.

The overall rate coefficient k_{rate} (with units sec^{-1} because this is a rate per collision pair) can thus be expressed entirely in terms of energetic, geometrical, and vibrational information about the reactants and the transition state. Even within the extensions to Eyring's original model, such is the case. The primary difference in the more modern theories is that the transition state is identified not as the point on the potential energy surface at which the gradient vanishes and there is one negative Hessian eigenvalue. Instead, a so-called variational transition state (see the above reference by Truhlar and Garrett) is identified. The geometry, energy, and local vibrational frequencies of this transition state are then used to compute, must like outlined above, k_{rate}.

3. Harmonic Vibrational Frequencies

It is possible (see, for example, J. Nichols, H. L. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. <u>92</u>, 340 (1990) and references therein) to remove from **H** the zero eigenvalues that correspond to rotation and translation and to thereby produce a Hessian matrix whose eigenvalues correspond only to internal motions of the system. After doing so, the number of negative eigenvalues of **H** can be used to characterize the nature of the

stationary point (local minimum or transition state), and **H** can be used to evaluate the local harmonic vibrational frequencies of the system.

The relationship between **H** and vibrational frequencies can be made clear by recalling the classical equations of motion in the Lagrangian formulation:

$$d/dt(\partial L/\partial \dot{q}_j) - (\partial L/\partial q_j) = 0,$$

where q_j denotes, in our case, the 3N cartesian coordinates of the N atoms, and $\dot{q}_j$ is the velocity of the corresponding coordinate. Expressing the Lagrangian L as kinetic energy minus potential energy and writing the potential energy as a local quadratic expansion about a point where **g** vanishes, gives

$$L = 1/2 \sum_j m_j \dot{q}_j^2 - E(0) - 1/2 \sum_{j,k} q_j H_{j,k} q_k .$$

Here, E(0) is the energy at the stationary point, m_j is the mass of the atom to which q_j applies, and the $H_{j,k}$ are the elements of **H** along the x, y, and z directions of the various atomic centers.

Applying the Lagrangian equations to this form for L gives the equations of motion of the q_j coordinates:

$$m_j \ddot{q}_j = - \sum_k H_{j,k} q_k.$$

To find solutions that correspond to local harmonic motion, one assumes that the coordinates q_j oscillate in time according to

$$q_j(t) = q_j \cos(\omega t).$$

Substituting this form for $q_j(t)$ into the equations of motion gives

$$m_j \omega^2 q_j = \sum_k H_{j,k} q_k.$$

Defining

$$q_j' = q_j (m_j)^{1/2}$$

and introducing this into the above equation of motion yields

$$\omega^2\, q_j' = \sum_k H'_{j,k}\, q_k'\ ,$$

where

$$H'_{j,k} = H_{j,k}\, (m_j m_k)^{-1/2}$$

is the so-called <u>mass-weighted Hessian</u> matrix.

     The squares of the desired harmonic vibrational frequencies ω^2 are thus given as eigenvalues of the mass-weighted Hessian $\mathbf{H'}$:

$$\mathbf{H'}\,\mathbf{q'} = \omega^2\,\mathbf{q'}$$

The corresponding eigenvector, $\{q'_{\omega,j}\}$ gives, when multiplied by $m_j^{-1/2}$, the atomic displacements that accompany that particular harmonic vibration. At a transition state, one of the ω^2 will be negative and 3N-6 or 3N-7 will be positive.

4. Reaction Path Following

     The Hessian and gradient can also be used to trace out 'streambeds' connecting local minima to transition states. In doing so, one utilizes a local harmonic description of the potential energy surface

$$E(\mathbf{x}) = E(\mathbf{0}) + \mathbf{x}\cdot\mathbf{g} + 1/2\,\mathbf{x}\cdot\mathbf{H}\cdot\mathbf{x} + \dots,$$

where $\mathbf{x}$ represents the (small) step away from the point $\mathbf{x}=\mathbf{0}$ at which the gradient $\mathbf{g}$ and Hessian $\mathbf{H}$ have been evaluated. By expressing $\mathbf{x}$ and $\mathbf{g}$ in terms of the eigenvectors $\mathbf{v}_\alpha$ of $\mathbf{H}$

$$\mathbf{H}\mathbf{v}_\alpha = \lambda_\alpha\,\mathbf{v}_\alpha\ ,$$

$$\mathbf{x} = \sum_\alpha \langle\mathbf{v}_\alpha|\mathbf{x}\rangle\,\mathbf{v}_\alpha = \sum_\alpha x_\alpha\,\mathbf{v}_\alpha\ ,$$

$$\mathbf{g} = \sum_\alpha \langle\mathbf{v}_\alpha|\mathbf{g}\rangle\,\mathbf{v}_\alpha = \sum_\alpha g_\alpha\,\mathbf{v}_\alpha\ ,$$

the energy change $E(\mathbf{x})$ - $E(\mathbf{0})$ can be expressed in terms of a sum of independent changes along the eigendirections:

$$E(\mathbf{x}) - E(\mathbf{0}) = \sum [\, x\ g\ +1/2\ x^2\ \,] + ...$$

Depending on the signs of g and of , various choices for the displacements x will produce increases or decreases in energy:

1. If is positive, then a step x 'along' g (i.e., one with $x\ g$ positive) will generate an energy increase. A step 'opposed to' g will generate an energy decrease if it is short enough that $x\ g$ is larger in magnitude than $1/2\ x^2$, otherwise the energy will increase.

2. If is negative, a step opposed to g will generate an energy decrease. A step along g will give an energy increase if it is short enough for $x\ g$ to be larger in magnitude than $1/2\ x^2$, otherwise the energy will decrease.

Thus, to proceed downhill in all directions (such as one wants to do when searching for local minima), one chooses each x in opposition to g and of small enough length to guarantee that the magnitude of $x\ g$ exceeds that of $1/2\ x^2$ for those modes with > 0. To proceed uphill along a mode with $' < 0$ and downhill along all other modes with > 0, one chooses x ' along g ' with x ' short enough to guarantee that $x\ '\ g\ '$ is larger in magnitude than $1/2\ x^2\ '$ ', and one chooses the other x opposed to g and short enough that $x\ g$ is larger in magnitude than $1/2\ x^2$.

Such considerations have allowed the development of highly efficient potential energy surface 'walking' algorithms (see, for example, J. Nichols, H. L. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. <u>92</u>, 340 (1990) and references therein) designed to trace out streambeds and to locate and characterize, via the local harmonic frequencies, minima and transition states. These algorithms form essential components of most modern <u>ab initio</u>, semi-empirical, and empirical computational chemistry software packages.

II. <u>Ab Initio</u>, Semi-Empirical and Empirical Force Field Methods

A. *Ab Initio* Methods

      Most of the techniques described in this Chapter are of the <u>ab initio</u> type. This means that they attempt to compute electronic state energies and other physical properties, as functions of the positions of the nuclei, from first principles without the use or knowledge of experimental input. Although perturbation theory or the variational method may be used to generate the working equations of a particular method, and although finite atomic orbital basis sets are nearly always utilized, these approximations do not involve 'fitting' to known experimental data. They represent approximations that can be systematically improved as the level of treatment is enhanced.

B. Semi-Empirical and Fully Empirical Methods

      Semi-empirical methods, such as those outlined in Appendix F, use experimental data or the results of <u>ab initio</u> calculations to determine some of the matrix elements or integrals needed to carry out their procedures. Totally empirical methods attempt to describe the internal electronic energy of a system as a function of geometrical degrees of freedom (e.g., bond lengths and angles) in terms of analytical 'force fields' whose parameters have been determined to 'fit' known experimental data on some class of compounds. Examples of such parameterized force fields were presented in Section III. A of Chapter 16.

C. Strengths and Weaknesses

      Each of these tools has advantages and limitations. <u>Ab initio</u> methods involve intensive computation and therefore tend to be limited, for practical reasons of computer time, to smaller atoms, molecules, radicals, and ions. Their CPU time needs usually vary with basis set size (M) as at least M^4; correlated methods require time proportional to at least M^5 because they involve transformation of the atomic-orbital-based two-electron integrals to the molecular orbital basis. As computers continue to advance in power and memory size, and as theoretical methods and algorithms continue to improve, <u>ab initio</u> techniques will be applied to larger and more complex species. When dealing with systems in which qualitatively new electronic environments and/or new bonding types arise, or excited electronic states that are unusual, <u>ab initio</u> methods are essential. Semi-empirical or empirical methods would be of little use on systems whose electronic properties have not been included in the data base used to construct the parameters of such models.

On the other hand, to determine the stable geometries of large molecules that are made of conventional chemical units (e.g., CC, CH, CO, etc. bonds and steric and torsional interactions among same), fully empirical force-field methods are usually quite reliable and computationally very fast. Stable geometries and the relative energetic stabilities of various conformers of large macromolecules and biopolymers can routinely be predicted using such tools if the system contains only conventional bonding and common chemical building blocks. These empirical potentials usually do not contain sufficient flexibility (i.e., their parameters and input data do not include enough knowledge) to address processes that involve rearrangement of the electronic configurations. For example, they can not treat:

1. Electronic transitions, because knowledge of the optical oscillator strengths and of the energies of excited states is absent in most such methods;

2. Concerted chemical reactions involving simultaneous bond breaking and forming, because to do so would require the force-field parameters to evolve from those of the reactant bonding to those for the product bonding as the reaction proceeds;

3. Molecular properties such as dipole moment and polarizability, although in certain fully empirical models, bond dipoles and lone-pair contributions have been incorporated (although again only for conventional chemical bonding situations).

Semi-empirical techniques share some of the strengths and weaknesses of ab initio and of fully empirical methods. They treat at least the valence electrons explicitly, so they are able to address questions that are inherently electronic such as electronic transitions, dipole moments, polarizability, and bond breaking and forming. Some of the integrals involving the Hamiltonian operator and the atomic basis orbitals are performed ab initio; others are obtained by fitting to experimental data. The computational needs of semi-empirical methods lie between those of the ab initio methods and the force-field techniques. As with the empirical methods, they should never be employed when qualitatively new electronic bonding situations are encountered because the data base upon which their parameters were determined contain, by assumption, no similar bonding cases.