Chapter 2 | Energy and Wavefunction Optimization Methods

A. INTRODUCTION

The total electronic energy of a system described by a state $|0\rangle$ is given as

$$E = \langle 0|H|0\rangle, \qquad \langle 0|0\rangle = 1 \tag{2.1}$$

In approximations commonly used to describe the true state function, $|0\rangle$ may depend on variational parameters $C_1, C_2, \ldots, C_j$, which may be expansion coefficients describing either the linear combination of configurations in $|0\rangle$ or the orbitals [Eq. (1.18)] appearing in these configurations. The total energy forms an energy hypersurface in these parameters $E(C_1, C_2, \ldots, C_j)$. We wish to determine stationary points or extrema of the energy hypersurface that, of course, occur when

$$\partial E(C_1, C_2, \ldots, C_j)/\partial C_i = 0, \qquad i = 1, 2, \ldots, j \tag{2.2}$$

In this chapter, the problem of making $E(C_1, C_2, \ldots, C_j)$ stationary will be treated for both linear and nonlinear parameters that arise in treating the most common quantum-chemical energy expressions. The first derivatives of the total energy determine the slope at a given point of the energy hypersurface, while the second derivatives of the total energy

$$\partial^2 E(C_1, C_2, \ldots, C_j)/\partial C_i \partial C_j \tag{2.3}$$

determine the curvature of the energy hypersurface and thus may be used to characterize the stationary point as a local minimum, a saddle point, or a local maximum. In attempting to find excited states of a given symmetry, one must use care to guarantee that the procedure does not permit a collapse to the lowest state of that symmetry. Procedures such as constraining the class of wavefunctions given by $\{C_i\}$ to be orthogonal to the ground state or monitoring the dominant components (largest C_i) of each wavefunction are commonly used to avoid this difficulty.

In the first optimization procedures we examine below, the parameters C_j characterize a unitary transformation of the wavefunction within the space of both orbital and configuration variations. To determine a stationary point (SP) on the energy hypersurface in this case, we derive an iterative scheme that is quadratically convergent both for ground and excited states. We use knowledge of the first and second derivatives of the total energy to determine the iterative step lengths that we have to take to reach the SP. If the energy hypersurface were parabolic in all of the parameters considered, we would reach the SP in one step. The iterative nature of the solution originates from the nonparabolic terms in the true energy hypersurface, whose description we truncate after quadratic terms.

To be more explicit about the kinds of variational parameters that commonly arise, we write the wavefunction $|0\rangle$ as a linear combination of the orthonormal basis states $\{|\phi_g\rangle\}$ that may originate from several electronic configurations:

$$|0\rangle = \sum_g |\phi_g\rangle C_{g0} \tag{2.4}$$

Each of the states $|\phi_g\rangle$ is formed from a single electronic configuration and is defined as

$$|\phi_g\rangle = \prod_{r \in g} r^+ |\text{vac}\rangle \tag{2.5}$$

where the product $\prod_{r \in g} r^+$ refers to an ordered set of creation operators. The coefficients C_{g0} are the expansion coefficients for the considered state $|0\rangle$ within this configuration basis $\{|\phi_g\rangle\}$. Variations of the spin-orbitals $\{\phi_r\}$ are commonly expressed in terms of variations in the linear expansion coefficients describing the $\{\phi_r\}$ within an atomic orbital basis. [Eq. (1.18)].

In a multiconfigurational self-consistent field (MCSCF) calculation (Dalgaard and Jørgensen, 1978; Schaefer and Miller, 1977, Chapters 3 and 4), we consider both the configuration expansion coefficients and the orbitals as variational parameters. The optimization techniques required to determine an MCSCF wavefunction are discussed in Section B. In a configuration interaction (CI) calculation, the coefficients C_{g0} are determined from Eq. (2.2) under the assumption that the orbitals are fixed. We discuss various approaches to the CI problem in more detail in Section D. The Hartree–Fock (HF) approximation assumes that the reference state refers to a single configuration but the orbitals (or creation operators) are allowed to vary and are determined from Eq. (2.2). Several techniques that have been put forth to generate optimal HF orbitals are considered in more detail in Section C.

B. MULTICONFIGURATIONAL SELF-CONSISTENT FIELD

1. Unitary Transformation of the Wavefunction

Let us now describe how one determines SPs on the energy hypersurface when the wavefunction has the form given in Eq. (2.4). We allow variations to occur in both the orbitals and the configuration expansion coefficients. In Eqs. (1.52) and (1.57) we have described how the orbital variations may be carried out by performing a unitary transformation among the orbitals. The variations in the expansion coefficients may be described in a similar manner (Dalgaard, 1980). The expansion coefficients for the state $|0\rangle$ form one column of a unitary matrix in which the remaining columns are the expansion coefficients for the orthogonal complement states within the configuration space being considered:

$$|n\rangle = \sum_g |\phi_g\rangle C_{gn} \tag{2.6}$$

The states $\{|0\rangle, |n\rangle\}$ and $\{|\phi_g\rangle\}$ thus are related through a unitary transformation matrix C. Variations in the expansion coefficients C_{gn} may be achieved either by a direct variation of these linear parameters or alternatively in terms of parameters S_{lm} describing a unitary transformation among the states $\{|l\rangle\}$. The operator

$$S = \sum_{l,m} S_{lm}|l\rangle\langle m| \tag{2.7}$$

when applied on the set of states $\{|k\rangle\}$ results in a general transformation among the states $\{|k\rangle\}$. The operator $\exp(iS)$ therefore may be used to describe a general unitary transformation among the states $\{|k\rangle\}$.

This unitary transformation shows great resemblance to the unitary transformation $\exp(i\lambda)$ in Eq. (1.36). The operator S is hermitian and the parameters S_{lm} form a hermitian matrix that determines the unitary transformation to be performed. Since we consider only real orbitals here, it becomes sufficient to use only the imaginary part of the variational parameters S_{lm}, denoted iP_{lm} [analogous to using only the $i\kappa_{rs}$ part of λ in Eq. (1.54)], and the S operator then takes the form

$$S = i \sum_{l>m} P_{lm}(|l\rangle\langle m| - |m\rangle\langle l|) \tag{2.8}$$

Further, because our interest is in optimizing the total energy for the state $|0\rangle$, we need only include the $m = 0$ parameter P_{l0} in Eq. (2.8), which then limits the operator S to be of the form

$$S = i \sum_{n\neq 0} P_{n0}(|n\rangle\langle 0| - |0\rangle\langle n|) \tag{2.9}$$

where the elements P_{n0} are real. The matrix P is a real antisymmetric matrix that, in its lower triangle, has all zero elements except for the elements P_{n0}:

$$\mathsf{P} = \begin{pmatrix} 0 & -P_{10} & -P_{20} & \cdots & -P_{n0} \\ P_{10} & 0 & 0 & \cdots & \\ P_{20} & 0 & 0 & & \\ \vdots & \vdots & & & \\ & & & & 0 \end{pmatrix} \tag{2.10}$$

The nonlinear variational parameters P_{n0} are one less in number than the linear expansion coefficients C_{g0}. This is due to the fact that a normalization condition has to be imposed on the linear expansion coefficients $\{C_{g0}\}$ if they are used as variational parameters, whereas variations described by the parameters P_{n0} automatically preserve the orthonormality of the states.

Let us now carry out the above unitary transformation. We obtain by expanding the exponential

$$\exp(iS)|m\rangle = \left[1 + iS + \frac{1}{2!}(iS)^2 + \frac{1}{3!}(iS)^3 + \cdots\right]|m\rangle. \tag{2.11}$$

The second term in the expansion may be written as

$$iS|m\rangle = ii\sum_{n\neq 0} P_{n0}(|n\rangle\langle 0| - |0\rangle\langle n|)|m\rangle = -\sum_{l}|l\rangle P_{lm}. \tag{2.12}$$

The last identity follows by the definition of the (sparse) P matrix in Eq. (2.10). The third term in the expansion in Eq. (2.11) may be determined through successive applications of Eq. (2.12) to be

$$+\tfrac{1}{2}iSiS|m\rangle = -\tfrac{1}{2}iS\sum_{l}|l\rangle P_{lm} = \tfrac{1}{2}\sum_{p,l}|p\rangle P_{pl}P_{lm} \tag{2.13}$$

Successive terms in the expansion of the exponential in Eq. (2.11) are determined in a similar manner, after which it becomes obvious that the terms may be summed to give an exponential matrix

$$\exp(iS)|m\rangle = \sum_{l}|l\rangle[\exp(-\mathsf{P})]_{lm}. \tag{2.14}$$

The actual evaluation of the exponential matrix in terms of the unitary transformation that diagonalizes $i\mathsf{P}$ may be carried out in a manner analogous to that described in Eq. (1.57) for $\exp(i\boldsymbol{\lambda})$.

Because of the especially simple nature of the above P matrix, the unitary transformation in Eq. (2.14) may be carried out analytically. We obtain by collecting together the terms arising in the $(1/n!)(iS)^n|m\rangle$ factors as sine and cosine components:

$$\exp(iS)|0\rangle = \cos x|0\rangle - \frac{1}{x}\sin x\sum_{n} P_{n0}|n\rangle \tag{2.15}$$

$$\exp(iS)|m\rangle = |m\rangle + P_{m0}\frac{1}{x}\sin x|0\rangle + \frac{1}{x^2}(\cos x - 1)P_{m0}\sum_n |n\rangle P_{n0} \quad (2.16)$$

where

$$x^2 \equiv \sum_n P_{n0}^2 \quad (2.17)$$

A unitary transformation of the reference state may now be described as

$$|\tilde{0}\rangle = \exp(i\lambda)\exp(iS)|0\rangle \quad (2.18)$$

Using the technique of Eq. (1.40) to transform all of the creation operators appearing in $|0\rangle$ and in $\exp(iS)$ (i.e., those in $|l\rangle$), we can write

$$|\tilde{0}\rangle = \exp(i\bar{S})|\bar{0}\rangle \quad (2.19)$$

where $\bar{S}$ and $|\bar{0}\rangle$ are defined as in Eqs. (2.9) and (2.4), respectively, with creation operators $\tilde{r}^+$ referring to the transformed set of orbitals. The unitary transformation of the state $|\tilde{0}\rangle$ can thus be thought of as first carrying out a unitary transformation among the orbitals in $|0\rangle$ and S and then performing a unitary transformation in the configuration space [Eq. (2.19)]. This same transformation can be viewed in a somewhat different manner. One may interpret it as first performing the configuration transformation involving all untransformed orbitals (or creation operators)

$$\exp(iS)|0\rangle = \sum_l [\exp(-\mathrm{P})]_{l0}|l\rangle \quad (2.20)$$

as given by Eq. (2.14) and then transforming the orbitals in the functions $|n\rangle$ to give

$$\exp(i\lambda)[\exp(iS)|0\rangle] = \sum_l [\exp(-P)]_{l0}|\bar{l}\rangle \quad (2.21)$$

where

$$|\bar{l}\rangle \equiv \exp(i\lambda)|l\rangle \quad (2.22)$$

Of course, both of these interpretations of Eq. (2.18) amount to nothing more than two ways of working at the same configuration and orbital transformation.

An alternative description of a unitary transformation of the reference state involves using the exponentials in Eq. (2.18) in the opposite order. This form implies that the reference state may be rewritten as

$$|\tilde{0}\rangle = \exp(iS)|\bar{0}\rangle \quad (2.23)$$

where the creation operators in $|\bar{0}\rangle$ refer to the set of transformed orbitals, while the creation operators in S correspond to the nontransformed set.

The subsequent evaluation of $\exp(iS)|\bar{0}\rangle$ would be very difficult because it would involve the computation of overlaps between states $\langle n|\bar{0}\rangle$ involving both transformed and nontransformed orbitals. This would complicate tremendously the determination of the transformed state $|\tilde{0}\rangle$; we therefore consider in the following only the unitary transformation of the reference state given in Eq. (2.18).

2. Variation of the Total Energy

The total energy corresponding to the transformed reference state is given as

$$\begin{aligned} E(\lambda, S) &= \langle 0|\exp(-iS)\exp(-i\lambda)H\exp(i\lambda)\exp(iS)|0\rangle \\ &= \langle 0|H|0\rangle - i\langle 0|[S+\lambda, H]|0\rangle + \tfrac{1}{2}\langle 0|[S,[H,S]]|0\rangle \\ &\quad + \tfrac{1}{2}\langle 0|[\lambda,[H,\lambda]]|0\rangle + \langle 0|[S,[H,\lambda]]|0\rangle + \cdots \end{aligned} \tag{2.24}$$

By introducing a matrix notation in which the variational parameters κ_{rs} and P_{n0} form row and column vectors, we can rewrite Eq. (2.24) as

$$E(\lambda, S) = E(0,0) - 2(\boldsymbol{\kappa}\mathbf{P})\begin{pmatrix}\mathbf{W}\\ \mathbf{V}\end{pmatrix} + (\boldsymbol{\kappa}\mathbf{P})(\mathsf{A}-\mathsf{B})\begin{pmatrix}\boldsymbol{\kappa}\\ \mathbf{P}\end{pmatrix} + \cdots \tag{2.25}$$

We have introduced in Eq. (2.25) the short-hand notation for the operators

$$Q^{+} = \{r^{+}s\}(r > s), \qquad R^{+} = \{|n\rangle\langle 0|\} \tag{2.26}$$

and defined the matrices

$$\mathbf{W} = \langle 0|[Q, H]|0\rangle \tag{2.27}$$

$$\mathbf{V} = \langle 0|[R, H]|0\rangle \tag{2.28}$$

$$\mathsf{A} = \begin{pmatrix}\mathsf{A}_{11} & \mathsf{A}_{12}\\ \mathsf{A}_{21} & \mathsf{A}_{22}\end{pmatrix} = \begin{pmatrix}\langle 0|[Q,H,Q^{+}]|0\rangle\langle 0|[[Q,H],R^{+}]|0\rangle\\ \langle 0|[R,[H,Q^{+}]]|0\rangle\langle 0|[R,H,R^{+}]|0\rangle\end{pmatrix} \tag{2.29}$$

$$\mathsf{B} = \begin{pmatrix}\mathsf{B}_{11} & \mathsf{B}_{12}\\ \mathsf{B}_{21} & \mathsf{B}_{22}\end{pmatrix} = \begin{pmatrix}\langle 0|[Q,H,Q]|0\rangle\langle 0|[[Q,H],R]|0\rangle\\ \langle 0|[R,[H,Q]|0\rangle\langle 0|[R,H,R]|0\rangle\end{pmatrix} \tag{2.30}$$

For convenience, we have introduced the double commutator, defined as

$$[Q,H,Q^{+}] = \tfrac{1}{2}\{[Q,[H,Q^{+}]] + [[Q,H],Q^{+}]\} \tag{2.31}$$

which arises naturally in A_{11}, A_{22}, B_{11}, and B_{22} because

$$\langle 0|[\lambda,[H,\lambda]]|0\rangle - \langle 0|[[\lambda,H],\lambda]|0\rangle = \langle 0|[[\lambda,\lambda],H]|0\rangle = 0 \tag{2.32}$$

and an analogous result for S. The matrices **W**, **V** determine the first-order variations of the energy function, which at a SP on the energy hypersurface are zero. The condition that **V** and **W** are zero at a SP is often referred to as the generalized Brillouin theorem (GBT).

The matrix $\mathsf{A} - \mathsf{B}$ defines the second-order variation of the energy function and is often referred to as the Hessian matrix. The double-commutator form of the Hessian matrix allows these second-order terms to be expressed as a quadratic form.

3. One-Step Second-Order Procedure

As stated previously, a SP on the energy hypersurface is obtained when $\delta E(\lambda, S) = 0$. Neglecting third- and higher-order terms in the energy function [which rigorously no longer makes $E(\lambda, S)$ a true expectation value] we obtain from Eq. (2.25), by differentiating with respect to $\boldsymbol{\kappa}$ and **P**,

$$-\begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix} + (\mathsf{A} - \mathsf{B})\begin{pmatrix} \boldsymbol{\kappa} \\ \mathbf{P} \end{pmatrix} = \mathbf{0} \tag{2.33}$$

or equivalently

$$\begin{pmatrix} \boldsymbol{\kappa} \\ \mathbf{P} \end{pmatrix} = (\mathsf{A} - \mathsf{B})^{-1}\begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix} \tag{2.34}$$

as the conditions for a SP. The matrices $\boldsymbol{\kappa}$ and **P** may then be determined from Eq. (2.34) and a set of transformed orbitals and states obtained from Eqs. (1.52) and (2.14), respectively. If the energy hypersurface contained no higher than quadratic terms, we would reach a SP in one iteration of the above procedure. The third- and higher-order terms in the energy function do, however, require that an iterative scheme be applied to determine a SP. The iterative scheme may be described as follows: From an initial guess of orbitals and a choice of the configuration space, we determine a set of approximate eigenstates $|n\rangle$ (e.g., by performing a configuration interaction calculation). The matrices **V**, **W**, A, and B are then determined and Eq. (2.34) is solved to give the matrices $\boldsymbol{\kappa}$ and **P**. A transformed set of orbitals and states may then be obtained from Eqs. (1.52) and (2.14) and the procedure repeated until the numerical values of **W** and **V** are smaller than a specific tolerance. The above described approach has included all terms in the energy function through second order and is therefore quadratically convergent. We therefore denote this scheme the one-step second-order approach (Yeager and Jørgensen, 1979).

4. Two-Step Procedure

Another approach, which differs slightly in its realization of the iterative procedure, has also been used and is referred to here as the two-step second-order scheme. It may be described as follows: After an initial guess of orbitals, a configuration interaction calculation (see Section D) is carried out to determine the starting set of CI eigenstates $|l\rangle$. We then have

$$\langle m|H|l\rangle = E_l\delta_{ml} \tag{2.35}$$

and the **V** matrix

$$V_n = \langle 0|[R_n, H]|0\rangle = \langle 0|[|0\rangle\langle n|, H]|0\rangle = 0 \tag{2.36}$$

becomes equal to zero. Equation (2.34) may then be partitioned (Löwdin, 1968) to give (using $\mathsf{B}_{22} = 0$)

$$\boldsymbol{\kappa} = [\mathsf{A}_{11} - \mathsf{B}_{11} - (\mathsf{A}_{12} - \mathsf{B}_{12})\mathsf{A}_{22}^{-1}(\mathsf{A}_{21} - \mathsf{B}_{21})]^{-1}\mathbf{W} \tag{2.37}$$

and the $\boldsymbol{\kappa}$ matrix can be determined from this set of linear equations.

A transformed set of orbitals may now be obtained using this $\boldsymbol{\kappa}$ in Eq. (1.52) and a new CI calculation (diagonalization of $\langle l|H|m\rangle$) carried out. This process is then continued until convergence is reached. In the two-step second-order procedure, Eq. (2.34) is thus always applied in a basis where the states are determined from a CI calculation. The matrix **P** is never explicitly calculated. In contrast, in the one-step procedure the configuration expansion coefficients of $|0\rangle$ and $|n\rangle$ are determined from the unitary transformation given in Eq. (2.14), where **P** is obtained from Eq. (2.34) rather than from a CI calculation.

The terms $\mathsf{A}_{12} - \mathsf{B}_{12}$ coupling the configuration and orbital space variation have been neglected in many calculations. In many cases, these terms show little effect on the convergence rate of the procedure. It should, however, be pointed out that a quadratically convergent scheme is only obtained when these coupling terms are included.

5. Explicit Hessian and Generalized Brillouin Matrix Elements

Let us now consider the evaluation of some of the matrix elements appearing above. In the one-step procedure we have to calculate

$$(A_{22})_{mn} = \langle 0|[R_m, H, R_n^+]|0\rangle = \langle m|H|n\rangle - \delta_{mn}\langle 0|H|0\rangle \tag{2.38}$$

and

$$V_n = \langle 0|[R_n, H]|0\rangle = \langle n|H|0\rangle \tag{2.39}$$

The elements A_{22} and **V** thus contain all matrix elements contained in a configuration interaction calculation within the considered configuration space. When the iterative MCSCF procedure has converged, all elements of **V** are zero and the interactions between the reference state $|0\rangle$ and the residual states are thus eliminated. The diagonal and off-diagonal matrix elements of the Hamiltonian in the residual space $\{|n\rangle\}$ may, however, all be nonvanishing.

In the two-step second-order approach, the CI calculation that is carried out in each step prior to the evaluation of the matrices simplifies the evaluation of **V** and A_{22}. The **V** matrix becomes, as stated earlier, zero and the A_{22} matrix

$$(A_{22})_{mn} = \delta_{mn}(E_m - E_0) \tag{2.40}$$

becomes diagonal. These simplifications remain in each step of the iterative process because a CI calculation is performed in each iteration.

Except for A_{22} and **V**, the form of the matrix element in the one- and two-step procedures are the same. The matrix elements of **W**, A_{11}, and B_{11} may be derived from Eqs. (2.41) and (2.42) by index substitution. The excitation operators in these equations have singlet spin symmetry, since they arise in the operator λ, which must preserve the symmetry of $|0\rangle$ in forming $\exp(i\lambda)|0\rangle$. These matrices can be expressed in terms of one- and two-electron integrals and the one- and two-electron density matrices as given below. Note that no more than two-electron density matrices appear in **W**, A_{11}, and B_{11}:

$$\begin{aligned}\langle 0|[t_\alpha^+ u_\alpha + t_\beta^+ u_\beta, H]|0\rangle &= \sum_{\sigma p} h_{up}\langle 0|t_\sigma^+ p_\sigma|0\rangle - \sum_{\sigma p} h_{pt}\langle 0|p_\sigma^+ u_\sigma|0\rangle \\ &\quad - \sum_{pqr}\langle pq|rt\rangle\rho_{qpru} + \sum_{qrs}\langle uq|rs\rangle\rho_{tqsr}\end{aligned} \tag{2.41}$$

$$(pr|qt)\rho_{qu,pr} \quad (ur|qs)\rho_{tr,qs}$$

$$\begin{aligned}&\langle 0|[l_\alpha^+ k_\alpha + l_\beta^+ k_\beta, [H, t_\alpha^+ u_\alpha + t_\beta^+ u_\beta]]|0\rangle \\ &\quad = h_{kt}\sum_\sigma\langle 0|l_\sigma^+ u_\sigma|0\rangle + h_{ul}\sum_\sigma\langle 0|t_\sigma^+ k_\sigma|0\rangle - \delta_{kt}\sum_{p\sigma} h_{up}\langle 0|l_\sigma^+ p_\sigma|0\rangle \\ &\qquad - \delta_{lu}\sum_{p\sigma} h_{pt}\langle 0|p_\sigma^+ k_\sigma|0\rangle - \delta_{lu}\sum_{pqr}\langle pq|rt\rangle\rho_{pqkr} - \delta_{kt}\sum_{qrs}\langle uq|rs\rangle\rho_{lqsr} \\ &\qquad - \sum_{pq}\langle pq|lt\rangle\rho_{pquk} - \sum_{rs}\langle uk|rs\rangle\rho_{tlsr} + \sum_{pr}\langle kp|rt\rangle\rho_{lpur} \\ &\qquad + \sum_{pr}\langle kp|tr\rangle\rho_{plur} + \sum_{qs}\langle uq|ls\rangle\rho_{tqsk} + \sum_{qs}\langle uq|sl\rangle\rho_{qtsk}\end{aligned} \tag{2.42}$$

where

$$\rho_{il,jk}$$

$$\rho_{ijkl} = \sum_{\sigma,\sigma'}\langle 0|i_\sigma^+ j_{\sigma'}^+ k_{\sigma'} l_\sigma|0\rangle \tag{2.43}$$

and σ and σ' run over the electron spin indices α and β.

The elements of A_{21} and B_{21} reduce as follows:

$$\langle 0|[|0\rangle\langle n|, [H, t_\alpha^+ u_\alpha + t_\beta^+ u_\beta]]|0\rangle = \langle n|[H, t_\alpha^+ u_\alpha + t_\beta^+ u_\beta]|0\rangle \tag{2.44}$$

and an explicit final formula for Eq. (2.44) may be obtained from Eq. (2.41) by replacing the one- and two-electron density matrices with the corresponding transition density matrix elements.

6. Mode Damping

The GBT matrix and the Hessian matrix arising in the one-step second-order procedure determine the energy slope and curvature, respectively, for a given point on the hypersurface. When a SP point has been reached, the eigenvalues of the Hessian matrix thus can be used to characterize this point. We have reached a local minimum if all eigenvalues are positive. Mixed positive and negative eigenvalues correspond to a saddle point on the energy hypersurface. In employing the two-step procedure outlined above, one no longer has the opportunity to characterize the state by its Hessian eigenvalues, because the full Hessian matrix is not employed and the partitioned Hessian of Eq. (2.37) does not have the same eigenvalues as the full Hessian.

Some insight into the step lengths $(\boldsymbol{\kappa}, \mathbf{P})$ that should be taken in second-order procedures may be obtained by transforming the second-order equation to a form in which the Hessian matrix is diagonal. Let us consider initially the diagonalization (by the unitary matrix U) of the full $\mathsf{A} - \mathsf{B}$ matrix appearing in the one-step second-order equation

$$\mathsf{A} - \mathsf{B} = \mathsf{U}\boldsymbol{\varepsilon}\mathsf{U}^{+} \tag{2.45}$$

Equation (2.34) then becomes

$$\begin{pmatrix} \tilde{\boldsymbol{\kappa}} \\ \tilde{\mathbf{P}} \end{pmatrix} = \boldsymbol{\varepsilon}^{-1} \begin{pmatrix} \tilde{\mathbf{W}} \\ \tilde{\mathbf{V}} \end{pmatrix} \tag{2.46}$$

where

$$\begin{pmatrix} \tilde{\boldsymbol{\kappa}} \\ \tilde{\mathbf{P}} \end{pmatrix} = \mathsf{U}^{+} \begin{pmatrix} \boldsymbol{\kappa} \\ \mathbf{P} \end{pmatrix} \tag{2.47}$$

$$\begin{pmatrix} \tilde{\mathbf{W}} \\ \tilde{\mathbf{V}} \end{pmatrix} = \mathsf{U}^{+} \begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix} \tag{2.48}$$

Each normal mode on the energy hypersurface is decoupled and hence may be described independently. This is particularly useful in the initial iterations of an MCSCF calculation, where third- and higher-order terms may be important and even dominate as a result of the poor initial guess of the orbitals. The second-order scheme may, in such cases, be forced to take step lengths $(\tilde{\boldsymbol{\kappa}}, \tilde{\mathbf{P}})$ that are too large. The normal mode analysis of Eq. (2.46), which displays the slopes $(\tilde{\mathbf{W}}, \tilde{\mathbf{V}})$ and curvature $(\boldsymbol{\varepsilon})$ of each mode independently, then becomes a convenient tool to use for changing the step length for those modes that take very large steps. This is done by restricting the allowed size of the $\tilde{\boldsymbol{\kappa}}$ and $\tilde{\mathbf{P}}$ matrix elements. In ground-state calculations where the Hessian matrix has to be positive definite, we may even change the direction of the step (i.e., change the sign of $\tilde{\boldsymbol{\kappa}}$) if small negative eigenvalues $\boldsymbol{\varepsilon}_j$ appear. This situation occurs frequently in the initial iterations of actual calculations. If the matrix elements that couple the orbital and coefficient optimization (the $\mathsf{A}_{12} - \mathsf{B}_{12}$ matrix) are very small, $\tilde{\boldsymbol{\kappa}}$ then predominantly refers to the orbital optimization while $\tilde{\mathbf{P}}$ refers to the coefficient optimization. In these cases, it is reasonable to impose some different limits upon the size of the maximum elements of the step length vectors $\tilde{\boldsymbol{\kappa}}$ and $\tilde{\mathbf{P}}$. At present, there is little experience on how to optimally make these restrictions although results of initial calculations indicate that the basic philosophy is correct. When strong coupling occurs between the configuration and the orbital space, more refined damping schemes may need to be introduced (Yeager *et al.*, 1980).

In the two-step second-order procedure, damping may only be performed in the space that is dominated by the orbital space. From applying the unitary transformation to Eq. (2.37) we get

$$\tilde{\boldsymbol{\kappa}} = \boldsymbol{\varepsilon}^{-1}\mathbf{W} \tag{2.49}$$

where

$$\mathsf{A}_{11} - \mathsf{B}_{11} - (\mathsf{A}_{12} - \mathsf{B}_{12})A_{22}^{-1}(\mathsf{A}_{21} - \mathsf{B}_{21}) = \mathsf{U}\boldsymbol{\varepsilon}\mathsf{U}^{+} \tag{2.50}$$

$$\tilde{\boldsymbol{\kappa}} = \mathsf{U}^{+}\boldsymbol{\kappa} \tag{2.51}$$

$$\tilde{\mathbf{W}} = \mathsf{U}^{+}\mathbf{W} \tag{2.52}$$

Because the reference state $|0\rangle$ and its orthogonal complement states $|n\rangle$ are determined from a CI calculation, it is not generally possible to impose constraints on the step lengths in the configuration space. Further, the CI steps are not necessarily taken along the normal modes. In particular, when strong coupling elements exist between the configuration and orbital spaces, large fluctuations in the amplitude of the dominant configuration may be encountered, which may lead to difficulties in converging to the state under consideration.

7. Elimination of Redundant Operators

Having now given a general discussion of quadratically convergent second-order MCSCF methods together with some analysis of how such techniques might best be implemented, we can move on to describe other MCSCF methods, as well as to give more detail about the numerical requirements of such calculations. Before doing so, however, it is important that we address a technical point that must be understood if one is to be successful in carrying out such MCSCF calculations.

The unitary transformation of the reference state given in Eq. (2.18) has as generators the operators r^+s of λ and $|n\rangle\langle 0|$ of S. It is possible that the operators r^+s and $|n\rangle\langle 0|$ span the same space. That is, the effects of the operators r^+s may be expressed in terms of those of the state projections in the configuration space. To determine whether the effects of a given operator r^+s can be expressed in terms of the kets $\{|m\rangle\}$, we examine the following difference ket:

$$|f\rangle \equiv r^+s|l\rangle - \sum_m |m\rangle\langle m|r^+s|l\rangle \tag{2.53}$$

If the norm of $|f\rangle$ vanishes, then $|f\rangle$ itself vanishes and hence $r^+s|l\rangle$ can be exactly represented as a sum of the $\{|m\rangle\}$ functions. The norm of $|f\rangle$ vanishes when

$$\langle f|f\rangle = 0 = \langle l|s^+rr^+s|l\rangle - \sum_m \langle l|s^+r|m\rangle\langle m|r^+s|l\rangle \tag{2.54}$$

or, in other words, when

$$\sum_m |\langle m|r^+s|l\rangle|^2 = \langle l|s^+rr^+s|l\rangle \tag{2.55}$$

When *both* the operators r^+s and s^+r fulfill Eq. (2.55), for any state $|l\rangle$ the variations described by the parameters λ_{rs} will be denoted as redundant.

The search for redundant variables may, of course, alternatively be performed in the configuration space $\{|\phi_g\rangle\}$ since this space is related to the space $\{|l\rangle\}$ through a unitary transformation. Because the states $\{r^+s|\phi_g\rangle\}$ are normalized to unity, the search for redundant variables may be achieved by investigating whether the sum

$$\sum_{g'} |\langle \phi_{g'}|r^+s|\phi_g\rangle|^2 \tag{2.56}$$

is equal to zero or one for any state $|\phi_g\rangle$.

We now show how orbital changes caused by redundant variables can be represented as configuration changes caused by S and can thus be eliminated from the energy optimization procedure. The redundant set of operators form a hermitian operator

$$\lambda'' = \sum_{rs} \lambda''_{rs} r^+s \tag{2.57}$$

The operator λ', which contains all of the r^+s that are not redundant also forms a hermitian operator

$$\lambda' = \sum_{rs} \lambda'_{rs} r^+s \tag{2.58}$$

Since the operators in Eq. (2.57) are generators of a subgroup of the unitary group, $\exp(i\lambda)$ may be factorized to give

$$\exp(i\lambda) = \exp(i\lambda')\exp(i\lambda'') \tag{2.59}$$

Equation (2.59) is thus a representation of an arbitrary group element expressed as a left coset of this subgroup. Expressed in other words, the unitary transformation that is described by $\exp(i\lambda)$ may alternatively be described by the unitary transformation $\exp(i\lambda')\exp(i\lambda'')$. It should be pointed out that there exist no simple relations between the λ_{rs} parameters and the λ'_{rs} and λ''_{rs} parameters. With the above factorization of the "redundant" part (λ''), the unitary transformation of the reference state may be written as

$$|\tilde{0}\rangle = \exp(i\lambda')\exp(i\lambda'')\exp(iS)|0\rangle$$

Since $|f\rangle$ in Eq. (2.53) is zero for any product of redundant operators,

$$\exp(i\lambda'')|l\rangle = \sum_p |p\rangle\langle p|\exp(i\lambda'')|l\rangle \tag{2.60}$$

Using this relation together with Eq. (2.14) gives

$$\begin{aligned} |\tilde{0}\rangle &= \exp(i\lambda')\exp(i\lambda'')\sum_l |l\rangle[\exp(-P)]_{l0} \\ &= \exp(i\lambda')\sum_{p,l}|p\rangle\langle p|\exp(i\lambda'')|l\rangle[\exp(-P)]_{l0} \end{aligned} \tag{2.61}$$

The matrix $\{\langle p|\exp(i\lambda'')|l\rangle\}$ is unitary since the scalar product of Eq. (2.60) with $\langle q|\exp(-i\lambda'')$ gives

$$\sum_p \langle q|\exp(-i\lambda'')|p\rangle\langle p|\exp(+i\lambda'')|l\rangle = \delta_{ql}$$

Therefore, the product matrix

$$\sum_l \langle p|\exp(i\lambda'')|l\rangle \cdot [\exp(-P)]_{l0} \tag{2.62}$$

must consequently also be unitary. Because a single unitary transformation of the form given in Eqs. (2.10) and (2.14) is sufficient for optimizing the total energy, the redundant variables may be left out when optimizing the energy. That is, the λ'' factors can do nothing more, in a wavefunction optimization, than can be done by the $\exp(iS)$ operator.

8. Practical Considerations

So far, no attention has been given to the spatial and spin symmetry features of the reference state. The theory we have outlined thus far may hence be described as unrestricted multiconfigurational HF. In most applications (Eyring *et al.*, 1967), we require the reference state to have a certain symmetry (i.e., the reference state should transform according to an irreducible representation of the Hamiltonian's point group). For the wavefunction symmetry to be conserved under a sequence of unitary transformations, the operators λ and S have to be tensor operators belonging to the totally symmetric irreducible representation. In this way, the symmetry of the wavefunction would be conserved during the iteration procedure.

Calculations of the matrix elements that are used to define the above procedures requires knowledge of the one- and two-electron integrals in the MCSCF spin-orbital basis. Therefore, a two-electron integral transformation (Schaefer and Miller, 1977, Chapt. 6) has to be performed in each step of the iterative procedure. MCSCF approaches, in general, require such repeated two-electron integral transformations to be performed. Since these transformations may, in many cases, be the computationally most demanding step of the calculation, it becomes very important to use MCSCF procedures that converge reliably in a minimum number of iterations. We have chosen to emphasize here the one- and two-step second-order procedures because they are quadratically convergent and because they allow a controlled (damped) "walk" to be performed on the energy hypersurface when cubic and higher-order terms and/or coupling between orbitals and configuration optimizations are important.

9. Generalized Brillouin-Theorem-Based Procedures

So far we have used the condition that the energy function be stationary to define MCSCF schemes. The existence of a stationary point on the energy hypersurface requires that the GBT be fulfilled at this point. Hence, iterative MCSCF procedures may alternatively be developed by insisting that the GBT be satisfied as the iterative procedure converges. A quadratically convergent scheme may be obtained by further insisting that the error in the GBT matrix in the $(n + 1)$th iteration should be the square of the error in the nth iteration. Denoting the operators and states in the $(n + 1)$th iteration with a tilde and those of the n'th with no tilde, using Eqs. (2.18) and (1.38) we obtain

$$\begin{aligned}\mathbf{W}_{n+1} &= \langle\tilde{0}|[\tilde{Q}, H]|\tilde{0}\rangle = \langle 0|[Q, H]|0\rangle + i\langle 0|[Q, [H, \lambda]]0\rangle \\ &\quad + i\langle 0|[[Q, H], S]|0\rangle + O(\kappa^2, P^2) \qquad (2.63)\end{aligned}$$

$$\begin{aligned}\mathbf{V}_{n+1} &= \langle\tilde{0}|[\tilde{R}, H]|\tilde{0}\rangle = \langle 0|[R, H]|0\rangle + i\langle 0|[R, [H, \lambda]]|0\rangle \\ &\quad + i\langle 0|[R, [H, S]]|0\rangle + O(\kappa^2, P^2) \qquad (2.64)\end{aligned}$$

since, for example,

$$\begin{aligned}\langle 0|[Q, [H, \lambda]]|0\rangle - \langle 0|[[Q, H], \lambda]|0\rangle &= \langle 0|[[Q, \lambda], H]|0\rangle \\ &= \boldsymbol{\kappa}\langle 0|[[Q, Q - Q^+], H]|0\rangle \\ &= O(\kappa^2) \qquad (2.65)\end{aligned}$$

The last identity arises because the GBT is not fulfilled until convergence is reached, and thus $\langle 0|[[Q, Q - Q^+], H]|0\rangle$ is of order κ itself. The double commutator may be introduced in Eq. (2.63), and Eqs. (2.63) and (2.64) may then be combined to give

$$\begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix}_{n+1} = \begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix}_{n} - (\mathsf{A} - \mathsf{B})_n \begin{pmatrix} \boldsymbol{\kappa} \\ \mathbf{P} \end{pmatrix} + O(\kappa^2, P^2) \tag{2.66}$$

where we have used Eqs. (2.27)–(2.30). A quadratically convergent scheme is thus obtained when the $\boldsymbol{\kappa}$ and $\mathbf{P}$ matrices are determined from

$$\begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix}_{n} = (\mathsf{A} - \mathsf{B})_n \begin{pmatrix} \boldsymbol{\kappa} \\ \mathbf{P} \end{pmatrix} \tag{2.67}$$

which is identical to the one-step second-order equation [Eq. (2.33)]. Hence the one-step second-order procedure described earlier can also be viewed as arising from the GBT.

Most MCSCF procedures that have been employed to date (Schaefer and Miller, 1977, Chapters 3 and 4) have concentrated on deriving iterative schemes based upon only insisting that

$$\langle 0|[H, r^+ s]|0\rangle = 0 \tag{2.68}$$

in each step of the iterative procedure. As successive sets of MCSCF orbitals are determined in each step of the iterative procedure, the configuration space equivalent of the GBT $\langle 0|[H, |n\rangle\langle 0|]|0\rangle = 0$ is achieved through performing a CI calculation within the limited configuration space.

To see how Eq. (2.68) can be used to define an iterative process, let us consider the first two terms in the expansion of $\exp(i\lambda)|0\rangle$:

$$|0\rangle - \sum_{r>s} \kappa_{rs}(r^+ s - s^+ r)|0\rangle \tag{2.69}$$

This first-order approximation to the true $\exp(i\lambda)|0\rangle$ then leads us to consider the variational wavefunction

$$|\tilde{0}\rangle \simeq X_0|0\rangle + \sum_{r>s} X_{rs}(r^+ s - s^+ r)|0\rangle \tag{2.70}$$

containing the linear variational parameters X_0 and $\{X_{rs}\}$. The optimal values of these parameters may then be determined from the superconfiguration interaction (SCI) secular problem (Banerjee and Grein, 1976)

$$\mathsf{H}\mathbf{X} = E\mathsf{S}\mathbf{X} \tag{2.71}$$

The SCI Hamiltonian matrix elements are defined as

$$H_{0,rs} = \langle 0|H(r^+ s - s^+ r)|0\rangle \tag{2.72}$$

which for real orbitals reduces to Eq. (2.68):

$$H_{0,rs} = \langle 0|[H, r^+s]|0\rangle = H_{rs,0} \tag{2.73}$$

The other matrix elements of H are

$$H_{00} = \langle 0|H|0\rangle \tag{2.74}$$

$$H_{rs,tu} = \langle 0|(r^+s - s^+r)^+ H(t^+u - u^+t)|0\rangle \tag{2.75}$$

The scalar product matrix S is defined in a similar fashion (e.g., $S_{0,rs} = \langle 0|r^+s - s^+r|0\rangle = 0$) and the eigenvector **X** has the components $\mathbf{X} = \{X_0, X_{rs}\}$. The GBT therefore states that in the so-called SCI secular problem [Eq. (2.71)], the state $|0\rangle$ should be noninteracting with its single excitations $(r^+s - s^+r)|0\rangle$. Once this occurs, Eq. (2.71) will have, as one of its eigenvalues, the MCSCF energy $\langle 0|H|0\rangle$. The other eigenvalues, as in all variational secular problems, represent upper bounds to other true energy levels.

The eigenvector **X** obtained from the SCI secular problem can be used to define a transformation of the orbital appearing in $|0\rangle$. To see how this transformation arises, we rewrite Eq. (2.70) as

$$|\tilde{0}\rangle = X_0^{1-N}\left[X_0^N + \sum_{r>s} X_0^{N-1} X_{rs}(r^+s - s^+r)\right]\sum_g C_{g0}|\phi_g\rangle \tag{2.76}$$

The effect of $\sum_{r>s}(r^+s - s^+r)X_{rs}$ on each configuration $|\phi_g\rangle$ results in two new configurations in which spin orbital ϕ_s is replaced by ϕ_r and vice versa. For example, the effect on $1^+2^+ \cdots N^+|\text{vac}\rangle$ is to give

$$\sum_{i=1}^{N}\left[\sum_{r>i} X_{ri}1^+2^+ \cdots (i-1)^+r^+(i+1)^+ \cdots N^+ \right.$$
$$\left. - \sum_{r<i} X_{ir}1^+2^+ \cdots (i-1)^+r^+(i+1)^+ \cdots N^+\right]|\text{vac}\rangle \tag{2.77}$$

If the spin-orbitals occupied in any configuration $|\phi_g\rangle$ are denoted by ϕ_t, then the above SCI wavefunction in Eq. (2.76) can be expressed as

$$|\tilde{0}\rangle = X_0^{1-N}\sum_g C_{g0}\prod_{t\in g}\left[X_0t^+ + \sum_{r>t} X_{rt}r^+ - \sum_{r<t} X_{tr}r^+\right]|\text{vac}\rangle + O(X_{rs}^2) \tag{2.78}$$

That is, the wavefunction used in the SCI calculation (Eq. (2.70)) is identical, through first order in the X_{rs} parameters, to a new linear combination of configurations with the same C_{g0} coefficients but with orbitals $\tilde{\phi}_t$ that can be expressed in terms of the original orbitals as

$$\tilde{\phi}_t = X_0\phi_t + \sum_{r>t} X_{rt}\phi_r - \sum_{r<t} X_{tr}\phi_r \tag{2.79}$$

Because this orbital transformation is not properly unitary (the $\{\tilde{\phi}_k\}$ are normalized and orthogonal only through first order in the X_{rs}), the set $\{\tilde{\phi}_k\}$ must, in each iteration, be orthonormalized (by, for example, the Schmidt or Löwdin procedure).

The SCI iterative procedure thus consists of guessing a starting set of orbitals and generating the $\{C_{g0}\}$ expansion coefficients from a CI calculation. The SCI secular problem is then constructed and solved (to give $\mathbf{X}$) after which the new orbitals $\{\tilde{\phi}_k\}$ are computed as in Eq. (2.79) and subsequently orthonormalized. These new orbitals are then used to perform a new CI calculation to generate new $\{C_{g0}\}$ coefficients and hence a new SCI secular problem. This iterative procedure is continued until convergence is achieved at which time the GBT is fulfilled. A significant drawback of most SCI procedures as now implemented is that they do not treat the coupling between orbital and configuration optimization. SCI methods that treat both optimizations on equal footing represent a significant improvement. In situations for which strong coupling exists between the orbital and configuration space, the above-described two-step SCI process might thus be expected to converge slowly. As we mentioned above, the quadratically convergent one-step second order procedure discussed in the preceding section could also be viewed as being defined, through Eq. (2.66), to make the GBT obeyed. It is then important to explore how the two iterative methods, both of which can be stated through the GBT, differ. The difference arises from terms in

$$\langle 0|\exp(-iS)\exp(-i\lambda)H\exp(i\lambda)\exp(iS)|0\rangle$$

that are quadratic in S or λ and that arise from the second-order components of the individual exponential operators. For example, $\langle 0|Hi\lambda i\lambda|0\rangle$ and $\langle 0|iSiSH|0\rangle$ arise in the exponential formulation but do not arise in the expectation value of the SCI wavefunction given in Eq. (2.69). The neglect of second-order terms and the requisite reorthogonalization of the MCSCF orbitals differentiate between the two methods and render the SCI approach not quadratically convergent.

Because SCI approaches to the MCSCF problem are not based upon extremizing the full second-order energy expression described above, their convergence rate is linear rather than quadratic, although in practice such SCI methods may sometimes demonstrate approximate quadratic convergence. Because the SCI energies result from solutions of an eigenvalue problem, each SCI energy is an upper bound to the respective true energies (ground and excited). The values of X_{ij} obtained from the SCI secular problem [Eq. (2.71)] when used to carry out orbital modifications [through Eq. (2.79)] yield a new multiconfigurational wavefunction whose Hamiltonian expectation value is, because of the subsequent orthonormalization needed, no longer identical to the eigenvalue E that was obtained from the SCI secular problem.

In the unitary second-order method, the energy expression $E(\lambda, S)$ given in Eq. (2.25), when truncated after terms linear and quadratic in λ and S, is no longer an expectation value of H and thus no longer bounds the ground-state–total energy. Thus, the stationary points of $E(\lambda, S)$ do not form rigorous upper bounds to the respective true ground- and excited-state energies. Of course, there are good reasons to believe that, in the neighborhood of an eigenstate, $E(\lambda, S)$ can be well approximated by this quadratic hypersurface. Moreover, the values of $\mathbf{P}$ and $\boldsymbol{\kappa}$ obtained from making $E(\lambda, S)$ stationary, when used in Eqs. (1.52) and (2.14) to obtain $|\tilde{0}\rangle$, do give a proper upper-bound energy through $\langle\tilde{0}|H|\tilde{0}\rangle$.

Having now discussed how one can go about optimizing the electronic energy of an MCSCF wavefunction, we turn our attention to two special subclasses of this procedure; the single-configuration SCF problem and the frozen-orbital CI problem. Because we choose to view these situations as special cases of the above MCSCF problem, we obtain a specialized view of SCF and CI theory. There already exist in the literature extensive and clear treatments of SCF and CI as they are more commonly treated within the linear variational framework. Hence we have not attempted to cover the more conventional aspects of these topics here.

C. SINGLE-CONFIGURATION SELF-CONSISTENT FIELD METHODS

1. Quadratically Convergent Scheme

Let us consider a situation in which we choose to work with a one-configuration wavefunction for which the orbitals are allowed to vary. This single configuration $|0\rangle$ may still consist of a linear combination of determinants whose (fixed) coefficients are determined by the space and spin symmetry imposed on $|0\rangle$. The orbital variations may be described by $\exp(i\lambda)$ and an optimal set of orbitals determined as in the previous section [by simply neglecting terms involving $\exp(iS)$]. The second-order Eq. (2.33) then reads

$$\mathbf{W} = (\mathsf{A}_{11} - \mathsf{B}_{11})\boldsymbol{\kappa} \tag{2.80}$$

where A_{11} and B_{11} are defined in Eqs. (2.29) and (2.30). A quadratically convergent scheme for optimizing orbitals may be described as follows. Given an initial guess for the "occupied" orbitals, we use Eq. (2.80) to determine $\boldsymbol{\kappa}$, and then we use Eq. (1.52) to generate a transformed set of orbitals. This process is repeated until convergence is reached. This process requires, even in the one-configuration case, a partial two-electron integral transformation in each step of the iterative procedure. For a single-configuration case, a more restricted two-electron integral transformation can be used in each step if, instead of the above quadratic procedure, one uses an approach that is based on the Brillouin condition alone. These so-called first-order BT-based self-consistent-field (SCF) procedures are, however, not quadratically convergent, much as the SCI method treated earlier is only a linearly convergent MCSCF method.

2. Brillouin-Theorem-Based Methods

The HF or SCF approaches based upon the BT itself,

$$\langle 0|[H, r^+ s]|0\rangle = 0 \tag{2.81}$$

introduce a decomposition of the Hamiltonian into a Fock operator (which the spin-orbital basis is chosen to diagonalize)

$$F = \sum_{r,s} (h_{rs} + V_{rs}) r^+ s = \sum_r \varepsilon_r r^+ r \tag{2.82}$$

where h_{rs} is the one-electron part of the Hamiltonian. A Fock potential

$$V = \sum_{r,s} V_{rs} r^+ s \tag{2.83}$$

and the electron repulsion term W combine with F so that

$$H = F - V + W \tag{2.84}$$

The one-electron Fock potential V is thus far arbitrary. Different choices for V correspond to different choices of the spin-orbitals $\{\phi_r\}$ and their corresponding orbital energies $\{\varepsilon_r\}$, since we require the ϕ_r and ε_r to obey

$$h_{rs} + V_{rs} = \delta_{rs}\varepsilon_r \tag{2.85}$$

The BT [Eq. (2.81)] can now be used to determine V and hence to determine the spin-orbitals ϕ_r. By inserting the H of Eq. (2.84) into Eq. (2.81) we obtain

$$\begin{aligned} 0 = \langle 0|[r^+ s, H]|0\rangle &= (\varepsilon_s - \varepsilon_r)\langle 0|r^+ s|0\rangle + \sum_j (V_{jr}\langle 0|j^+ s|0\rangle - V_{sj}\langle 0|r^+ j|0\rangle) \\ &+ \sum_{k,l,j} (\langle sk||jl\rangle\langle 0|r^+ k^+ lj|0\rangle + \langle lk||jr\rangle\langle 0|l^+ k^+ js|0\rangle) \end{aligned} \tag{2.86}$$

where

$$\langle ij||kl\rangle \equiv \langle ij|kl\rangle - \langle ij|lk\rangle \tag{2.87}$$

Since the one-electron density matrix is diagonal for the single-configuration case considered here, we have

$$\langle 0|r^+ s|0\rangle = \delta_{rs}\nu_s \tag{2.88}$$

where ν_s denotes the occupation number for orbital ϕ_s in $|0\rangle$. Because $|0\rangle$ may consist of a linear combination of determinants, the ν_s are not necessarily zero or unity. The Fock potential determined from Eq. (2.86) is then

$$V_{rs}(\nu_r - \nu_s) = \sum_{k,l,j} (\langle sk||jl\rangle\langle 0|r^+k^+lj|0\rangle + \langle lk||jr\rangle\langle 0|l^+k^+js|0\rangle) \tag{2.89}$$

which is only defined from the Brillouin condition when $\nu_r - \nu_s$ is nonzero. Notice that the symmetry of the Fock operator defined in Eq. (2.82) is determined by the symmetry of the above Fock potential. This in turn depends upon the symmetry of the density matrices appearing in Eq. (2.89). As a result, the Fock operator may not have the same symmetry as the full electronic Hamiltonian for specific choice of the reference state $|0\rangle$.

Before discussing various possibilities for how to choose the part of the Fock potential that is not determined from the BT, let us describe the iterative procedure that can be used for obtaining a set of optimized orbitals given any final choice for the form of the full Fock potential. From an initial guess of orbtials, we use Eq. (2.89) together with one of many choices of the remainder of the V to determine a Fock potential. The Fock matrix $\mathsf{F} = \mathsf{h} + \mathsf{V}$ (which is hermitian) is then diagonalized, and a new set of orbitals is determined, which are then used to set up a new Fock potential. This (first-order) process is continued until convergence. The above HF iteration process is nothing but a variant of the commonly used Roothaan SCF procedure (Roothaan, 1951, 1960).

3. Choices of the Nondefined Blocks of the Fock Potential

The part of the Fock potential not defined through the Brillouin condition is often chosen on physical ground [e.g., to have the resultant orbital energies represent ionization potentials and electron affinities (via Koopmans' theorem)] (McWeeney and Sutcliffe, 1976). For a reference state containing a set of occupied spin-orbitals that we denote by α, β, γ, δ and a set of unoccupied spin-orbitals denoted m, n, p, q, the Fock potential in Eq. (2.89) is defined by the BT only between occupied and unoccupied orbitals. From Eq. (2.89) we get

$$V_{m\alpha} = \sum_{\gamma} \langle m\gamma||\alpha\gamma\rangle \tag{2.90}$$

One commonly used choice of the remaining blocks of V results in a Fock potential that can be expressed as

$$V = \sum_{\gamma,r,s} \langle r\gamma || s\gamma \rangle r^{+} s \tag{2.91}$$

where r and s run over all spin-orbitals. With this choice, the orbital energies represent (through Koopmans' theorem) the ionization potentials and electron affinities of $|0\rangle$. Of course, other choices of the nondetermined part of V have been made in the literature. For example, the (unoccupied–unoccupied) part of the Fock potential (V_{mn}) has been chosen to correspond to a so-called V^{N-1} potential (Kelly, 1964), thereby making the virtual orbitals more suitable for use in the calculation of excitation energies.

Calculations such as the one discussed above do not involve imposed symmetry restrictions on the reference wavefunction. Hence this approach is referred to as the unrestricted Hartree–Fock (UHF) method. When symmetry restrictions are imposed upon the reference wavefunction the resulting calculation is denoted a restricted Hartree–Fock (RHF) calculation. When the simplest RHF type calculation is carried out for a closed-shell reference state (i.e., one having doubly occupied orbitals), the nondefined part of the Fock potential (the occupied–occupied) and (empty–empty) part is often chosen to have the same form as the (occupied–empty) part defined from the BT. We then would obtain for the entire Fock potential

$$V = \sum_{\substack{\gamma \\ r,s}} (2\langle r\gamma | s\gamma \rangle - \langle r\gamma | \gamma s \rangle)(r_{\alpha}^{+} s_{\alpha} + r_{\beta}^{+} s_{\beta}) \tag{2.92}$$

where the indices r, s, and γ refer to orbital indices and the subscripts α, β denote the electron spin m_s component. The orbital energies ε_r then correspond to approximate ionization energies. For a state that has some doubly occupied and some partially filled orbitals, the choice of the nondefined blocks of the Fock potential is less obvious. The BT defines the blocks that connect (occupied–partly occupied), (occupied–empty), and (partly occupied–empty) orbitals. The (occupied–occupied), (partly occupied–partly occupied), and (empty–empty) blocks of the Fock potential are not defined through the BT and many choices have been suggested. One common feature of any of these choices is that the sets of orbitals one obtains in a converged calculation using any arbitrary choice of the nondefined Fock matrix blocks would represent the same SP on the energy hypersurface. The physical interpretations of the orbital energies do, of course, depend on the actual choices made for these "diagonal blocks" of V. For this reason, much work has been devoted to finding particular choices of diagonal blocks that are optimal for particular physical situations. It is not our intention to provide a lengthy discussion of the merits and weaknesses of numerous such methods. Rather, we merely wish to stress that the undetermined blocks of V represent a certain freedom or flexibility that can be exploited to generate orbitals whose orbital energies have some approximate physical meaning.

4. Practical Considerations

Although it is not obvious from Eq (2.89) that a two-electron integral transformation is not required to set up the Fock potential matrix for a general reference state, it becomes clear upon actually working out the matrix elements for a particular case. For example, for either a spin-unrestricted reference state or a closed-shell reference state, the Fock potentials of Eqs. (2.91) and (2.92), respectively, are seen to involve only a two-index transformation [e.g., sum over γ in Eq. (2.92)].

From the above discussion it should be clear that the first-order procedures based upon using the Brillouin condition to define V suffer from some drawbacks. They involve arbitrary choices of certain elements of V (this is related to the invariance of $|0\rangle$ under certain orbital rotations). They are not quadratically convergent and may thus suffer from convergency difficulties. On the other hand, the freedom in choosing elements of V (including the diagonal blocks) is useful when one wishes to cause the resultant orbital energies to have certain physical interpretations (e.g., Koopmans' theorem of ionization energies or excitation energies). The $\exp(i\lambda)$ approach to HF orbital optimization is quadratically convergent but contains no orbital energies for use in physical interpretation. It avoids the problems related to arbitrary choices by simply eliminating from the orbital optimization operator space those operators $(r^+s - s^+r)$ that are redundant and that therefore have no effect on the energy to be extremized.

D. CONFIGURATION INTERACTION METHOD

1. Connection with Second-Order MCSCF Theory

Next we consider the optimization of the total energy when orbital relaxation is not explicitly accounted for in the calculation. The optimization of the total energy may then be carried out either in terms of the configuration expansion coefficients C_{g0} of Eq. (2.4) or in terms of the parameters P of Eq (2.9). Let us consider initially the optimization of the total energy when the configuration expansion coefficients of Eq. (2.4) are used as linear variational parameters. The total energy then becomes

$$E(C_{10}, C_{20}, \ldots) = \sum_{g,g'} C_{g'0} C_{g0} \langle \phi_{g'} | H | \phi_g \rangle \Big/ \sum_g |C_{g0}|^2 \qquad (2.93)$$

where we have assumed that the configuration basis states $|\phi_g\rangle$ are orthonormal. Because the energy function contains no higher than quadratic terms in the C_{g0}, determination of stationary points of the energy function

$$\delta E(C_{10}, C_{20}, \ldots) = 0 \tag{2.94}$$

leads to a set of eigenvalue equations in the configuration expansion coefficients

$$\mathsf{H}\mathbf{C}_0 = E\mathbf{C}_0 \tag{2.95}$$

where H is the matrix representative of the Hamiltonian

$$H_{gg'} = \langle\phi_g|H|\phi_{g'}\rangle \tag{2.96}$$

and the eigenvector

$$\mathbf{C}_0 = \{C_{10}C_{20} \cdots C_{g0}\} \tag{2.97}$$

determines the values of the set of parameters at the SP, where the value of E is E_0. In fact, the same eigenvalue equation, Eq. (2.95), can be used to determine all extrema of the energy within a given configuration space because the energy function contains no more than quadratic terms in $\mathbf{C}$. Equation (2.95) is referred to as the CI eigenvalue equation.

The optimization of the total energy might alternatively be expressed in terms of the variation parameters P [in $\exp(iS)$]. The energy function $E(S)$ would not be quadratic in these parameters P but would contain cubic, quartic, etc. terms in P. An explicit solution from which to determine a SP of the energy function when this unitary $\exp(iS)$ operator is used is very difficult to establish; hence an iterative procedure is required to determine SPs of the energy hypersurface. One iterative scheme that is quadratically convergent is obtained if the terms that refer to the orbital optimization [$\exp(i\lambda)$] are neglected in the MCSCF derivation performed in Section B. The second-order Eq. (2.33) then would read

$$\mathbf{V} = \mathsf{A}_{22}\mathbf{P} \tag{2.98}$$

where A_{22} is defined in Eq. (2.29) as

$$(A_{22})_{mn} = \langle m|H|n\rangle - \delta_{mn}\langle 0|H|0\rangle \tag{2.99}$$

$$\mathbf{V}_n = \langle 0|[R_n, H]|0\rangle = \langle n|H|0\rangle \tag{2.100}$$

and the indices n, m are different from 0. The iterative procedure may be described as follows. For an initial set of configurations ($|0\rangle, |n\rangle$) the matrices A_{22} and $\mathbf{V}$ can be formed. The matrix $\mathbf{P}$ then is determined from Eq. (2.98), and Eq. (2.14) is used to obtain a transformed set of states [one can use alternatively Eqs. (2.15) and (2.16)], and the whole process is repeated until convergence is obtained. The state $|0\rangle$ that would be determined in this iterative procedure would, of course, be the same as that obtained by solving the CI eigenvalue problem.

In the derivation of Section B we considered the energy function to depend on both orbital variation parameters and the configuration expansion coefficients. By freezing the orbital variation parameters, we prohibit orbital relaxation effects from being considered explicitly. To obtain with a CI calculation, which does not permit such orbital relaxation, the same quality as in an MCSCF calculation would require the inclusion of many more configurations, whose purpose would be to compensate for the neglect of explicit orbital relaxation. These additional functions would include a large number of singly excited configurations, but some double, triple, etc. excited configurations would also be needed to fully compensate. If all configurations arising from a given orbital basis were included in a CI calculation (full CI), the need for considering orbital relaxation effects explicitly would, of course, not be present because all orbital variation parameters (λ) would then be redundant variables. However, the number of configurations required to perform a full CI calculation is usually prohibitively large even for systems of modest size. Because CI expansions converge very slowly (as a function of the dimension of the CI secular problem) and the requisite computer time increases very rapidly as more and more configurations are included, efforts must be made to optimize the convergence of a CI calculation by facing two major problems. First, we must make a reasonable choice of orbitals to use in the calculation, and second, the configurations that are to be included in the calculation must be picked by some physically motivated procedure.

2. Choice of Orbitals for Use in CI

The most commonly used set of spin-orbitals for setting up a CI matrix eigenvalue problem is the set of orbitals obtained in a RHF calculation. These orbitals form a particularly convenient set in the sense that they fulfill the BT (i.e., there are no matrix elements connecting the HF ground state and singly excited configurations). However, these orbitals are not especially well suited for use in the CI problem if one desires a reasonably short CI expansion to give high precision. One major problem with the HF orbitals comes from the fact that the electrons in the virtual canonical HF orbitals "feel" an N-electron potential and not an $N - 1$ electron potential, as would be physically more proper.

One partial solution to this problem is to use a set of orbitals obtained in a MCSCF calculation for setting up a CI matrix problem whose dimension is much larger than that of the MCSCF calculation. The orbitals would then be relaxed with respect to the configurations included in the MCSCF calculation. Since this MCSCF function normally would include the dominant configurations, a very large fraction of the orbital relaxation effects would would thus explictly be accounted for in the following CI calculations.

Another set of orbitals that has been used as a basis for CI calculations is the so-called iterative natural orbitals (INO) (Bender and Davidson, 1967), which are obtained in the following manner: From a limited number of configurations (the same in all iterations) a reference state $|0\rangle$ is determined by the CI procedure. This reference state is then used to set up the one-electron density matrix $\langle 0|r^+ s|0\rangle$, which upon diagonalization gives a set of "natural orbitals." These orbitals are then used for setting up a new CI problem, a new reference state $|0\rangle$ is then determined, and the procedure is continued until a self-consistent set of natural orbitals is determined. Clearly, the INOs are not identical to the MCSCF orbitals discussed earlier. The former are obtained by diagonalizing the first-order density matrix, whereas the MCSCF orbitals are determined by minimizing the electronic energy. The use of INOs in CI calculations is motivated by Löwdin's (Löwdin, 1955) analysis, which showed that such orbitals result in the most compact configuration expansion of $|0\rangle$ (i.e., the fewest configurations being required to generate a wavefunction of a given overlap with the true wavefunction).

The choice of configurations to include in an INO calculation requires particular attention. If the configuration list only includes configurations that are doubly excited with respect to each other, any set of orbitals would be natural orbitals. To make the natural orbital concept useful, the list of configurations has to contain configurations that are singly excited with respect to each other. For example, for the ground state of the beryllium atom, a natural choice of configurations in an MCSCF calculation would be $1s^2 2s^2$ and $1s^2 2p^2$. In the INO calculation, the configuration list would further have to include $1s^2 2sns$ and $1s^2 2pnp$, $n = 3, 4, 5, \ldots$. These configurations would then, to a certain degree, simulate the orbital optimization parameters κ_{ns2s} and κ_{np2p} contained in the MCSCF calculation.

3. Selection of Configurations

Let us now move on to discuss some basic ideas (Schaefer and Miller, 1977, Chapter 6) behind selecting the number of configurations to be included in the CI calculation. With a well-chosen set of orbitals, it is thought that a very small fraction of all possible configurations gives the most important contributions to the total energy. Estimates of the importance of the individual configurations may be obtained from a perturbation theory analysis of the CI secular problem [Eq. (2.95)]. An order analysis based upon Rayleigh–Schrödinger perturbation theory (RSPT) shows immediately the order in which any particular class of configurations enters into the wavefunction. For example, for a set of HF orbitals of a closed-shell system, only the doubly excited configurations contribute to the first-order wavefunction (see Section 3.F). Estimation of the coefficients of the individual configurations through perturbation theory may then be used to select the important configurations by specifying a certain tolerance for the coefficient (or the energy contribution) below which the configurations are not included. For cases in which several configurations are very important to the description of the system, these configurations may be used to form a so-called reference space whose coupling with other configurations can then be estimated through perturbation theory. Another approach is based on performing a series of $(n + 1)$-dimensional CI calculations among the n-dimensional reference space and a sequence of configurations that are obtained as low-order excitations out of these reference functions. The criterion for rejecting configurations tested in this manner usually has to do with the energy lowering of one or more of the n reference-state energies caused by the "added configuration" (Buenker and Peyerimhoff, 1974).

4. Treating Large CI Matrices—Direct Methods

When any such preselection of configurations has been performed, one is often faced with the problem that 10–300,000 configurations have to be included in the final CI calculation. Conventional matrix diagonalization routines such as the one used in the Householder algorithm, which modifies the elements of the matrix as it proceeds, cannot be used to determine the eigenvalues and eigenvectors of the CI matrix. For this reason, specialized approaches have been developed (Schaefer and Miller, 1977, Chapters 7 and 8) to determine a few selected roots (usually the lowest) of such very large CI matrices. One very important feature of these methods is that they do not entail modification of the CI matrix while determining a particular root. To clarify this point, we describe two such techniques, which are referred to as the power method and the perturbation theory method. Although much more efficient approaches have become available, we have chosen to discuss these techniques because they stress, in a simple manner, the basic principles underlying the direct determination of particular eigenstates. In the power method one considers a sequence of operations of the Hamiltonian matrix on an, in principle, arbitrary initial guess of the state vector $\mathbf{C}^0$:

$$\{\mathbf{C}^0, \mathsf{H}\mathbf{C}^0, \mathsf{H}^2\mathbf{C}^0, \ldots, \mathsf{H}^n\mathbf{C}^0\} \tag{2.101}$$

The information content of the converged vector $\lim_{n\to\infty} \mathsf{H}^n\mathbf{C}^0$ can be understood by expanding the vector $\mathbf{C}^0$ in terms of the (unknown) exact eigenvectors $\mathbf{C}_i$ of the Hamiltonian matrix

$$\mathsf{H}\mathbf{C}_i = E_i\mathbf{C}_i \tag{2.102}$$

$$\mathbf{C}^0 = \sum_n a_n\mathbf{C}_n, \qquad a_n = \langle \mathbf{C}^0 | \mathbf{C}_n \rangle \tag{2.103}$$

By assuming that the eigenvalues of H are ordered such that

$$|E_0| \geq |E_1| \geq |E_2| \geq \cdots \geq 0 \tag{2.104}$$

we obtain the formal result

$$\mathsf{H}^n\mathbf{C}^0 = E_0^n\left\{a_0\mathbf{C}_0 + \sum_{j=1} a_j\left(\frac{E_j}{E_0}\right)^n \mathbf{C}_j\right\} \tag{2.105}$$

which, because $|E_j/E_0| < 1$, reduces for large n to

$$\mathsf{H}^n\mathbf{C}^0 = a_0E_0^n\mathbf{C}_0 \tag{2.106}$$

Of course, to arrange the energy ordering assumed above, one might have to subtract from all diagonal elements of H a constant that depends on the largest positive diagonal H_{ii} element. This constant would then be added back onto the resultant E_0 value to obtain the true lowest desired eigenvalue. Hence we see that, for large enough n, the vectors $\mathsf{H}^{n+1}\mathbf{C}^0$ and $\mathsf{H}^n\mathbf{C}^0$ should be proportional, with their proportionality constant equal to E_0, and $\mathbf{C}_0$ should be the eigenvector of the Hamiltonian matrix having the largest eigenvalue E_0. Notice from Eq. (2.106) that the norm of $\mathsf{H}^n\mathbf{C}^0$ grows with n; therefore, normalization of the eigenvector $\mathbf{C}^0$ may be required during the above iterative scheme. E_0 and $\mathbf{C}_0$ are obtained without ever modifying the elements of matrix H; only simple row-by-row multiplication of H with a vector is involved. In fact, as we show below, one can even circumvent the explicit reference to elements of H by using integral-driven matrix multiplication techniques. Such steps become advantageous when one must avoid having to read through the integrals many times. The convergency rate of the power method is governed by the ratio E_1/E_0 and by the choice of $\mathbf{C}^0$. An inappropriate choice of the initial state vector $\mathbf{C}^0$ may lead to slow convergence (e.g., if a_0 vanishes, the power method, in principle, cannot converge to $\mathbf{C}_0$). Once one has obtained the desired E_0 and $\mathbf{C}_0$, the next eigenvalue of H can be found by employing $\mathsf{H} + |E_0||\mathbf{C}_0\rangle\langle\mathbf{C}_0|$ instead of H in the next application of the power method. The lowest root of this

$$(\mathsf{H} + |E_0||\mathbf{C}_0\rangle\langle\mathbf{C}_0|)$$

matrix should then be E_1.

The power method as outlined above is not very widely used in large-scale CI calculations because it is not usually very rapidly convergent. In contrast, variants of the perturbation method described below do constitute viable approaches to finding eigenvalues of large CI matrices. In the basic perturbation method one introduces a decomposition of the full CI Hamiltonian matrix

$$\mathsf{H} = \mathsf{H}^0 + \mathsf{V} \tag{2.107}$$

In what follows, we make the simplest possible choice of H^0; we take it to be the diagonal part of H. Another choice of H^0 that has been widely used. (Davidson, 1975) involves taking H^0 to be a certain small subblock of H (with H_{kk} elements filling the remaining diagonal entries of H^0), which involves the Hamiltonian matrix elements of the most dominant configurations in the desired eigenvector. Given a choice of H^0, the CI secular problem becomes

$$(\mathsf{H}^0 - E)\mathbf{C} = -\mathsf{V}\mathbf{C} \tag{2.108}$$

By iterating on this equation according to the prescription

$$\mathbf{C}^{(n)} = (E - \mathsf{H}^0)^{-1}\mathsf{V}\mathbf{C}^{(n-1)} \tag{2.109}$$

one generates successively higher approximations to the desired $\mathbf{C}$ vector. Corrections to the eigenvalue E are achieved at each iteration by premultiplying Eq. (2.108) on the left by the transpose of $\mathbf{C}^0$ to yield

$$(E^0 - E)(\mathbf{C}^0)^{\mathrm{T}}\mathbf{C} = -(\mathbf{C}^0)^{\mathrm{T}}\mathsf{V}\mathbf{C} \tag{2.110}$$

Initial estimates $\mathbf{C}^0$ and E^0 must, of course, be made consistent with the choice of H^0. For the diagonal choice of H^0, $\mathbf{C}^0$ would correspond to a unit vector $\mathbf{C}^0 = (1, 0, 0, \ldots, 0)$ and E^0 to the diagonal element of H ($E^0 = H_{11}$). If H^0 were taken to be a small subblock of a very large H matrix, Eq. (2.108) could still be solved perturbatively since the dimension of the matrix $(\mathsf{H}^0 - E)$ to be inverted would not be large. The iterative scheme contained in Eqs. (2.109) and (2.110) generates successively higher-order corrections to the desired energy and eigenvector.

To demonstrate how such perturbative methods lead to so-called direct CI techniques, let us consider a simple application of Eqs. (2.109) and (2.110) to a CI wavefunction consisting of a dominant HF configuration $|\phi_{\mathrm{HF}}\rangle$ plus all pair excitations of the form $|\phi_\mu^m\rangle = m_\alpha^+ m_\beta^+ \mu_\beta \mu_\alpha |\phi_{\mathrm{HF}}\rangle$. The elements of the V matrix can be easily written in terms of two-electron integrals

$$V_{\mathrm{HF},m\mu} = \langle \mu\mu | mm \rangle \tag{2.111}$$

$$V_{m\mu,n\nu} = \delta_{mn}\langle \nu\nu | \mu\mu \rangle + \delta_{\mu\nu}\langle mm | nn \rangle \tag{2.112}$$

whereas

$$\langle \phi_{\mathrm{HF}} | H^0 | \phi_{\mathrm{HF}} \rangle = E_{\mathrm{HF}} \tag{2.113}$$

and $\langle \phi_\mu^m | H^0 | \phi_\mu^m \rangle$ is the Hamiltonian expectation value for the doubly excited configuration $|\phi_\mu^m\rangle$. The matrix product $\mathsf{V}\mathbf{C}^{(n-1)}$ appearing in Eq. (2.109) can be written as follows:

$$\sum_{kv} V_{\mathrm{HF},kv} C_{kv}^{(n-1)} = \sum_{kv} \langle vv | kk \rangle C_{kv}^{(n-1)} \tag{2.114}$$

$$\sum_{kv} V_{m\mu,kv} C_{kv}^{(n-1)} = \sum_{v} \langle vv | \mu\mu \rangle C_{mv}^{(n-1)} + \sum_{k} \langle mm | kk \rangle C_{k\mu}^{(n-1)} \tag{2.115}$$

Using these results in Eq. (2.109) we obtain an explicit formula for the elements of $\mathbf{C}^{(n)}$:

$$C_{\mathrm{HF}}^{(n)} = \sum_{kv} (E - E_{\mathrm{HF}})^{-1} \langle vv | kk \rangle C_{kv}^{(n-1)} \tag{2.116}$$

$$C_{m\mu}^{(n)} = (E - \langle \phi_\mu^m | H^0 | \phi_\mu^m \rangle)^{-1} \left(\sum_{v} \langle vv | \mu\mu \rangle C_{mv}^{(n-1)} + \sum_{k} \langle mm | kk \rangle C_{k\mu}^{(n-1)} \right) \tag{2.117}$$

By writing out the elements of V and H^0 in terms of the integrals, we see that the iterative scheme for the evaluation of $\mathbf{C}$ and E can be written entirely in terms of sums over integrals and $\mathbf{C}^{(n-1)}$ and E values from the preceding iteration. This fact allows this perturbation scheme to be programmed on a computer in an integral-driven manner. That is, as the two-electron integrals $\langle ij | kl \rangle$ are brought into the core memory of the computer, all contributions of each successive integral to all of the sums appearing in Eqs. (2.116) and (2.117) can be evaluated, multiplied by appropriate factors, and added to the appropriate expansion coefficients. In this way, the computer is required to read through the (presumably long) list of two-electron integrals only once for each iteration. In this way, one avoids the explicit construction and storage of the Hamiltonian matrix, which may be very large and much larger than the number of two-electron integrals.

Techniques that permit the working numerical equations [e.g., Eqs. (2.116) and (2.117)] to be expressed as sums over explicit two-electron integrals are referred to as integral-driven direct CI methods. The perturbation solution described above is only a simple example of such methods. For more general classes of CI wavefunctions, the expressions for the V matrix elements are more involved. However, the basic structure and philosophy of the direct CI techniques remain as outlined. These techniques have proven to be quite useful in carrying out large-scale CI calculations, and such integral-driven strategies have been used to efficiently implement the graphical unitary group approach (GUGA) for electronic structure calculations (Paldus and Boyle, 1980; Shavitt, 1978; Brooks and Schaefer, 1979).

5. Size Consistency

Thus far, we have concentrated on describing how the CI procedure is used in practical applications and how it can be viewed as relating to the MCSCF method. It is important to realize that even though difficulties having to do with large CI matrices may be overcome, a serious problem remains inherent in nearly all of the above methods. To understand the difficulty, consider how one might perform a calculation of a potential-energy curve for the diatomic Be_2. Assume that a prior calculation on a single beryllium atom indicated that the $2s^2$ and $2p^2\ {}^1S$ configuration should be included in order to describe the electron correlation in beryllium. Then to describe the correlation in Be_2 in a balanced manner (i.e., such as to yield a $2s^2 + 2p^2$ level description of both beryllium atoms upon dissociation), one must include the $2s_A^2 2s_B^2$, $2s_A^2 2p_B^2$, $2p_A^2 2s_B^2$, and $2p_A^2 2p_B^2$ configurations, where A and B label the two beryllium nuclei. Hence, although a double-excitation CI or MCSCF could be employed for Be, one needs to include (certain) quadruple excitation (relative to $2s_A^2 2s_B^2$) for Be_2. Clearly, for more complex molecular clusters one would need to include even higher level excitations (e.g., eightfold for Be_4) to achieve a qualitatively balanced description of the complex and its fragments. This is, of course, essential if one is trying to compute energy changes (bond energies and energies of formation) for chemical reactions. Then one must use a method that yields the same value for the molecular complex energy (e.g., Be_2) when evaluated at large interfragment separation as the sum of the fragment (e.g., two beryllium atoms) energies evaluated separately within the same method. Such methods are said to be size consistent (Pople *et al.*, 1977). The use of a restricted CI or MCSCF wavefunction (e.g., doubly excited for Be_2) could indeed yield a smooth potential-energy curve free of obvious pathological behavior. However, such a wavefunction would preferentially describe the electron correlation in the complex (Be_2) near its equilibrium geometry and would dissociate to yield fragments that are described to a lower correlation level (e.g., the $2s_A^2 2s_B^2$ configuration would dominate).

The size consistency problem may be less significant if an appropriate configuration selection is performed at each geometry on the molecular potential surface, but the problem still remains as to how to efficiently choose configurations that describe equally well an entire potential energy surface. It may in fact be more straightforward to achieve this goal using an MCSCF wavefunction, since the orbital optimization thereby included can make the configuration expansion length short enough to be physically understood and hence correctly chosen. As we discuss in more detail in Chapter 3, this question relating to achieving a balanced description of a molecule and its fragments is important and not easily addressed within a variational framework.

6. Discussion

Because the CI technique has been the most widely used approach to treating electron correlation problems, many advances have been made in matrix storage techniques, two-electron integral transformations, the use of unitary group tools, matrix eigenvalue and eigenvector determinations, and configuration selection processes. We by no means intend to treat these advances here; many of them are reviewed well in Chapters 6–8 of Schaefer and Miller (1977). It is essential that one realize that the monumental development of exactly these same data management methods is what makes it possible to implement not only efficient CI computer programs but also highly efficient MCSCF, HF, coupled-cluster, and Green's function routines. To implement any of the above quantum-chemical methods in a state-of-the-art manner, one must make extensive use of many of the advances in numerical methods and data handling that the scientists who have been instrumental in developing efficient CI programs have made.

PROBLEMS

2.1 Using the one- and two-electron integrals given below, carry out an SCF calculation for the $1\sigma^2\mathrm{HeH}^+$ ground state using a first-order procedure.

1. By expanding the molecular orbitals $\{\phi_k\}$ as linear combinations of atomic orbitals $\{\chi_\mu\}$,

$$\phi_k = \sum_\mu c_{\mu k}\chi_\mu$$

and using the definition of the closed-shell Fock operator given in Eq. (2.92), show that the Fock eigenvalue equation can be written in terms of the atomic orbital basis as

$$\mathsf{F}\mathsf{c} = \mathsf{S}\mathsf{c}\boldsymbol{\varepsilon}$$

where the overlap matrix is

$$S_{\mu\nu} = \langle\mu|\nu\rangle$$

the elements of the Fock matrix are

$$F_{\mu\nu} = \langle\mu|h|\nu\rangle + \sum_{\rho\sigma} P_{\rho\sigma}\{2\langle\mu\rho|\nu\sigma\rangle - \langle\mu\rho|\sigma\nu\rangle\} \tag{A}$$

h is the one-electron operator in the Hamiltonian, and the charge bond order matrix P is defined as

$$P_{\rho\sigma} = \sum_k c^*_{\rho k}c_{\sigma k}$$

2. Show that the HF total energy for a closed-shell system may be written in terms of integrals over the orthonormal HF orbitals as

$$E(\mathrm{SCF}) = 2\sum_{k}^{\mathrm{occ}} \langle \phi_k|h|\phi_k\rangle + \sum_{kl}^{\mathrm{occ}} \{2\langle kl|kl\rangle - \langle kl|lk\rangle\} + \sum_{\mu>\nu} (Z_\mu Z_\nu/R_{\mu\nu}) \quad \text{(B)}$$

3. Show that the HF total energy may alternatively be expressed as

$$E(\mathrm{SCF}) = \sum_{k}^{\mathrm{occ}} \{\varepsilon_k + \langle \phi_k|h|\phi_k\rangle\} + \sum_{\mu>\nu} (Z_\mu Z_\nu/R_{\mu\nu}) \quad \text{(C)}$$

where the $\{\varepsilon_k\}$ refer to the HF orbital energies.

To carry out an SCF calculation on the ground state of HeH^+ at $R =$ 1.4 a.u., the following information is to be used. The orbital exponents of the 1s, Slater orbitals of the He and H are 1.6875 and 1.0, respectively. The atomic integrals required to carry out the HF calculation are (in a.u.)

$$S_{11} = S_{22} = 1.0, \qquad S_{12} = 0.5784$$

$$h_{11} = -2.6442, \qquad h_{22} = -1.7201, \qquad h_{12} = -1.5113, \qquad (h_{ij} \equiv \langle i|h|j\rangle),$$

$$\langle 11|11\rangle = 1.0547, \qquad \langle 11|21\rangle = 0.4744, \qquad \langle 12|12\rangle = 0.5664,$$

$$\langle 22|11\rangle = 0.2469, \qquad \langle 22|21\rangle = 0.3504, \qquad \langle 22|22\rangle = 0.6250$$

where 1 refers to $1s_{\mathrm{He}}$ and 2 to $1s_{\mathrm{H}}$. In this and the following problems we shall employ the indices 1 and 2 to label either the molecular orbitals or the atomic orbitals whenever doing so is not confusing. We shall reserve the notation 1σ and 2σ primarily for describing the orbital occupancies arising in the wavefunctions. As an initial guess for the occupied molecular orbital use $\phi_1 \simeq 1s_{\mathrm{He}}$.

4. Form, with this initial guess of the occupied molecular orbital, a 2×2 Fock matrix, using Eq. (A) for $F_{\mu\nu}$.

5. Solve the Fock matrix eigenvalue equations given above to obtain the orbital energies and an improved occupied molecular orbital. In so doing, note that the normalization condition $\langle \phi_1|\phi_1\rangle = 1 = \mathbf{C}_1^\dagger \mathbf{S} \mathbf{C}_1$ gives the needed normalization condition for the expansion coefficients of the ϕ_1 in the atomic orbital basis.

6. Determine the total SCF energy using Eq. (C) at this step of the iterative procedure. When will this energy agree with that obtained by using the alternative expression for $E(\mathrm{SCF})$ given in Eq. (B)?

7. Use the ϕ_1 molecular orbital from question 5 to determine a new Fock matrix.

8. Determine a new set of orbital energies and an improved occupied molecular orbital.

9. Determine from Eq. (C) the SCF total energy at this step of the iterative procedure.

The iterative process may be continued until convergence. As a convergence criterion, assume that the difference between the SCF total energy in two successive iterations must be less than 10^{-6} a.u. Listed below are the HF total energies (in a.u.) obtained during the iterative procedure beyond the two iterations performed above:

$$-2.842151, \quad -2.843221, \quad -2.843393,$$
$$-2.843420, \quad -2.843425, \quad -2.843425$$

10. Show, by comparing the difference between the SCF total energy at one iteration and the converged SCF total energy, that the convergence of the above SCF approach is linear or first order.

11. Is the SCF total energy listed above in each iteration of the SCF procedure an upper bound to the exact ground-state total energy?

The converged self-consistent set of molecular orbitals ϕ_1 and ϕ_2 is

$$\phi_1 = 0.9000\ 1s_{\text{He}} + 0.1584\ 1s_{\text{H}}, \qquad \phi_2 = -0.8324\ 1s_{\text{He}} + 1.2156\ 1s_{\text{H}}$$

12. Show, using the one- and two-electron integrals in the molecular orbital basis,

$$\langle 1|h|1\rangle = -2.6158, \qquad \langle 1|h|2\rangle = 0.1954, \qquad \langle 2|h|2\rangle = -1.3154$$
$$\langle 11|11\rangle = 0.9596, \qquad \langle 11|21\rangle = -0.1954, \qquad \langle 12|12\rangle = 0.6063,$$
$$\langle 12|21\rangle = 0.1261, \qquad \langle 22|21\rangle = -0.0045, \qquad \langle 22|22\rangle = 0.6159$$

hat the converged values of the orbital energies are

$$\varepsilon_1 = -1.6562, \qquad \varepsilon_2 = -0.2289$$

13. Does this SCF wavefunction give rise (at $R \to \infty$) to proper dissociaion products?

2.2 Now carry out an SCF calculation for the same closed-shell HeH$^+$ ystem using a second-order SCF procedure. Some of the integrals used in roblem 2.1 will be useful again here.

1. Show that the one- and two-electron density matrices decouple as ollows for a closed-shell reference state:

$$\sum_{\sigma} \langle 0|r_\sigma^+ s_\sigma|0\rangle = \delta_{rs} 2v_r, \qquad \sum_{\sigma\sigma'} \langle 0|r_\sigma^+ s_{\sigma'}^+ t_{\sigma'} u_\sigma|0\rangle = (4\delta_{ru}\delta_{st} - 2\delta_{rt}\delta_{su}) v_r v_s$$

here v_r is the occupation of orbital ϕ_r. That is, if ϕ_r is an occupied orbital $= 1$, and if ϕ_r is unoccupied $v_r = 0$.

2. Show that the A_{11}, B_{11}, and W matrices of Eqs. (2.29), (2.30), and (2.27), respectively, may be written for a closed-shell system as

$$\begin{aligned}(A_{11})_{n\beta,m\alpha} &= \langle 0|[\beta_\alpha^+ n_\alpha + \beta_\beta^+ \overset{m}{\cancel{n}}_\beta, H, m_\alpha^+ \alpha_\alpha + m_\beta^+ \alpha_\beta]|0\rangle \\ &= 2\Big[-h_{\alpha\beta}\delta_{mn} + h_{nm}\delta_{\alpha\beta} + \delta_{\alpha\beta}\sum_\gamma \{2\langle n\gamma|m\gamma\rangle - \langle n\gamma|\gamma m\rangle\} \\ &\quad - \delta_{mn}\sum_\gamma \{2\langle \alpha\gamma|\beta\gamma\rangle - \langle \alpha\gamma|\gamma\beta\rangle\} \\ &\quad + 2\langle n\alpha|\beta m\rangle - \langle n\alpha|m\beta\rangle\Big]\end{aligned}$$

$$\begin{aligned}(B_{11})_{n\beta,m\alpha} &= \langle 0|[\beta_\alpha^+ n_\alpha + \beta_\beta^+ n_\beta, H, \alpha_\alpha^+ m_\alpha + \alpha_\beta^+ m_\beta]|0\rangle \\ &= 2[\langle mn|\beta\alpha\rangle - 2\langle mn|\alpha\beta\rangle]\end{aligned}$$

$$\begin{aligned}W_{\alpha m} &= \langle 0|[\alpha_\alpha^+ m_\alpha + \alpha_\beta^+ m_\beta, H]|0\rangle = 2\Big[h_{m\alpha} \\ &\quad + \sum_\gamma \{2\langle m\gamma|\alpha\gamma\rangle - \langle m\gamma|\gamma\alpha\rangle\}\Big]\end{aligned}$$

Again use as the initial guess of the occupied molecular orbital $1s_{\mathrm{He}}$.

3. Given this guess for ϕ_1, determine the virtual or unoccupied molecular orbital ϕ_2 using a Schmidt orthogonalization procedure. The atomic integrals required are given in Problem 2.1.

The second-order SCF procedure requires knowledge of the integrals in the basis of the set of initial orthonormal molecular orbitals (ϕ_1 and ϕ_2 obtained above). The one- and two-electron integrals in this basis are given below (in a.u.):

$$\begin{array}{lll} h_{11} = -2.6442, & h_{22} = -1.2870, & h_{12} = 0.0223 \\ \langle 11|11\rangle = 1.0547, & \langle 11|21\rangle = -0.1663, & \langle 12|12\rangle = 0.5567, \\ \langle 22|11\rangle = 0.0765, & \langle 22|21\rangle = 0.0171, & \langle 22|22\rangle = 0.6200 \end{array}$$

where, as before, 1 denotes the occupied and 2 the unoccupied molecular orbital.

4. Determine the SCF total energy that corresponds to this initial guess of molecular orbitals.

5. Determine the A_{11}, B_{11}, and W matrix elements.

6. Determine the $\boldsymbol{\kappa}$ matrix and the unitary matrix $\mathsf{X} = e^{-\boldsymbol{\kappa}}$.

7. Determine the new improved set of orthonomal molecular orbitals resulting from applying X to ϕ_1 and ϕ_2.

The one- and two-electron integrals may now be evaluated in the set of improved molecular orbitals and the iterative procedure thus may be continued until convergence is obtained. The HF total energies obtained during the iterative procedure become

$$-2.80504513, \quad -2.84303574, \quad -2.84342526, \quad -2.84342527$$

8. Show by comparing the difference between the SCF total energies at successive iterations and the converged SCF total energy that the convergence of the above SCF approach is quadratic or second order.

2.3 Given the one- and two-electron integrals in the SCF orbital basis found in Problem 2.1, carry out a two-configuration CI calculation on HeH$^+$ using the $1\sigma^2$ and $2\sigma^2$ configurations.

1. First obtain expressions for the CI matrix elements $H_{ij}\,(i, j = 1\sigma^2, 2\sigma^2)$ in terms of one- and two-electron integrals.

2. Show that the resultant CI matrix is (ignoring the nuclear repulsion term)

$$\begin{pmatrix} -4.2720 & 0.1261 \\ 0.1261 & -2.0149 \end{pmatrix}$$

3. Obtain the two CI energies and eigenvectors for the above matrix.

4. Show that the lowest-energy CI wavefunction is equivalent to the following two-determinant (single configuration) wavefunction:

$$\tfrac{1}{2}\big[|(a^{1/2}\phi_1 + b^{1/2}\phi_2)\alpha(a^{1/2}\phi_1 - b^{1/2}\phi_2)\beta| \\ + |(a^{1/2}\phi_1 - b^{1/2}\phi_2)\alpha(a^{1/2}\phi_1 + b^{1/2}\phi_2)\beta|\big]$$

involving the polarized orbitals $a^{1/2}\phi_1 \pm b^{1/2}\phi_2$, where $a = 0.9984$ and $b = 0.0556$.

2.4 Using the same information as in Problem 2.3, carry out a three-configuration CI calculation on HeH$^+$ at $R = 1.4$ a.u. using the $1\sigma^2$, $2\sigma^2$, and $1\sigma 2\sigma$ electronic configurations.

1. First express the proper singlet spin-coupled $1\sigma 2\sigma$ configuration as a combination of Slater determinants.

2. Compute all elements of the 3 × 3 CI matrix.

3. Obtain the eigenenergies and corresponding normalized eigenvectors for this problem.

2.5 Use the perturbative method described in Section D.4 on the CI matrix eigenvalue problem of Problem 2.4 to find the lowest eigenenergy and its corresponding eigenvalues. Use as the initial guess for the eigenvector $\mathbf{C}^0 = (1.0000, 0.0, 0.0)$ and take

$$\mathsf{H}^0 = \begin{pmatrix} -4.2720 & 0 & 0 \\ 0 & -2.0149 & 0 \\ 0 & 0 & -3.1988 \end{pmatrix}$$

and $E^0 = -4.2720$ for the first iteration. Use the energy computed using Eq. (2.110) to start the second iteration, but notice that the $\mathbf{C}^{(1)}$ vector you then obtain is so much different from $\mathbf{C}^{(0)}$ that convergence of the process is not likely. Therefore, average these $\mathbf{C}^{(0)}$ and $\mathbf{C}^{(1)}$ to obtain a new damped $\mathbf{C}^{(1)}$ for use in determining $\mathbf{C}^{(2)}$.

2.6 Perform a one-step second-order multiconfiguration HF calculation on HeH$^+$, using the minimum Slater basis of Problem 2.1. The multiconfiguration reference state will include the two configurations $1\sigma^2$ and $2\sigma^2$. As an initial guess of orbitals use the set of single-configuration HF orbitals of the principal configuration $1\sigma^2$. The HF orbitals were determined in Problem 2.1, and the one- and two-electron integrals in the HF basis are given there. The initial guess of the configuration state functions (denoted $|0\rangle$ and $|1\rangle$) will be the ones determined in the two-configuration CI calculation given in Problem 2.3.

1. Determine all of the nonvanishing one- and two-electron density matrix elements

$$\langle 0|r^\dagger s|0\rangle, \langle 0|r^\dagger s^\dagger tu|0\rangle$$

and the nonvanishing one- and two-electron transition density matrix elements

$$\langle 1|r^\dagger s|0\rangle, \langle 1|r^\dagger s^\dagger tu|0\rangle$$

2. Determine the $\mathbf{V}$, $\mathbf{W}$, A, and B matrix elements.

3. Determine the $\boldsymbol{\kappa}$ and the $\mathbf{P}$ matrix elements via the one-step second-order MCSCF method.

4. Determine the transformed set of orbitals and states ($|0\rangle$ and $|1\rangle$).

5. Discuss whether the orbitals and states obtained after the first iteration of the one-step second-order MCSCF procedure (question 4) differ from the orbitals and states that would be obtained after the first iteration of the two-step second order MCSCF procedure. If they differ, describe how they would be obtained in the two-step procedure.

From the orbitals and states obtained in question 4 new one- and two-electron integrals and one- and two-electron density and transition density matrix elements may now be evaluated, and the iterative procedure thus continued. The multiconfigurational HF total energies obtained during this iterative procedure are

$$-2.85044942, \quad -2.85066435, \quad -2.85066436$$

6. Show by comparing the difference between the MCSCF total energies at each iteration and the converged MCSCF total energy that the convergence rate of the used MCSCF approach is second order.

7. How must the converged MCSCF ground-state total energy compare with the ground-state total energy obtained in the full CI calculation?

8. Write a priori the ground-state total energy that would be obtained if we used the three configurations $1\sigma^2$, $2\sigma^2$, and $1\sigma^1 2\sigma^1$ in an MCSCF calculation.

9. Write a priori the ground-state total energy that would be obtained from a converged two-configuration MCSCF calculation that used the $|1\sigma^2\rangle$ and $|1\sigma 2\sigma\rangle$ configurations.

2.7 Consider n HeH$^+$ molecular ions, which do not interact because they are infinitely far from one another.

1. Write the electronic Hamiltonian for this system in a basis consisting of orthonormal orbitals that are localized on each of the HeH$^+$ molecules. Retain only those contributions that are nonzero. In so doing, describe each HeH$^+$ molecule with a bonding and antibonding SCF orbital pair.

2. Show that a CI calculation that includes the HF ground-state wavefunction consisting of the antisymmetrized product of orbitals localized on the n ions having $1\sigma^2$ occupancy, and all doubly excited configurations leads to the following CI matrix of dimension $n + 1$:

$$\begin{pmatrix} nE_{\text{HF}} & B & B & B & \cdots & B \\ B & C & 0 & 0 & & \\ B & 0 & C & 0 & & \\ B & 0 & 0 & C & & \\ \vdots & & & & \ddots & \\ B & & & & & C \end{pmatrix}$$

where

$$C = E^*_{\text{HF}} - E_{\text{HF}} + nE_{\text{HF}}, \qquad B = \langle 11|22\rangle = \langle 1\sigma^2|H|2\sigma^2\rangle$$

$$E_{\text{HF}} = 2h_{11} + \langle 11|11\rangle = \langle 1\sigma^2|H|1\sigma^2\rangle,$$

$$E^*_{\text{HF}} = 2h_{22} + \langle 22|22\rangle = \langle 2\sigma^2|H|2\sigma^2\rangle$$

As in other problems, 1 and 2 denote the bonding and antibonding SCF molecular orbitals, respectively, for an isolated HeH$^+$ molecule.

3. Show that the correlation energy for n infinitely separated HeH$^+$ molecules is

$$E_{\text{corr}} = \frac{-E_{\text{HF}} + E^*_{\text{HF}}}{2} - \left[\frac{(-E_{\text{HF}} + E^*_{\text{HF}})^2}{4} + n\langle 11|22\rangle^2\right]^{1/2}$$

4. Use the HeH$^+$ SCF orbitals and results from Problems 2.1 and 2.3 to evaluate for $n = 2$, 4, 10, 100, and 1000 the correlation energy obtained for n infinitely separated HeH$^+$ molecules. Show that the correlation energy increases as $n^{1/2}$ when n becomes large. How would the correlation energy increase in a size-consistent model?

Bartlett and Purvis (1981) have found that for H_2 and He the percentage errors caused by size inconsistency in double-excitation CI calculation vary as follows:

n	Error H_2 (%)	Error He (%)
2	1.5	0.8
4	4.8	2.4
10	12.3	6.5
100	48.0	34.8
1000	79.1	70.8

5. Argue why the two-basis function HeH^+ problem is likely to underestimate the non-size-consistent contributions when compared with results obtained in more accurate calculations on HeH^+.

SOLUTIONS

2.1

1. $V_{ij} = \sum_{\gamma} (2\langle i\gamma | j\gamma\rangle - \langle i\gamma | \gamma j\rangle)$

Let

$$\phi_i = \sum_{\mu} C_{\mu i}\chi_{\mu}, \qquad \phi_{\gamma} = \sum_{\nu} C_{\nu\gamma}\chi_{\nu}$$

Then

$$\begin{aligned} V_{ij} &= \sum_{\gamma\mu'\mu\nu\nu'} (C_{\nu\gamma}C_{\nu'\gamma})(C_{\mu i}C_{\mu' j})(2\langle \mu\nu | \mu'\nu'\rangle - \langle \mu\nu | \nu'\mu'\rangle) \\ &= \sum_{\mu\mu'} C_{\mu i}C_{\mu' j}V_{\mu\mu'} \end{aligned}$$

where

$$V_{\mu\mu'} = \sum_{\nu\nu'} P_{\nu\nu'}(2\langle \mu\nu | \mu'\nu'\rangle - \langle \mu\nu | \nu'\mu'\rangle), \qquad P_{\nu\nu'} = \sum_{\gamma} C_{\nu\gamma}C_{\nu'\gamma}$$

Likewise

$$\langle \phi_i | -\tfrac{1}{2}\nabla^2 - \sum_{A}(Z_A/|\mathbf{r} - \mathbf{R}_A|) | \phi_j\rangle \equiv h_{ij} = \sum_{\mu\mu'} C_{\mu i}C_{\mu' j}h_{\mu\mu'}$$

$$h_{\mu\mu'} = \langle \chi_{\mu} | -\tfrac{1}{2}\nabla^2 - \sum_{A}(Z_A/|\mathbf{r} - \mathbf{R}_A|) | \chi_{\mu'}\rangle$$

As a result $F\phi_i = \varepsilon_i\phi_i$ can, by expanding ϕ_i as above, be expressed as

$$h_{ij} + V_{ij} = \delta_{ij}\varepsilon_i = \sum_{\mu\mu'} C_{\mu i}C_{\mu' j}(h_{\mu\mu'} + V_{\mu\mu'})$$

Then using

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} = \sum_{\mu\mu'} C_{\mu i} S_{\mu\mu'} C_{\mu' j}$$

we have

$$\sum_{\mu\mu'} C_{\mu i} [\varepsilon_j S_{\mu\mu'} - h_{\mu\mu'} - V_{\mu\mu'}] C_{\mu' j} = 0, \qquad \text{for all} \quad i, j$$

This can only be true if

$$\sum_{\mu'} (h_{\mu\mu'} + V_{\mu\mu'} - \varepsilon_j S_{\mu\mu'}) C_{\mu' j} = 0$$

This is $\mathsf{FC} = \mathsf{SC}\boldsymbol{\varepsilon}$.

2. The Slater–Condon rules tell us that the Hamiltonian expectation value for a single Slater determinant in which spin orbitals $\phi_1, \ldots, \phi_N$ are occupied is

$$E = \sum_{k=1}^{N} \langle \phi_k | -\tfrac{1}{2}\nabla^2 - \sum_A (Z_A / |\mathbf{r} - \mathbf{R}_A|) | \phi_k \rangle + \frac{1}{2} \sum_{k,l} [\langle kl | kl \rangle - \langle kl | lk \rangle]$$

For a closed-shell system the orbitals are doubly occupied and therefore $\phi_1 = \bar{\Phi}_1 \alpha$, $\phi_2 = \bar{\Phi}_1 \beta$, $\phi_3 = \bar{\Phi}_2 \alpha$, $\phi_4 = \bar{\Phi}_2 \beta$, etc., where $\bar{\Phi}_1$, $\bar{\Phi}_2$, etc. label the occupied orbitals (not spin-orbitals). Hence by carrying out the spin integration in the above energy expression and using the fact that each orbital is doubly occupied, we obtain

$$E = 2 \sum_{k}^{\text{occ}} \langle \phi_k | h | \phi_k \rangle + \sum_{kl}^{\text{occ}} \{ 2\langle kl | kl \rangle - \langle kl | lk \rangle \}$$

where labels now refer to orbital index. The term $\sum_{\mu > \nu} (Z_\mu Z_\nu / R_{\mu\nu})$ must then be added on to obtain the total energy (including nuclear repulsion).

3. If the occupied orbitals ϕ_k obey $F\phi_k = \varepsilon_k \phi_k$ then the above expression for E can be rearranged to give

$$E = \sum_{k}^{\text{occ}} \{ \langle \phi_k | h | \phi_k \rangle + \sum_{l}^{\text{occ}} [2\langle kl | kl \rangle - \langle kl | lk \rangle] \} + \sum_{k}^{\text{occ}} \langle \phi_k | h | \phi_k \rangle$$

The first two terms in this expression can be recognized as $\langle \phi_k | F | \phi_k \rangle$, where F is the closed-shell Fock operator whose potential is defined in Eq. (2.92). Hence

$$E = \sum_{k}^{\text{occ}} \langle \phi_k | F | \phi_k \rangle + \sum_{k}^{\text{occ}} \langle \phi_k | h | \phi_k \rangle$$

4. $\mathsf{P} = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 0.0 \end{pmatrix}, \qquad \mathsf{F} = \begin{pmatrix} -1.5895 & -1.0369 \\ -1.0369 & -0.8342 \end{pmatrix}$

5. $\varepsilon_1 = -1.6048$, $\varepsilon_2 = -0.2348$, $\phi_1 = 0.9194\ 1s_{\mathrm{He}} + 0.1296\ 1s_{\mathrm{H}}$

6. $E_{\mathrm{SCF}} = -2.8005$. The two expressions will agree only upon convergence of the SCF when $F\phi_i = \varepsilon_i \phi_i$, which was assumed in writing the expression for E containing the orbital energies.

7. $\mathsf{P} = \begin{pmatrix} 0.8453 & 0.1192 \\ 0.1192 & 0.0168 \end{pmatrix}$, $\mathsf{F} = \begin{pmatrix} -1.6246 & -1.0836 \\ -1.0836 & -0.8772 \end{pmatrix}$

8. $\varepsilon_1 = -1.6469$, $\varepsilon_2 = -0.2289$, $\phi_1 = 0.9032\ 1s_{\mathrm{He}} + 0.1537\ 1s_{\mathrm{H}}$

9. $E_{\mathrm{SCF}} = -2.8356$

10.

$E_{\mathrm{SCF}} - E_{\mathrm{SCF}}^{\mathrm{conv}}$	$(E_{\mathrm{SCF}} - E_{\mathrm{SCF}}^{\mathrm{conv}})^2$
0.001274	0.000002
0.000204	0.000000
0.000032	
0.000005	
0.000000	

Second-order convergence requires that the error in the $(n + 1)$th iteration is the square of the error in the nth iteration, In the first iteration above the error is 0.001274; thus in the next iteration the error should be $(0.001274)^2 = 0.0000016$ if we used a second-order procedure. Since the second iteration's error is 0.000204, the convergency of the above SCF procedure is linear rather than quadratic.

11. The *converged* SCF total energy calculated from Eq. (C) is an upper bound to the ground-state energy, whereas the SCF total energy from Eq. (C) during the iterative procedure is not a bound. It is only at convergence that the expectation value of the Hamiltonian for the HF determinant is given by Eq. (C).

12. The SCF orbital energies are determined to be

$$\varepsilon_k = \langle k|h|k\rangle + \sum_l \{2\langle kl|kl\rangle - \langle kl|lk\rangle\}$$

from which the orbital energies follow straightforwardly.

13. Yes, the $1\sigma^2$ configuration does dissociate properly because at $R \to \infty$, the lowest-energy state is He + H^+, which also has a $1\sigma^2$ orbital occupancy.

2.2

1. Since ϕ_r and ϕ_s are either occupied or unoccupied $\sum_\sigma \langle 0|r_\sigma^+ s_\sigma|0\rangle$ vanishes unless both ϕ_r and ϕ_s are in $|0\rangle$. Hence $\sum_\sigma \langle 0|r_\sigma^+ s_\sigma|0\rangle = \delta_{rs} 2v_r$. Likewise, in $\sum_{\sigma\sigma'} \langle 0|r_\sigma^+ s_{\sigma'}^+ t_{\sigma'} u_\sigma|0\rangle$ all four spin-orbitals must be in $|0\rangle$. Then

$$\begin{aligned} \langle 0|r_\sigma^+ s_{\sigma'}^+ t_{\sigma'} u_\sigma|0\rangle &= \delta_{st}\langle 0|r_\sigma^+ u_\sigma|0\rangle - \langle 0|r_\sigma^+ t_{\sigma'} s_{\sigma'}^+ u_\sigma|0\rangle \\ &= \delta_{st}\delta_{ru} - \delta_{su}\delta_{\sigma\sigma'}\langle 0|r_\sigma^+ t_{\sigma'}|0\rangle + \langle 0|r_\sigma^+ t_{\sigma'} u_\sigma s_{\sigma'}^+|0\rangle \\ &= \delta_{st}\delta_{ru} - \delta_{su}\delta_{rt}\delta_{\sigma\sigma'} \end{aligned}$$

where all orbitals are occupied. Clearly $\sum_{\sigma\sigma'}\langle 0|r_\sigma^+ s_{\sigma'}^+ t_{\sigma'} u_\sigma|0\rangle$ vanishes when one or more of the four orbitals are unoccupied, and its equals $4\delta_{st}\delta_{ru} - 2\delta_{su}\delta_{rt}$ when all four orbitals are occupied.

2. Equation (2.42), when combined with the results of question 1, give

$$(A_{11})_{n\beta,m\alpha} = 2\delta_{\alpha\beta}h_{nm} + 0 - \delta_{nm}2h_{\alpha\beta} - 0 - 0 - \delta_{nm}\sum_\gamma[\langle\alpha\gamma|\beta\gamma\rangle 4 - 2\langle\alpha\gamma|\gamma\beta\rangle]$$

$$- 0 - 0 - \sum_\gamma\langle n\gamma|\gamma m\rangle\delta_{\alpha\beta}2 + 4\langle n\alpha|\beta m\rangle$$

$$+ \delta_{\alpha\beta}\sum_\gamma\langle n\gamma|m\gamma\rangle 4 - 2\langle n\alpha|m\beta\rangle + 0 + 0$$

$$(B_{11})_{n\beta,m\alpha} = 0 + 0 - 0 - 0 - 0 - 0 - 0 - 4\langle mn|\alpha\beta\rangle + 2\langle mn|\beta\alpha\rangle$$

$$+ 0 + 0 + 0 + 0$$

Using Eq. (2.41) we find

$$W_{\alpha m} = 2h_{m\alpha} - 0 - 0 + \sum_\gamma[4\langle m\gamma|\alpha\gamma\rangle - 2\langle m\gamma|\gamma\alpha\rangle]$$

3. $\phi_1 = 1s_{\mathrm{He}}$

$\phi_2 = 1s_{\mathrm{H}} - \langle 1s_{\mathrm{H}}|1s_{\mathrm{He}}\rangle 1s_{\mathrm{He}} = 1s_{\mathrm{H}} - 0.5784\ 1s_{\mathrm{He}}$

Normalizing ϕ_2 requires that we divide by the square root of $1 + \langle 1s_{\mathrm{H}}|1s_{\mathrm{He}}\rangle^2 - 2\langle 1s_{\mathrm{H}}|1s_{\mathrm{He}}\rangle^2$ to obtain

$$\phi_2 = 1.2259\ 1s_{\mathrm{H}} - 0.7091\ 1s_{\mathrm{He}}$$

The coefficient matrix C whose elements $C_{\mu k}$ are the orbital expansion coefficients then becomes

$$\mathsf{C} = \begin{pmatrix} 1.0 & -0.7091 \\ 0.0 & 1.2259 \end{pmatrix}$$

4. -2.8050

5. The excitation from molecular orbital 1 to 2 is nonredundant. The relevant matrices thus become one dimensional:

$$A_{11} = 1.8713, \qquad B_{11} = -0.1530, \qquad W = -0.2880$$

6. $\boldsymbol{\kappa} = \begin{pmatrix} 0.0 & 0.1423 \\ -0.1423 & 0.0 \end{pmatrix}, \qquad \mathsf{X} = \begin{pmatrix} 0.9899 & -0.1418 \\ 0.1418 & 0.9899 \end{pmatrix}$

7. $\mathsf{C} = \begin{pmatrix} 0.8893 & -0.8437 \\ 0.1739 & 1.2135 \end{pmatrix}$

8.

$E_{\mathrm{SCF}} - E_{\mathrm{SCF}}^{\mathrm{conv}}$	$(E_{\mathrm{SCF}} - E_{\mathrm{SCF}}^{\mathrm{conv}})^2$
0.03838013	0.00147303
0.00038952	0.00000015
0.00000001	

The error in the $(n+1)$th iteration is the square of the error in the nth iteration.

2.3

1. $\langle 1\sigma^2|H|1\sigma^2\rangle = 2h_{11} + \langle 1\,1|1\,1\rangle$, $\quad \langle 2\sigma^2|H|2\sigma^2\rangle = 2h_{22} + \langle 2\,2|2\,2\rangle$, $\quad \langle 1\sigma^2|H|2\sigma^2\rangle = \langle 1\,1|2\,2\rangle$

2. $\langle 1\sigma^2|H|1\sigma^2\rangle = (-2.6158)2 + 0.9596 = -4.2720$
$\langle 2\sigma^2|H|2\sigma^2\rangle = (-1.3154)2 + 0.6159 = -2.0149$
$\langle 1\sigma^2|H|2\sigma^2\rangle = 0.1261$

3. $E_- = -4.2790$, $\quad E_+ = -2.0079$, $\quad \mathbf{C}_- = (0.9984, -0.0556)$, $\mathbf{C}_+ = (0.0556, 0.9984)$

4. $$\begin{aligned} &\tfrac{1}{2}\big[|(a^{1/2}\phi_1 + b^{1/2}\phi_2)\alpha(a^{1/2}\phi_1 - b^{1/2}\phi_2)\beta| \\ &\quad + |(a^{1/2}\phi_1 - b^{1/2}\phi_2)\alpha(a^{1/2}\phi_1 + b^{1/2}\phi_2)\beta|\big] \\ &= \frac{1}{2\sqrt{2}}\big[(a^{1/2}\phi_1 + b^{1/2}\phi_2)(a^{1/2}\phi_1 - b^{1/2}\phi_2) \\ &\quad + (a^{1/2}\phi_1 - b^{1/2}\phi_2)(a^{1/2}\phi_1 + b^{1/2}\phi_2)\big](\alpha\beta - \beta\alpha) \\ &= (a\phi_1\phi_1 - b\phi_2\phi_2)(\alpha\beta - \beta\alpha)/\sqrt{2} \\ &= a|\phi_1\alpha\phi_1\beta| - b|\phi_2\alpha\phi_2\beta| \end{aligned}$$

2.4

1. The singlet function with $1\sigma 2\sigma$ occupancy is
$$\frac{1}{\sqrt{2}}\big[|1\,\alpha 2\,\beta| - |1\,\beta 2\,\alpha|\big]$$

2. $$\langle 1\sigma 2\sigma|H|1\sigma^2\rangle = \frac{1}{\sqrt{2}}[2h_{12} + 2\langle 11|21\rangle] = \sqrt{2}[0.1954 - 0.1954] = 0$$
$$\begin{aligned} \langle 1\sigma 2\sigma|H|2\sigma^2\rangle &= \frac{1}{\sqrt{2}}[2h_{12} + 2\langle 22|21\rangle] \\ &= \sqrt{2}[0.1954 - 0.0045] = 0.2699 \end{aligned}$$
$$\begin{aligned} \langle 1\sigma 2\sigma|H|1\sigma 2\sigma\rangle &= h_{11} + h_{22} + \langle 12|12\rangle + \langle 12|21\rangle \\ &= -2.6158 - 1.3154 + 0.6063 + 0.1261 = -3.1988 \end{aligned}$$

The other matrix elements are derived in Problem 2.3.

3. The 3 × 3 CI matrix

$$\begin{pmatrix} -4.2720 & 0.1261 & 0.000 \\ & -2.0149 & 0.2699 \\ & & -3.1988 \end{pmatrix}$$

then has as its eigenvalues $E_1 = -4.2792$, $E_2 = -3.2567$, and $E_3 = -1.9497$. The corresponding eigenvectors have, as coefficients of $|1\sigma^2|$, $|2\sigma^2|$, and $|1\sigma 2\sigma|$, (0.9982, −0.0573, 0.0143), (−0.0261, −0.2098, 0.9772), and (0.0530, 0.9761, −0.2109), respectively. The ground-state total energy, including nuclear repulsion, is $-4.2792 + (2/1.4) = -2.8506$.

2.5 First iteration:

$$E^0 - E = -(\mathbf{C}^0)^{\mathrm{T}}\mathsf{V}\mathbf{C}^0 = 0$$

and so we use $E = -4.2720$ in the first iteration to calculate $\mathbf{C}^{(1)}$:

$$\mathbf{C}^{(1)} = (E - \mathsf{H}^0)^{-1}\mathsf{V}\mathbf{C}^0, \qquad (E\mathsf{1} - \mathsf{H}^0)^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.4430 & 0 \\ 0 & 0 & -0.9318 \end{pmatrix}$$

(Actually, the 1,1 element of this matrix is ill defined. However, this does not cause trouble here since $\mathsf{V}\mathbf{C}^0$ has zero as its first entry. In general, however, this trouble arises wherever E^0 is taken to a diagonal element of H^0):

$$\mathsf{V}\mathbf{C}^0 = \begin{pmatrix} 0.0 \\ 0.1261 \\ 0.0 \end{pmatrix}$$

Therefore,

$$\mathbf{C}^{(1)} = \begin{pmatrix} 0.0 \\ -0.0559 \\ 0.0 \end{pmatrix}$$

A problem now arises: $\mathbf{C}^{(1)}$ does not obey intermediate normalization, which was assumed in deriving Eq. (2.110) for E. Hence we must damp the iteration process by averaging $\mathbf{C}^{(1)}$ and $\mathbf{C}^0$ to obtain a better $\mathbf{C}^{(1)}$ (which must then be intermediate normalized):

$$\mathbf{C}^{(1)} = \frac{1}{2}\begin{pmatrix} 1.0000 \\ 0.0 \\ 0.0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0.0 \\ -0.0559 \\ 0.0 \end{pmatrix} \rightarrow \begin{pmatrix} 1.0000 \\ -0.0559 \\ 0.0 \end{pmatrix}$$

One finds that by not damping, the successive $\mathbf{C}^{(n)}$ computed (using $E = -4.2790$) are wildly oscillating.

$C^{(n)} = [(0.9989, 0, 0.0140), (0.0, -0.0573, -0.0035), (1.0000, 0.0004, 0.0139),$
$(1.000, 0.0797, 0.0001), (1.0000, 0.0386, 0.0138)]$

However, once the damping is introduced, we can proceed to find $\mathbf{C}^{(n)}$ values in a stable manner.

$$E^0 - E = -(\mathbf{C}^0)^+ \mathsf{V}\mathbf{C} = -(0, 0.1261, 0)\begin{pmatrix} 1.0000 \\ -0.0559 \\ 0.0000 \end{pmatrix} = 0.0070$$

and so $E = E^0 - 0.0070 = -4.2790$. Now compute $\mathbf{C}^{(2)}$:

$$(E\mathsf{1} - \mathsf{H}^0)^{-1} = \begin{pmatrix} -142.9 & 0 & 0 \\ 0 & -0.4417 & 0 \\ 0 & 0 & -0.9258 \end{pmatrix}$$

and

$$\mathbf{C}^{(2)} = \begin{pmatrix} -142.9 & 0 & 0 \\ 0 & -0.4417 & 0 \\ 0 & 0 & -0.9258 \end{pmatrix}\begin{pmatrix} 0 & 0.1261 & 0 \\ 0.1261 & 0 & 0.2699 \\ 0 & 0.2699 & 0 \end{pmatrix}\begin{pmatrix} 1.0000 \\ -0.0559 \\ 0.0000 \end{pmatrix}$$

Then

$$\mathbf{C}^{(2)} = \begin{pmatrix} 1.0003 \\ -0.0557 \\ 0.0140 \end{pmatrix}$$

which upon intermediate normalization becomes

$$\mathbf{C}^{(2)} = \begin{pmatrix} 1.0000 \\ -0.0557 \\ 0.0140 \end{pmatrix}$$

For the third iteration,

$$E^0 - E = -(0, 0.1261, 0)\begin{pmatrix} 1.0000 \\ -0.0557 \\ 0.0140 \end{pmatrix} = 0.0070$$

Therefore, $E = -4.2790$ and so $(\mathrm{E}\mathsf{1} - \mathsf{H}^0)^{-1}$ is unchanged:

$$\mathbf{C}^{(3)} = (E\mathsf{1} - \mathsf{H}^0)^{-1}\begin{pmatrix} 0 & 0.1261 & 0 \\ 0.1261 & 0 & 0.2699 \\ 0 & 0.2699 & 0 \end{pmatrix}\begin{pmatrix} 1.0000 \\ -0.0557 \\ 0.0140 \end{pmatrix} = \begin{pmatrix} 1.0003 \\ -0.0574 \\ 0.0139 \end{pmatrix}$$

which when renormalized becomes

$$\begin{pmatrix} 1.0000 \\ -0.0574 \\ 0.0139 \end{pmatrix}$$

To compare these results with those of the CI Problem 2.4, we must normalize $\mathbf{C}^{(3)}$ so that $1 = (\mathbf{C}^{(3)})^{\mathrm{T}}(\mathbf{C}^{(3)})$. This procedure gives

$$\mathbf{C}^{(3)} = \begin{pmatrix} 0.9983 \\ -0.0573 \\ 0.0139 \end{pmatrix}$$

which is to be compared with the CI eigenvector

$$\mathbf{C}_{\mathrm{CI}} = \begin{pmatrix} 0.9982 \\ -0.0573 \\ 0.0143 \end{pmatrix}$$

The CI energy -4.2792 compares well with our third iterate $E = -4.2790$.

2.6

1. Let 1 and 2 denote the 1σ and 2σ orbitals, respectively:

	$\langle i\vert = \langle 0\vert$	$\langle i\vert = \langle 1\vert$
$\langle i\vert 1_\alpha^+ 1_\alpha\vert 0\rangle$	0.9970	0.0555
$\langle i\vert 2_\alpha^+ 2_\alpha\vert 0\rangle$	0.0031	-0.0555
$\langle i\vert 1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha\vert 0\rangle$	0.9970	0.0555
$\langle i\vert 1_\alpha^+ 1_\beta^+ 2_\beta 2_\alpha\vert 0\rangle$	-0.0555	-0.0031
$\langle i\vert 2_\alpha^+ 2_\beta^+ 1_\beta 1_\alpha\vert 0\rangle$	-0.0555	0.9970
$\langle i\vert 2_\alpha^+ 2_\beta^+ 2_\beta 2_\alpha\vert 0\rangle$	0.0031	-0.0555

Since the states $|0\rangle$, $|1\rangle$ have singlet symmetry, interchange of α and β spin gives the same matrix elements, e.g., $\langle i|1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha|0\rangle = \langle i|1_\beta^+ 1_\alpha^+ 1_\alpha 1_\beta|0\rangle$. Change of the sequence of the creation (or the annihilation) operators does, of course, change the sign of the matrix elements.

2. $V_{|1\rangle} = \langle 1|H|0\rangle = 0$, since $|1\rangle$ and $|0\rangle$ are determined from a CI calculation:

$$\begin{aligned} W_{12} = \langle 0|[1_\alpha^+ 2_\alpha + 1_\beta^+ 2_\beta, H]|0\rangle &= 2[h_{12}\langle 0|1_\alpha^+ 1_\alpha - 2_\alpha^+ 2_\alpha|0\rangle \\ &\quad + \langle 22|21\rangle\langle 0|2_\beta^+ 2_\alpha^+ 2_\beta 2_\alpha + 1_\alpha^+ 1_\beta^+ 2_\beta 2_\alpha|0\rangle \\ &\quad + \langle 11|21\rangle\langle 0|1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha + 1_\beta^+ 1_\alpha^+ 2_\beta 2_\alpha|0\rangle] \\ &= -0.0224 \end{aligned}$$

$$\begin{aligned}A_{21,21} &= \langle 0|[1_\alpha^+ 2_\alpha + 1_\beta^+ 2_\beta, H, 2_\alpha^+ 1_\alpha + 2_\beta^+ 1_\beta]|0\rangle \\ &= 2[(h_{22} - h_{11})\langle 0|1_\alpha^+ 1_\alpha - 2_\alpha^+ 2_\alpha|0\rangle - \langle 22|22\rangle\langle 0|2_\beta^+ 2_\alpha^+ 2_\alpha 2_\beta|0\rangle \\ &\quad - \langle 11|11\rangle\langle 0|1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha|0\rangle - 2\langle 11|22\rangle\langle 0|1_\beta^+ 1_\alpha^+ 2_\alpha 2_\beta|0\rangle \\ &\quad - \langle 21|12\rangle\langle 0|1_\beta^+ 1_\alpha^+ 1_\beta 1_\alpha + 2_\beta^+ 2_\alpha^+ 2_\beta 2_\alpha|0\rangle \\ &\quad + \langle 21|21\rangle\langle 0|1_\beta^+ 1_\alpha^+ 1_\alpha 1_\beta + 2_\alpha^+ 2_\beta^+ 2_\beta 2_\alpha|0\rangle] \\ &= 2.1606\end{aligned}$$

$$\begin{aligned}B_{21,21} &= \langle 0|[1_\alpha^+ 2_\alpha + 1_\beta^+ 2_\beta, H, 1_\alpha^+ 2_\alpha + 1_\beta^+ 2_\beta]|0\rangle \\ &= 2[(2\langle 22|11\rangle + 2\langle 21|21\rangle - \langle 11|11\rangle - \langle 22|22\rangle)\langle 0|1_\alpha^+ 1_\beta^+ 2_\beta 2_\alpha|0\rangle \\ &\quad - \langle 22|11\rangle\langle 0|2_\alpha^+ 2_\beta^+ 2_\beta 2_\alpha + 1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha|0\rangle] \\ &= -0.2400\end{aligned}$$

$$\begin{aligned}(\mathsf{A}_{21})_{|1\rangle,21} &= \langle 1|[H, 2_\alpha^+ 1_\alpha + 2_\beta^+ 1_\beta]|0\rangle = 2[h_{12}\langle 1|1_\alpha^+ 1_\alpha - 2_\alpha^+ 2_\alpha|0\rangle \\ &\quad - \langle 11|12\rangle\langle 1|1_\beta^+ 1_\alpha^+ 1_\beta 1_\alpha + 2_\alpha^+ 2_\beta^+ 1_\beta 1_\alpha|0\rangle \\ &\quad - \langle 22|12\rangle\langle 1|2_\beta^+ 2_\alpha^+ 1_\beta 1_\alpha + 2_\alpha^+ 2_\beta^+ 2_\beta 2_\alpha|0\rangle] \\ &= 0.4020\end{aligned}$$

$$\begin{aligned}(\mathsf{B}_{21})_{|1\rangle,21} &= \langle 1|[H, 1_\alpha^+ 2_\alpha + 1_\beta^+ 2_\beta]|0\rangle \\ &= 2[h_{12}\langle 1|2_\alpha^+ 2_\alpha - 1_\alpha^+ 1_\alpha|0\rangle - \langle 22|21\rangle\langle 1|2_\beta^+ 2_\alpha^+ 2_\beta 2_\alpha + 1_\alpha^+ 1_\beta^+ 2_\beta 2_\alpha|0\rangle \\ &\quad - \langle 11|21\rangle\langle 1|1_\alpha^+ 1_\beta^+ 1_\beta 1_\alpha + 1_\beta^+ 1_\alpha^+ 2_\beta 2_\alpha|0\rangle] \\ &= -0.0198\end{aligned}$$

$$(\mathsf{A}_{22})_{|1\rangle,|1\rangle} = -2.0079 + 4.2790 = 2.2711, \qquad (\mathsf{B}_{22})_{|1\rangle,|1\rangle} = 0$$

3. Since $\mathsf{A} - \mathsf{B}$ is a 2×2 matrix, we invert it easily and obtain the numerical value of κ_{21} and P_{10} through $\binom{\boldsymbol{\kappa}}{\mathsf{P}} = (\mathsf{A} - \mathsf{B})^{-1}\binom{\mathsf{W}}{\mathsf{V}}$. Hence the $\boldsymbol{\kappa}$ and P matrices are

$$\boldsymbol{\kappa} = \begin{pmatrix} 0.0 & 0.0097 \\ -0.0097 & 0.0 \end{pmatrix}, \qquad \mathsf{P} = \begin{pmatrix} 0.0 & -0.0018 \\ 0.0018 & 0.0 \end{pmatrix}$$

4. By applying $\exp(-\boldsymbol{\kappa})$ as shown in Problem 1.5 to the HF orbitals, we obtain

$$\phi_1 = 0.8919\ 1s_{\mathrm{He}} + 0.1701\ 1s_{\mathrm{H}}, \qquad \phi_2 = -0.8410\ 1s_{\mathrm{He}} + 1.2140\ 1s_{\mathrm{H}}$$

Likewise, transformation of the two CI eigenstates through $\exp(-\mathsf{P})$ gives rise to two new MC state vectors whose expansion coefficients are given by

$$\mathsf{C} = \begin{pmatrix} 0.9984 & 0.0574 \\ -0.0574 & 0.9984 \end{pmatrix}$$

5. Since we used the states of a CI calculation as the initial guess for the reference state in the one-step MCSCF procedure, the orbitals of the one- and two-step MCSCF approach become identical when the first iteration is carried out. The states obtained in the next iteration of the two-step MCSCF approach would, however, be determined from a new CI calculation (which requires transformed integrals) and would thus differ from the states obtained in the one-step MCSCF procedure.

6.

$E_{\mathrm{MCSCF}} - E_{\mathrm{MCSCF}}^{\mathrm{CONV}}$	$(E_{\mathrm{MCSCF}} - E_{\mathrm{MCSCF}}^{\mathrm{CONV}})^2$
2.1494×10^{-4}	4.6×10^{-8}
1×10^{-8}	1×10^{-16}

The error in the $(n + 1)$th iteration is the square of the error in the nth iteration.

7. The MCSCF and the full CI calculation have the same number and kind of variational parameters; hence the total energies obtained in the two calculations should become identical. The $|1\sigma 2\sigma\rangle$ configuration included in the CI wavefunction is treated in the two-configuration MCSCF function through the 2^+1 κ_{21} orbital optimization parameter.

8. In an MCSCF calculation that uses $1\sigma^2$, $2\sigma^2$, and $1\sigma^1 2\sigma^1$, the $2\sigma^+ 1\sigma$ excitation operator becomes a redundant excitation operator and hence the orbital optimization step need not be included. The three-configuration MCSCF calculation thus becomes identical to the three-configuration CI calculation.

9. Again, two configurations plus one degree of orbital optimization freedom span all of the configuration space needed to generate the full CI wavefunction. Hence the converged MCSCF energy would equal the full CI energy here.

2.7

1. In the one- and two-electron integrals appearing in the second-quantized form of H, we neglect all integrals involving orbitals on different HeH^+ ions. Hence

$$H = \sum_{A=1}^{n} \left(\sum_{\substack{i,j=1\sigma,2\sigma \\ \in A}} h_{ij} i^+ j + \frac{1}{2} \sum_{\substack{ijkl=1\sigma,2\sigma \\ \in A}} \langle ij|kl\rangle i^+ j^+ lk \right) = \sum_{A=1}^{n} H_A$$

where A labels the n HeH^+ ions.

2. Let us denote the reference HF determinant by HF. Then doubly excited configurations involving excitation of the Ath HeH^+ ion can be represented as $(2_\alpha^+ 2_\beta^+ 1_\beta 1_\alpha)_A|\mathrm{HF}\rangle \equiv |A\rangle$. Doubly excited configurations in which one orbital is excited on each of two HeH^+ ions will not give rise to

nonvanishing CI matrix elements since the one- and two-electron integrals that arise in evaluating such CI elements would vanish because of the large separation between the two ions.

The CI matrix elements arising from the functions $|\mathrm{HF}\rangle$ and $\{|A\rangle\}$ are

$$\langle \mathrm{HF}|H|\mathrm{HF}\rangle = \sum_{A=1}^{n} \langle \mathrm{HF}|H_A|\mathrm{HF}\rangle = n\langle 1\sigma^2|H|1\sigma^2\rangle \equiv nE_{\mathrm{HF}}$$

(This result follows since $|\mathrm{HF}\rangle = \prod_{A=1}^{n} (1_\beta^+ 1_\alpha^+)_A|\mathrm{vac}\rangle$);

$$\langle \mathrm{HF}|H|A\rangle = \langle 1\sigma^2|H|2\sigma^2\rangle \qquad \text{(the same for all } A)$$

$$\langle A|H|A'\rangle = \delta_{AA'}[\langle 2\sigma^2|H|2\sigma^2\rangle + (n-1)E_{\mathrm{HF}}] \qquad \text{(the same for all } A)$$

3. The components of the eigenvalue problem ($\mathbf{HC} = E\mathbf{C}$) for the matrix shown in question 2 can be written as

$$nE_{\mathrm{HF}}C_{\mathrm{HF}} + \sum_{A=1}^{n} BC_A = EC_{\mathrm{HF}}, \qquad BC_{\mathrm{HF}} + CC_{A'} = EC_{A'}, \quad A' = 1, \ldots, n$$

Solving for $C_{A'}$ in terms of C_{HF} and substituting into the first equation gives

$$nE_{\mathrm{HF}}C_{\mathrm{HF}} + \sum_{A=1}^{n} B[E - C]^{-1}BC_{\mathrm{HF}} = EC_{\mathrm{HF}}$$

This equation will have a nontrivial solution for C_{HF} only if

$$nE_{\mathrm{HF}} + B^2 n(E - C)^{-1} = E$$

This quadratic equation can be written as

$$(E - C)(nE_{\mathrm{HF}} - C) + nB^2 = (E - C)^2$$

the solutions of which are

$$E - C = \tfrac{1}{2}\{nE_{\mathrm{HF}} - C \pm [(nE_{\mathrm{HF}} - C)^2 + 4nB^2]^{1/2}\}$$

Using the definitions of C and B, the ground-state energy becomes

$$E = nE_{\mathrm{HF}} + \tfrac{1}{2}\{-(E_{\mathrm{HF}}^* - E_{\mathrm{HF}}) - [(E_{\mathrm{HF}}^* - E_{\mathrm{HF}})^2 + 4n\langle 11|22\rangle^2]^{1/2}\} + E_{\mathrm{HF}}^* - E_{\mathrm{HF}}$$

The correlation energy then becomes

$$E_{\mathrm{corr}} = E - nE_{\mathrm{HF}} = \tfrac{1}{2}\{E_{\mathrm{HF}}^* - E_{\mathrm{HF}} - [(E_{\mathrm{HF}}^* - E_{\mathrm{HF}})^2 + 4n\langle 11|22\rangle^2]^{1/2}\}$$

4. From Problem 2.1 we find $\langle 11|22\rangle = 0.1261$ and from Problem 2.3, $\langle 1\sigma^2|H|1\sigma^2\rangle = E_{\mathrm{HF}} = -4.2720$ and $\langle 2\sigma^2|H|2\sigma^2\rangle = E_{\mathrm{HF}}^* = -2.0149$. Therefore ($E_{\mathrm{HF}}^* - E_{\mathrm{HF}} = 2.2571$),

$$E_{\mathrm{corr}} = \frac{2.2571}{2} - \left[\left(\frac{2.2571}{2}\right)^2 + n(0.1261)^2\right]^{1/2}$$

Evaluating E_{corr} for $n = 2, 4, 10, 100$, and 1000, and comparing it to n times the correlation energy of 1 HeH^+ ion, we find

n	E_{corr}	nE_{corr} $(n = 1)$	Error (%)
1	−0.0070	−0.0070	0.0
2	−0.0140	−0.0140	0.0*
4	−0.0278	−0.0280	0.7
10	−0.0684	−0.0700	2.3
100	−0.5637	−0.7000	19.5
1000	−3.0156	−7.0000	56.9

* If one were to carry more significant figures, this result would be 0.3%.

For large n, the analytical expression of question 3 clearly varies as $n^{1/2}$. Comparing our results for $n = 100$ and $n = 1000$, we find a ratio of $3.0156/0.5637 = 5.35$, which is not $(1000/100)^{1/2} = 3.16$. Thus $n = 100$ is not yet in the large-n range. The ratio for $n = 10$ and $n = 100$ is $0.5637/0.0684 = 8.24$, which is even further from $(100/10)^{1/2} = 3.16$. Hence one must go beyond $n = 100$ before this large-n behavior is realized.

5. Within our small basis the HeH^+ is undercorrelated because the 2σ orbital is much higher in energy than would be expected for the lowest excited σ orbital of HeH^+. Therefore, our correlation energy, which arises from the $1\sigma^2 \to 2\sigma^2$ excitation is smaller (because $\langle 11|22 \rangle$ is smaller and $2\varepsilon_2 - 2\varepsilon_1$ is larger) than one would obtain if one were to use a better atomic orbital basis on HeH^+. As a result the $\langle 11|22 \rangle$ appearing in the above expression for E_{corr} is "too small" and $(E^*_{\text{HF}} - E_{\text{HF}})$ is "too large." This leads to an underestimate of E_{corr}.

References

Banerjee, A., and Grein, F. (1976). *Int. J. Quantum Chem.* **10**, 123.
Bartlett, R. J., and Purvis, G. D. (1981). *Ann. N.Y. Acad. Sci.* **62**, 367.
Bender, C. F., and Davidson, E. R. (1967). *J. Chem. Phys.* **47**, 4972.
Brooks, B. R., and Schaefer, H. F. (1979). *J. Chem. Phys.* **70**, 5092.
Buenker, R. J., and Peyerimhoff, S. D. (1974). *Theor. Chim. Acta* **35**, 33.
Dalgaard, E., and Jørgensen, P. (1978). *J. Chem. Phys.* **69**, 3833.
Dalgaard, E. (1979). *Chem. Phys. Lett.* **65**, 559.
Davidson, E. R. (1975). *J. Comput. Phys.* **17**, 87.
Eyring, H., Walter, J., and Kimball, G. E. (1944). "Quantum Chemistry." Wiley, New York.
Kelly, H. P. (1964). *Phys. Rev. B* **136**, 896.
Löwdin, P. O. (1955). *Phys. Rev.* **97**, 1474.
Löwdin, P. O. (1968). *Int. J. Quantum Chem.* **2**, 867.

McWeeney, R., and Sutcliffe, B. T. (1976). "Methods of Molecular Quantum Mechanics." Academic Press, London.

Paldus, J., and Boyle, M. J. (1980). *Phys. Scripta* **21**, 295.

Pople, J. A., Seeger, R., and Krishnan, R. (1977). *Int. J. Quantum Chem.* **S11**, 149.

Roothaan, C. C. J. (1951). *Rev. Mod. Phys.* **23**, 69.

Roothaan, C. C. J. (1960). *Rev. Mod. Phys.* **32**, 179.

Schaefer, H. F., and Miller, W. H., Eds. (1977). "Modern Theoretical Chemistry," Vol. 3. Plenum, New York.

Shavitt, I. (1978). *Int. J. Quantum Chem.* **S12**, 5.

Yeager, D. L., Albertsen, P., and Jørgensen, P. (1980). *J. Chem. Phys.* **73**, 2811.

Yeager, D. L., and Jørgensen, P. (1979). *J. Chem. Phys.* **71**, 755.