

The Theory and Computational Implementation of Quadratically Convergent Hartree–Fock Orbital Optimization Methods and Their Relationship to the Time-dependent Hartree–Fock and RPA Methods

George B. Bacskay

Department of Theoretical Chemistry, University of Sydney, N.S.W. 2006.

Abstract

The theory of quadratically convergent Hartree–Fock or self-consistent field (QC–SCF) orbital optimization is presented using the language of second quantization. Two methods that are appropriate for the computational implementation of QC–SCF are described: the Newton–Raphson method and an approximate super configuration interaction (CI) approach, both of which can be implemented such that no four-index transformation is necessary. The Newton–Raphson formulation of QC–SCF is shown to be equivalent to solving the frequency-independent coupled perturbation Hartree–Fock equations, and consequently a close relationship exists between QC–SCF and the more general time-dependent coupled perturbation Hartree–Fock (TDCPHF) theory and the related theories of random phase approximation (RPA) and time-dependent Hartree–Fock. Matrix element expressions that are needed for RPA naturally arise in QC–SCF too, and a list of these for open-shell ground state wavefunctions is also given. Computational techniques that are believed to be useful for the solution of the TDCPHF and RPA problems are also briefly discussed.

1. Introduction

The Hartree–Fock self-consistent field (SCF) model has proved to be of crucial importance in physics and chemistry, since it provides a simple yet quite accurate description of the electronic structure of atoms and molecules as well as of nuclear structure (Slater 1930; Fock 1930*a*, 1930*b*; Roothaan 1951, 1960; Hartree 1957; Rowe 1970). In quantum chemistry especially the SCF molecular orbital method is still the most widely used approach to obtain *ab initio* estimates of molecular properties. Even when the Hartree–Fock method is inapplicable, for example, in describing molecular dissociation, the calculation of an SCF wavefunction is often a useful and convenient starting point.

The Hartree–Fock method is most readily applicable to the calculation of ground state wavefunctions and that is indeed what constitutes the majority of the reported calculations. The extension of the Hartree–Fock model to the calculation of excited states is most easily and elegantly achieved through the formalism of time-dependent Hartree–Fock (TDHF) theory, which is formally equivalent to the RPA (Rowe 1970). Intimately connected with these latter theories is of course the problem of calculating second and higher order induced properties, for example, electric susceptibilities, which in the Hartree–Fock framework gives rise to time-dependent coupled perturbation Hartree–Fock (TDCPHF) theory (Dalgarno and Victor 1966).

The most widely used method of calculating SCF orbitals, for molecules especially, is based on the finite basis set expansion technique, i.e. linear combinations of (atomic) orbitals, as proposed by Roothaan (1951, 1960). The usual and traditional method

of orbital optimization is to write the relevant Fock equations as matrix pseudo-eigenvalue equations which are solved iteratively until the required degree of self-consistency is achieved. In practical applications, convergence forcing and accelerating methods are often vitally important and many papers have addressed the latter problem alone. The Roothaan iterative scheme is essentially a simple relaxation method, and consequently its convergence is approximately linear (Pulay 1980). The alternative to the above orbital optimization method is to minimize the SCF energy directly with respect to the orbital coefficients, an approach that is utilized in the steepest descent and conjugate gradient methods (McWeeny 1956; Fletcher and Powell 1963; Hillier and Saunders 1970). Unfortunately, little is known about the efficiency of these techniques, and consequently they have not seemed to achieve widespread popularity or acceptance. The philosophy of the direct approaches has been adopted in the development of multiconfigurational SCF (MC-SCF) methods, culminating in the quadratically convergent MC-SCF theories of recent years (Dalgaard and Jørgensen 1978; Yeager and Jørgensen 1979a; Roothaan *et al.* 1979; Siegbahn *et al.* 1980; Werner and Meyer 1980; Shepard and Simons 1980; Shepard *et al.* 1982).

The basic principle of quadratically convergent orbital optimization schemes is to expand the energy as a quadratic hypersurface with respect to changes in the orbital expansion coefficients, for which the minimum is then located by solving the resulting Newton-Raphson (i.e. inhomogeneous linear) set of equations. The key to the successful application of the Newton-Raphson method is provided by the unitary (exponential) transformation of the orbitals and the removal of the redundant parameters, as discussed in detail by Douady *et al.* (1980).

In the present paper the theory of quadratically convergent orbital optimization will be briefly presented and a practical and efficient computational scheme will be discussed that has been recently developed and applied to both closed- and open-shell systems (Bacskay 1981, 1982). As will become evident, a fundamental relationship exists between the quadratically convergent SCF (QC-SCF) and TDCPHF methods, and consequently the TDHF and RPA methods. This relationship will be explored and discussed in some detail. At the same time matrix element expressions will be derived and listed that are needed in open-shell TDHF-RPA calculations. Finally, possible computational schemes will be described that could facilitate the solution of the TDCPHF and TDHF-RPA equations.

2. Theory of QC-SCF Method

Consider an arbitrary single determinantal wavefunction $|\Psi^0\rangle$ written as the antisymmetrized product of a set of occupied spin orbitals $\{\phi_i^0\}$:

$$|\Psi^0\rangle = \hat{A} \prod_i \phi_i(i), \quad (1)$$

where $\hat{A}$ is the antisymmetrizer and i labels the occupied spin orbitals and the corresponding electron coordinates. In the finite basis set formulation of SCF theory there exists also a set of unoccupied or virtual orbitals $\{\phi_a^0\}$, and the full orbital space written as the row vector Φ^0 is assumed to be real and orthonormal.

Consider now a unitary transformation of the orbitals according to

$$\Phi = \Phi^0 U, \quad (2)$$

$$U = \exp X = 1 + X + \frac{1}{2}X^2 + \dots, \quad (3)$$

where X is a real skew symmetric matrix

$$X_{tu} = -X_{ut}. \quad (4)$$

(The subscripts s, t, u, v will be used to label *any* orbital, irrespective of occupancy, while $i, j, (a, b)$ label the occupied (unoccupied) orbitals.) Now let

$$X_{ij} = X_{ab} = 0, \quad \forall i, j, a, b, \quad (5)$$

that is, to first order only, mixing of the occupied and unoccupied subspaces is allowed. This restriction is necessary so that the Hessian will contain no singularities (Douady *et al.* 1980; Siegbahn *et al.* 1980). A given transformed orbital can thus be written as, for example,

$$\phi_i = \phi_i^0 + \sum_a X_{ai} \phi_a^0 + \frac{1}{2} \sum_{j,b} X_{jb} X_{bi} \phi_j^0 + \dots, \quad (6)$$

while the total single determinantal wavefunction $|\Psi^0\rangle$ is transformed according to

$$|\Psi\rangle = \exp(\hat{T})|\Psi^0\rangle = (1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \dots)|\Psi^0\rangle. \quad (7)$$

Here $\hat{T}$ is conveniently written in second quantized form

$$\hat{T} = \sum_{a,i} X_{ai} \hat{\Omega}_{ai}, \quad (8)$$

where

$$\hat{\Omega}_{ai} = \hat{c}_a^\dagger \hat{c}_i - \hat{c}_i^\dagger \hat{c}_a, \quad (9)$$

$\hat{c}_i, \hat{c}_a$ ($\hat{c}_i^\dagger, \hat{c}_a^\dagger$) being the fermion annihilation (creation) operators that act on the relevant particle (hole) states of the Hartree-Fock (vacuum) state $|\Psi^0\rangle$. Equation (7) is basically equivalent to the Thouless (1961) theorem (see also Rowe 1970), according to which any other single determinantal wavefunction $|\Psi\rangle'$, not actually orthogonal to $|\Psi^0\rangle$, can be written in the (unnormalized) form

$$|\Psi\rangle' = \exp\left(\sum_{i,a} \hat{c}_a^\dagger \hat{c}_i\right)|\Psi^0\rangle. \quad (10)$$

The energy expectation value corresponding to the transformed wavefunction is simply given as

$$E = \langle\Psi^0|\exp(\hat{T}^\dagger)\hat{H}\exp(\hat{T})|\Psi^0\rangle, \quad (11)$$

which can then be rewritten as

$$E = \langle\Psi^0|\hat{H} + [\hat{H}, \hat{T}] - \frac{1}{2}[\hat{T}, \hat{H}, \hat{T}] + \dots|\Psi^0\rangle, \quad (12)$$

since

$$\hat{T}^\dagger = -\hat{T}, \quad (13)$$

$$[\hat{T}, \hat{H}, \hat{T}] = \frac{1}{2}([[\hat{T}, \hat{H}], \hat{T}] + [\hat{T}, [\hat{H}, \hat{T}]] = -[[\hat{H}, \hat{T}], \hat{T}]. \quad (14)$$

Using the expansion (8) we can write the energy as

$$E = \langle \Psi^0 | \hat{H} | \Psi^0 \rangle + \sum_{a,i} X_{ai} \langle \Psi^0 | [\hat{H}, \hat{\Omega}_{ai}] | \Psi^0 \rangle - \frac{1}{2} \sum_{a,i} \sum_{b,j} X_{ai} \langle \Psi^0 | [\hat{\Omega}_{ai}, \hat{H}, \hat{\Omega}_{bj}] | \Psi^0 \rangle X_{bj}. \quad (15)$$

The first and second derivatives of the energy at $\mathbf{X} = \mathbf{0}$ are thus given by

$$E'_{ai} = \langle \Psi^0 | [\hat{H}, \hat{\Omega}_{ai}] | \Psi^0 \rangle, \quad (16)$$

$$E''_{ai,bj} = -\langle \Psi^0 | [\hat{\Omega}_{ai}, \hat{H}, \hat{\Omega}_{bj}] | \Psi^0 \rangle, \quad (17)$$

and hence the energy correct to second order is conveniently written as a Taylor series expansion

$$E^{(2)} = E_0 + \mathbf{X} \mathbf{E}' + \frac{1}{2} \mathbf{X} \mathbf{E}'' \mathbf{X}, \quad (18)$$

where $\mathbf{X}$ and $\mathbf{E}'$ are written as (super)vectors and $\mathbf{E}''$ as a supermatrix, and E_0 is the energy expectation value corresponding to the zeroth order reference state $|\Psi^0\rangle$.

For $\hat{H}$ Hermitean and a real orbital basis ϕ^0 , the first and second derivatives of the energy can be simplified to

$$E'_{ai} = 2\langle \Psi^0 | \hat{H} \hat{c}_a^\dagger \hat{c}_i | \Psi^0 \rangle = 2\langle \Psi^0 | \hat{H} | \Psi_i^a \rangle, \quad (19)$$

$$E''_{ai,bj} = 2\langle \Psi^0 | \hat{c}_i^\dagger \hat{c}_a (\hat{H} - E_0 \delta_{ij} \delta_{ab}) \hat{c}_b^\dagger \hat{c}_j | \Psi^0 \rangle + \langle \Psi^0 | \hat{H} (\hat{c}_a^\dagger \hat{c}_i \hat{c}_b^\dagger \hat{c}_j + \hat{c}_b^\dagger \hat{c}_j \hat{c}_a^\dagger \hat{c}_i) | \Psi^0 \rangle \quad (20a)$$

$$= 2\langle \Psi_i^a | \hat{H} - E_0 \delta_{ij} \delta_{ab} | \Psi_j^b \rangle + 2\langle \Psi^0 | \hat{H} | \Psi_{ij}^{ab} \rangle, \quad (20b)$$

where $\{|\Psi_i^a\rangle\}$ and $\{|\Psi_{ij}^{ab}\rangle\}$ are the familiar singly and doubly excited determinants with respect to $|\Psi^0\rangle$. For future use the two terms in (20a) will be denoted by

$$A_{ai,bj} = 2\langle \Psi^0 | \hat{c}_i^\dagger \hat{c}_a (\hat{H} - E_0 \delta_{ij} \delta_{ab}) \hat{c}_b^\dagger \hat{c}_j | \Psi^0 \rangle \quad (21a)$$

$$= \langle \Psi^0 | [\hat{c}_i^\dagger \hat{c}_a, \hat{H}, \hat{c}_b^\dagger \hat{c}_j] | \Psi^0 \rangle, \quad (21b)$$

$$B_{ai,bj} = \langle \Psi^0 | \hat{H} (\hat{c}_a^\dagger \hat{c}_i \hat{c}_b^\dagger \hat{c}_j + \hat{c}_b^\dagger \hat{c}_j \hat{c}_a^\dagger \hat{c}_i) | \Psi^0 \rangle, \quad (22a)$$

$$= \langle \Psi^0 | [\hat{c}_a^\dagger \hat{c}_i, \hat{H}, \hat{c}_b^\dagger \hat{c}_j] | \Psi^0 \rangle, \quad (22b)$$

$$= -\langle \Psi^0 | \hat{c}_i^\dagger \hat{c}_a, \hat{H}, \hat{c}_j^\dagger \hat{c}_b | \Psi^0 \rangle, \quad (22c)$$

given a real orbital basis and a Hermitean Hamiltonian $\hat{H}$.

Equation (18) represents a quadratic expansion of the energy with respect to the orbital 'rotations' $\{X_{ai}\}$, and provided the (Hessian) matrix of second derivatives is positive definite a minimum can be found that yields the Newton-Raphson equations:

$$\partial E^{(2)} / \partial X_{ai} = E'_{ai} + \sum_{b,j} E''_{ai,bj} X_{bj} = 0, \quad \forall i, a, \quad (23)$$

and hence the solution is simply written as

$$\mathbf{X} = -\mathbf{E}''^{-1} \mathbf{E}'. \quad (24)$$

An alternative approach to obtain X is by a matrix diagonalization technique, which is referred to as an approximate super configuration interaction method, i.e. it uses the technique of CI. Thus, an energy functional $\varepsilon[X]$ is defined as

$$\varepsilon[X] = (E^{(2)} + E_0 D^{(2)}) / (1 + D^{(2)}), \quad (25)$$

where $1 + D^{(2)}$ is an energy denominator defined by

$$1 + D^{(2)} = \langle \Psi^0 | (1 + \hat{T} + \frac{1}{2} \hat{T}^2) (1 + \hat{T} + \frac{1}{2} \hat{T}^2) | \Psi^0 \rangle, \quad (26)$$

which for a orthonormal spin-orbital basis reduces to

$$1 + D^{(2)} = 1 + \sum_{i,a} \langle \Psi^0 | \hat{Q}_{ai}^\dagger \hat{Q}_{ai} | \Psi^0 \rangle X_{ai}^2 \quad (27)$$

$$= 1 + \sum_{i,a} S_{ai,ai} X_{ai}^2, \quad (28)$$

where S is a diagonal matrix, in fact, a unit matrix. It is readily shown that $\varepsilon[X]$ contains third and higher order errors in comparison with the exact quadratic expression $E^{(2)}$, and consequently convergence is still essentially quadratic, but the radius of convergence is larger than for other second order procedures such as the Newton-Raphson scheme (Yarkony 1981; Bacskay 1982; Shepard *et al.* 1982). The approximate super CI approach can most simply be developed by noting that

$$\varepsilon[X] = \langle \Psi | \hat{H} | \Psi \rangle / \langle \Psi | \Psi \rangle, \quad (29)$$

where

$$| \Psi \rangle = | \Psi^0 \rangle + \sum_{i,a} X_{ai} | \Psi_i^a \rangle + \frac{1}{2} \sum_{i,a} \sum_{j,b} X_{ai} X_{bj} | \Psi_{ij}^{ab} \rangle, \quad (30)$$

and in (29) both numerator and denominator are evaluated up to quadratic terms in $\{X_{ai}\}$. Minimization of equation (25), i.e. of (29), leads to the secular equations

$$\begin{bmatrix} E_0 - \varepsilon & \vdots & \frac{1}{2} E'^\dagger \\ \vdots & \ddots & \vdots \\ \frac{1}{2} E' & \vdots & \frac{1}{2} E'' + (E_0 - \varepsilon) S \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ X \end{bmatrix} = 0, \quad (31)$$

which, in principle, are readily soluble by a suitable matrix diagonalization program. Since for ground state wavefunctions only the lowest eigenvalue is required, the iterative eigenvalue routines that are widely used in most CI packages, such as the variation-perturbation (Brändas and Goscinski 1970) and Davidson (1975) methods, are readily applicable.

The key to the practical computational implementation of the QC-SCF method is the use of the direct CI technique (Roos 1972; Roos and Siegbahn 1977), coupled with one of the above iterative diagonalization methods. As discussed in detail elsewhere (Bacskay 1981, 1982), the iterative solution of equation (31) can be carried out directly in terms of the *atomic* integrals or supermatrices, that is, avoiding the time consuming four-index transformation needed for the construction of the Hessian E . Thus E is not constructed explicitly, only vectors of the form

$$q_I = E b_I, \quad (32)$$

where $\{b_i\}$ represents a set of suitable basis vectors in terms of which the column vector in equation (31) is expanded. The construction of one q_i requires approximately the same computer time as the setting up of the Fock matrix in the traditional schemes. The same iterative approach discussed above can be used to solve the related inhomogeneous linear, that is, Newton–Raphson equations (23) and (24), as discussed by Pople (1968) and Bacskay (1981, 1982).

In the iterative Hartree–Fock orbital optimization process, once X is obtained the unitary matrix U of equation (3) is constructed according to (16):

$$U = \exp X = V(\cos |d|)V^\dagger + XV(\sin |d|)|d|^{-1}V, \quad (33)$$

where V is the unitary matrix that diagonalizes X^2 , i.e.

$$V^\dagger X^2 V = d^2, \quad (34)$$

noting that X^2 is a negative semi-definite Hermitean matrix, hence its eigenvalues d^2 are all negative or zero.

Having obtained a set of updated orbitals we repeat the process until the first derivatives of the energy, i.e. the Brillouin matrix elements $\{E_{ai}\}$, all fall in magnitude below a suitably chosen threshold value.

Test calculations that have been carried out to date indicate that QC–SCF compares favourably with the traditional Roothaan scheme, but is most useful in calculations where a high degree of convergence is required, for example in finite field and force field calculations (Bacskay 1981, 1982). In typical cases the computer time requirements are $\sim 30\%$ less for QC–SCF.

In order to derive the working equations that are most suitable for *restricted* (open or closed shell) Hartree–Fock (RHF) calculations it is more convenient to write the operators $\hat{T}$ and $\{\hat{\Omega}_{ts}\}$ of equations (8) and (9) in terms of the generators of the unitary group $\{\hat{E}_{ts}\}$ (Pople 1968; Paldus 1974, 1975, 1976; Shavitt 1978; Matsen 1979; Matsen and Nelin 1979):

$$\hat{T} = \sum_{s < t} X_{ts} \hat{\Omega}_{ts}, \quad (35)$$

$$\hat{\Omega}_{ts} = \hat{E}_{ts} - \hat{E}_{st}, \quad (36)$$

where

$$\hat{E}_{ts} = \sum_{\sigma} \hat{c}_{t\sigma}^\dagger \hat{c}_{s\sigma}, \quad (37)$$

the summation over σ implying summation over α and β spins. The operators $\{\hat{E}_{ts}\}$ satisfy the commutation relations

$$[\hat{E}_{ts}, \hat{E}_{uv}] = \delta_{su} \hat{E}_{tv} - \delta_{tv} \hat{E}_{us}, \quad (38)$$

which readily follow from the anticommutation properties of the fermion creation and annihilation operators.

For the more general open-shell RHF calculations the (spatial) orbital space is split into three subspaces D, P, V spanned by the sets of doubly partially occupied and unoccupied orbitals $\{\phi_i^0\}$, $\{\phi_p^0\}$, $\{\phi_a^0\}$ respectively. The condition for an optimized RHF wavefunction that represents a ground state is that the energy must be at its minimum with respect to any unitary transformation of the orbitals. Consequently, the wavefunction $|\Psi^0\rangle$ must be invariant under any unitary mixing of the D and

P subspaces, and the energy must be stable under unitary transformations that result in the mixings $V|D$, $D|P$ and $V|P$. Thus, after setting to zero the redundant parameters $\{X_{ij}\}$, $\{X_{pq}\}$, $\{X_{ab}\}$ (where i, j label D orbitals, p, q label P orbitals and a, b label V orbitals), three distinct types of first derivatives are obtained, namely $\{E'_{ai}\}$, $\{E'_{pi}\}$, $\{E'_{ap}\}$, and six types of second derivatives $\{E''_{ai,bj}\}$, $\{E''_{ai,pj}\}$ etc. Explicit formulae for these have been given elsewhere (Bacskay 1982) in terms of one- and two-electron integrals and the coupling constants w, x, y (Guest and Saunders 1974) that are necessary to write down the average energy expression over all degenerate configurations where appropriate.

3. Relationship of QC-SCF to TDCPHF, TDHF and RPA Theories

In the equations of motion method (Rowe 1970; McCurdy *et al.* 1977), given the ground state wavefunction $|0\rangle$ for the system, an excitation operator $\hat{O}_\lambda^\dagger$ for state $|\lambda\rangle$ is defined by

$$\hat{O}_\lambda^\dagger |0\rangle = |\lambda\rangle, \quad \hat{O}_\lambda |0\rangle = 0. \quad (39a, b)$$

Here $\hat{O}_\lambda^\dagger$ and its Hermitean conjugate $\hat{O}_\lambda$ satisfy the equations of motion

$$[\hat{H}, \hat{O}_\lambda^\dagger] |0\rangle = \omega_\lambda \hat{O}_\lambda^\dagger |0\rangle, \quad [\hat{H}, \hat{O}_\lambda] |0\rangle = -\omega_\lambda \hat{O}_\lambda |0\rangle = 0, \quad (40a, b)$$

and also

$$\langle 0 | [\hat{O}_\kappa, \hat{H}, \hat{O}_\lambda^\dagger] | 0 \rangle = \omega_\lambda \langle 0 | [\hat{O}_\kappa, \hat{O}_\lambda^\dagger] | 0 \rangle, \quad (41)$$

where ω_λ is the excitation energy and the double commutator is defined as in equation (14):

$$[\hat{a}, \hat{b}, \hat{c}] = \frac{1}{2}([\hat{a}, [\hat{b}, \hat{c}]] + [[\hat{a}, \hat{b}], \hat{c}]). \quad (42)$$

Equation (41) is the equation of motion for $\hat{O}_\lambda^\dagger$ and is particularly useful for a variational or finite basis solution.

In the RPA $|0\rangle$ is the Hartree-Fock ground state and $\hat{O}_\lambda^\dagger$ is expanded in the simple particle-hole (spin-orbital) form

$$\hat{O}_\lambda^\dagger = \sum_{i,a} \{Y_{ai}(\lambda) \hat{c}_a^\dagger \hat{c}_i - Z_{ai}(\lambda) \hat{c}_i^\dagger \hat{c}_a\}. \quad (43)$$

In order to apply RPA to open-shell RHF ground states, the excitation operator is written in terms of the generators of the unitary group:

$$\begin{aligned} \hat{O}_\lambda^\dagger = & \sum_{i,a} \{Y_{ai}(\lambda) \hat{E}_{ai} - Z_{ai}(\lambda) \hat{E}_{ia}\} + \sum_{i,p} \{Y_{pi}(\lambda) \hat{E}_{pi} - Z_{pi}(\lambda) \hat{E}_{ip}\} \\ & + \sum_{p,a} \{Y_{ap}(\lambda) \hat{E}_{ap} - Z_{ap}(\lambda) \hat{E}_{pa}\}, \end{aligned} \quad (44)$$

where, as in Section 2, the indices i, j label doubly occupied orbitals, p, q partially occupied orbitals, a, b unoccupied (virtual) orbitals and s, t, u, v are used to label any orbital where necessary. Given the excitation operator according to equation (44), the variational solution of (41) yields the well-known equations

$$\begin{bmatrix} \mathbf{A} & \vdots & \mathbf{B} \\ \cdots & \ddots & \cdots \\ -\mathbf{B}^* & \vdots & -\mathbf{A}^* \end{bmatrix} \begin{bmatrix} \mathbf{Y} \\ \cdots \\ \mathbf{Z} \end{bmatrix} = \omega_\lambda \begin{bmatrix} \mathbf{D} & \vdots & 0 \\ \cdots & \ddots & \cdots \\ 0 & \vdots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{Y} \\ \cdots \\ \mathbf{Z} \end{bmatrix}, \quad (45)$$

where

$$A_{st,uv} = g^{-1} \sum_d \langle 0_d | [\hat{E}_{ts}, \hat{H}, \hat{E}_{uv}] | 0_d \rangle, \quad (46a)$$

$$B_{st,uv} = -g^{-1} \sum_d \langle 0_d | [\hat{E}_{ts}, \hat{H}, \hat{E}_{vu}] | 0_d \rangle, \quad (46b)$$

$$D_{st,uv} = g^{-1} \sum_d \langle 0_d | [\hat{E}_{ts}, \hat{E}_{uv}] | 0_d \rangle, \quad (46c)$$

and where g is the degeneracy of the ground state, one component of which is $|0_d\rangle$. The above generalization of RPA to cover degenerate ground state wavefunctions is consistent with the QC-SCF and TDCPHF theories.

Explicit formulae for the **A**, **B**, **D** matrices have been derived inductively and explicitly verified for non-degenerate and doubly degenerate (for example π^1, π^3) ground states. While the formulae given in the Appendix are expected to be correct, they should be verified for other cases as well, for example the p^1, p^2, p^4, p^5 configurations.

The TDHF theory (Rowe 1970; Jamieson 1971; Jamieson and Watts 1980) is equivalent to the RPA although often preferred for its physical appeal. Briefly, in TDHF we look at the time development of an arbitrary wavefunction $|\Phi(t)\rangle$ that at every instant in time is a single determinant. Thus, utilizing the Thouless theorem we write

$$|\Phi(t)\rangle = \exp\{-(i/\hbar)E_0 t\} \exp\left(\sum_{a,i} c_{ai}(t) \hat{c}_a^\dagger c_i\right) |0\rangle, \quad (47)$$

where $|0\rangle$ is the ground state Hartree-Fock wavefunction with energy E_0 . Application of the variational principle to leading order yields

$$\sum_{j,b} \left(A_{ai,bj} c_{bj}(t) + B_{ai,bj} c_{bj}^*(t) \right) = (i/\hbar)(\partial/\partial t)c_{ai}, \quad \forall i, a. \quad (48)$$

By putting

$$c_{ai}(t) \propto Y_{ai} \exp(-i\omega t) + Z_{ai}^* \exp(i\omega t), \quad (49)$$

and substituting (49) into (48) yields, after equating positive and negative frequency components, the RPA equations (45) with **A**, **B**, **D** defined in the appropriate spin-orbital basis.

In the context of this paper it is however the TDCPHF approach (Dalgarno and Victor 1966) which is of more interest. This method involves solving the perturbed TDHF equations in the presence of an external time-dependent perturbation:

$$\hat{\mathcal{V}}(t) = \{\exp(i\epsilon t) + \exp(-i\epsilon t)\} \hat{V},$$

where $\hat{V}$ is a time-independent operator, for example the dipole moment operator if the perturbation is due to an oscillating electric field. Application of the variational principle to the perturbed TDHF problem yields to leading order the inhomogeneous equations

$$(\mathbf{A} - \epsilon \mathbf{D}) \mathcal{Y}(\epsilon) + \mathbf{B} \mathcal{Z}(\epsilon) = -V, \quad (50a)$$

$$\mathbf{B}^* \mathcal{Y}(\epsilon) + (\mathbf{A}^* + \epsilon \mathbf{D}) \mathcal{Z}(\epsilon) = -V, \quad (50b)$$

where in a spin-orbital basis

$$V_{ai} = \langle 0 | [\hat{V}, \hat{c}_a^\dagger \hat{c}_i] | 0 \rangle = \langle 0 | \hat{V} \hat{c}_a^\dagger \hat{c}_i | 0 \rangle, \quad (51)$$

or in a spatial-orbital basis for an RHF ground state

$$V_{st} = g^{-1} \sum_d \langle 0_d | [\hat{V}, \hat{E}_{st}] | 0_d \rangle. \quad (52)$$

Note also that the poles in the second order response function,

$$E_2(\varepsilon) = -\frac{1}{2} \sum_{i,a} \langle 0 | [\hat{V}, \hat{c}_a^\dagger \hat{c}_i] | 0 \rangle \{ \mathcal{Y}_{ai}(\varepsilon) + \mathcal{Z}_{ai}(\varepsilon) \}, \quad (53a)$$

correspond to the RPA excitation energies $\{\omega_\lambda\}$, since in the RPA formalism

$$E_2(\varepsilon) = - \sum_\lambda f_{0\lambda} / (\omega_\lambda^2 - \varepsilon^2), \quad (53b)$$

where

$$f_{0\lambda} = \omega_\lambda \left(\sum_{i,a} \langle 0 | [\hat{V}, \hat{c}_a^\dagger \hat{c}_i] | 0 \rangle \{ Y_{ai}(\lambda) + Z_{ai}(\lambda) \} \right)^2. \quad (54)$$

If the external perturbation is time independent, given a real-orbital basis, i.e. real **A**, **B**, **D** and **V**, then equations (50) can be rewritten as

$$(\mathbf{A} + \mathbf{B}) \{ \mathcal{Y}(0) + \mathcal{Z}(0) \} = -2\mathbf{V}, \quad (\mathbf{A} - \mathbf{B}) \{ \mathcal{Y}(0) - \mathcal{Z}(0) \} = 0. \quad (55a, b)$$

Hence for $\varepsilon = 0$ we have according to (55b)

$$\mathcal{Y}(0) = \mathcal{Z}(0), \quad (56)$$

whereas the solution of (55a) yields $2\mathcal{Y}(0)$, i.e. the first correction to the Hartree-Fock unperturbed ground state wavefunction in the presence of the static perturbation. Equations (55a) are known as the coupled perturbation Hartree-Fock (CPHF) equations and have exactly the same form as the Newton-Raphson equations (23), since $\mathbf{A} + \mathbf{B}$ is in fact the Hessian matrix $\mathbf{E}''$, while $\mathbf{V}$ takes the place of the vector of first derivatives $\mathbf{E}'$. Thus, as discussed in more detail elsewhere (Bacskay 1981), the Newton-Raphson formulation is equivalent to solving the CPHF equations, treating $\mathbf{E}'$ as an 'external' perturbation.

As discussed in Section 2, as well as elsewhere (Bacskay 1981, 1982), the solution of the CPHF equations (55a) is readily accomplished by an iterative scheme that does not require $\mathbf{A} + \mathbf{B}$ explicitly but rather a set of vectors

$$\mathbf{q}_I = (\mathbf{A} + \mathbf{B}) \mathbf{b}_I, \quad (57)$$

the solution to (55a) being expanded as

$$\mathcal{Y}(0) = \sum_I \alpha_I \mathbf{b}_I. \quad (58)$$

As already noted $\{\mathbf{q}_I\}$ can be constructed directly from the *atomic* two-electron integrals, thus no four-index transformation is required, making this scheme very attractive computationally. The CPHF calculations of dipole polarizabilities of N_2 , O_2 and NO have demonstrated that the iterative scheme discussed above is efficient, requiring about six iterations to achieve a six significant figure accuracy in the calculated polarizabilities.

The question that naturally arises is whether the above type of iterative scheme could be applied to the solution of the TDCPHF equations, that is, to the calculation

of dynamic properties. If we assume $\mathbf{A}, \mathbf{B}, \mathbf{D}$ to be real, the TDCPHF equations can be written in the form

$$(\mathbf{A} + \mathbf{B})\{\mathcal{Y}(\varepsilon) + \mathcal{Z}(\varepsilon)\} - \varepsilon \mathbf{D}\{\mathcal{Y}(\varepsilon) - \mathcal{Z}(\varepsilon)\} = -2V, \quad (59a)$$

$$(\mathbf{A} - \mathbf{B})\{\mathcal{Y}(\varepsilon) - \mathcal{Z}(\varepsilon)\} - \varepsilon \mathbf{D}\{\mathcal{Y}(\varepsilon) + \mathcal{Z}(\varepsilon)\} = 0, \quad (59b)$$

and an iterative solution could well be feasible using suitable sets of basis vectors to expand $\mathcal{Y}(\varepsilon) + \mathcal{Z}(\varepsilon)$ and $\mathcal{Y}(\varepsilon) - \mathcal{Z}(\varepsilon)$. Numerical work is to be shortly undertaken to examine the feasibility of such a scheme. Whether such an approach is ultimately preferable to the standard method, i.e. the explicit construction of the $\mathbf{A}, \mathbf{B}, \mathbf{D}$ matrices from the *molecular* two-electron integrals, depends of course on the efficiency of the proposed iterative scheme and on the number of frequencies ε for which the calculations are to be carried out.

On the other hand, if $\mathbf{A}, \mathbf{B}, \mathbf{D}$ are to be constructed explicitly a useful and efficient algorithm would be to carry out partial four-index transformations of *supermatrices* rather than of two-electron integrals. For example, all elements of $\mathbf{A} + \mathbf{B}$, i.e. of $\mathbf{E}''$, are very simply related to the elements of the supermatrices $\mathbf{P}$ and $\mathbf{K}$ (Bacskay 1981, 1982) and are readily constructed in the same order as the supermatrix elements. For the construction of $\mathbf{A} - \mathbf{B}$, however, another supermatrix $\mathbf{T}$ is needed, where (see the Appendix)

$$T_{st,uv} = [su | tv] - [sv | tu]. \quad (60)$$

To summarize this section, the equivalence of the CPHF and the Newton-Raphson formulation of QC-SCF has been demonstrated. While the solution of the TDCPHF and TDHF-RPA equations is certainly more complex, it is possible that the techniques that were developed to solve the QC-SCF equations could be generalized so as to be useful for these problems also.

4. Summary and Conclusions

In this paper the theory of quadratically convergent orbital optimization for Hartree-Fock wavefunctions has been presented. Two schemes for numerical implementation have been discussed, the Newton-Raphson and the approximate super CI methods, both of which can be used in iterative approaches, thus avoiding the four-index transformation. In certain applications QC-SCF was found to have significant advantages over the traditional scheme, yet hardly more difficult to implement on the computer.

The connection between QC-SCF and CPHF has been discussed, thus establishing a relationship with RPA and TDHF. The matrix element expressions that are needed for open-shell QC-SCF calculations are also those that occur in RPA-TDHF and are listed in the Appendix. Finally, the feasibility of making use of existing QC-SCF computational techniques for TDCPHF and TDHF-RPA calculations has been commented on.

It should be noted at this point that the ideas outlined in this paper have recently been extended to multiconfigurational SCF (MC-SCF) ground state wavefunctions. Thus MC-SCF wavefunctions are nowadays routinely obtained by quadratically convergent methods (Dalgaard and Jørgensen 1978; Yeager and Jørgensen 1979a; Roothaan *et al.* 1979; Siegbahn *et al.* 1980; Werner and Meyer 1980; Shepard and

Simons 1980; Shepard *et al.* 1982), while the theories of TDHF-RPA, TDCPHF and CPHF have been suitably generalized for MC-SCF reference states (Yeager and Jørgensen 1979b; Albertson *et al.* 1980a, 1980b). Recent calculations using these techniques have indicated that they are capable of good accuracy and in some cases are preferable to the traditional RPA type calculations. Nevertheless, for a wide class of atomic and molecular problems the techniques based on a Hartree-Fock reference state remain valid and most useful.

References

- Albertson, P., Jørgensen, P., and Yeager, D. L. (1980a). *Mol. Phys.* **41**, 409.
 Albertson, P., Jørgensen, P., and Yeager, D. L. (1980b). *Int. J. Quantum Chem. Symp.* **14**, 249.
 Bacskey, G. B. (1981). *Chem. Phys.* **61**, 385.
 Bacskey, G. B. (1982). *Chem. Phys.* **65**, 383.
 Brändas, E., and Goscinski, O. (1970). *Phys. Rev. A* **1**, 552.
 Dalgaard, E., and Jørgensen, P. (1978). *J. Chem. Phys.* **69**, 3833.
 Dalgarno, A., and Victor, G. A. (1966). *Proc. R. Soc. London A* **291**, 291.
 Davidson, E. R. (1975). *J. Comput. Phys.* **17**, 87.
 Douady, J., Ellinger, Y., Sebra, R., and Levy, B. (1980). *J. Chem. Phys.* **72**, 1452.
 Fletcher, R., and Powell, M. J. D. (1963). *Comput. J.* **6**, 163.
 Fock, V. (1930a). *Z. Phys.* **61**, 126.
 Fock, V. (1930b). *Z. Phys.* **62**, 795.
 Guest, M. J., and Saunders, V. R. (1974). *Mol. Phys.* **28**, 819.
 Hartree, D. R. (1957). 'The Calculation of Atomic Structures' (Wiley: New York).
 Hillier, I. H., and Saunders, V. R. (1970). *Proc. R. Soc. London A* **320**, 161.
 Jamieson, M. J. (1971). *Int. J. Quantum Chem.* **4**, 103.
 Jamieson, M. J., and Watts, R. S. (1980). *Chem. Phys. Lett.* **76**, 287.
 McCurdy, Jr., C. W., Rescigno, T. N., Yeager, D. L., and McKoy, V. (1977). In 'Methods of Electronic Structure Theory' (Ed. H. F. Schaefer III), Ch. 9 (Plenum: New York).
 McWeeny, R. (1956). *Proc. R. Soc. London A* **235**, 496.
 Matsen, F. A. (1979). *Adv. Quantum Chem.* **15**, 751.
 Matsen, F. A., and Nelin, C. J. (1979). *Int. J. Quantum Chem.* **15**, 751.
 Paldus, J. (1974). *J. Chem. Phys.* **61**, 5321.
 Paldus, J. (1975). *Int. J. Quant. Chem.* **9**, 165.
 Paldus, J. (1976). *Phys. Rev. A* **14**, 1620.
 Pople, J. A. (1968). In 'Group Theory and the Many Body Problem' (Ed. M. Moshinsky), (Gordon and Breach: New York).
 Pulay, P. (1980). *Chem. Phys. Lett.* **73**, 393.
 Roos, B. (1972). *Chem. Phys. Lett.* **15**, 153.
 Roos, B., and Siegbahn, P. E. M. (1977). In 'Methods of Electronic Structure Theory' (Ed. H. F. Schaefer III), Ch. 7 (Plenum: New York).
 Roothaan, C. C. J. (1951). *Rev. Mod. Phys.* **23**, 69.
 Roothaan, C. C. J. (1960). *Rev. Mod. Phys.* **32**, 179.
 Roothaan, C. C. J., Detrich, J., and Hooper, D. G. (1979). *Int. J. Quantum Chem. Symp.* **13**, 93.
 Rowe, D. J. (1970). 'Nuclear Collective Motion' (Methuen: London).
 Shavitt, I. (1978). *Int. J. Quantum Chem. Symp.* **12**, 5.
 Shepard, R., Shavitt, I., and Simons, J. (1982). *J. Chem. Phys.* (to be published).
 Shepard, R., and Simons, J. (1980). *Int. J. Quantum Chem. Symp.* **14**, 211.
 Siegbahn, P., Heiberg, A., Roos, B., and Levy, B. (1980). *Phys. Scripta* **21**, 323.
 Slater, J. C. (1930). *Phys. Rev.* **35**, 210.
 Thouless, D. J. (1961). 'Quantum Mechanics of Many Body Systems' (Academic: New York).
 Werner, H.-J., and Meyer, W. (1980). *J. Chem. Phys.* **73**, 2342.
 Yarkony, D. (1981). *Chem. Phys. Lett.* **77**, 634.
 Yeager, D. L., and Jørgensen, P. (1979a). *J. Chem. Phys.* **71**, 755.
 Yeager, D. L., and Jørgensen, P. (1979b). *Chem. Phys. Lett.* **65**, 77.

Appendix. RPA Matrix Elements for Open-shell RHF Ground States

For definitions of the Fock matrices $\mathbf{F}^{DV}$, $\mathbf{F}^{PV}$, $\mathbf{F}^{DP}$ see Bacskay (1982). The notation for two-electron integrals is $[st|uv] = \langle s(1)u(2)|r_{12}^{-1}|t(1)v(2)\rangle$:

$$A_{ai,bj} = 4[ai|bj] - 2[ab|ij] + 2F_{ab}^{DV} \delta_{ij} - 2F_{ij}^{DV} \delta_{ab},$$

$$B_{ai,bj} = 4[ai|bj] - 2[aj|bi],$$

$$A_{ai,pj} = (2-w)(2[ai|pj] - [aj|pi]) + F_{pa}^{DP},$$

$$B_{ai,pj} = (2-w)(2[ai|pj] - [ap|ij]),$$

$$A_{ai,bp} = w(2[ai|bp] - [ab|pi]) - F_{ip}^{PV} \delta_{ab},$$

$$B_{ai,bp} = w(2[ai|bp] - [ap|bi]),$$

$$A_{ap,bq} = 2x[ap|bq] - 2y[ab|pq] + w(F_{ab}^{PV} \delta_{pq} - F_{pq}^{PV} \delta_{ab}),$$

$$B_{ap,bq} = 2x[ap|bq] - 2y[aq|bp],$$

$$A_{pi,qj} = (4-4w+2x)[pi|qj] - (2-2w+2y)[pq|ij] + (2-w)(F_{pq}^{DP} \delta_{ij} - F_{ij}^{DP} \delta_{pq}),$$

$$B_{pi,qj} = (4-4w+2x)[pi|qj] - (2-2w+2y)[pj|qi],$$

$$A_{pi,aq} = 2(w-x)[pi|aq] - (w-2y)[pa|qi],$$

$$B_{pi,aq} = 2(w-x)[pi|aq] - (w-2y)[pq|ai] + \frac{1}{2}(2-w)F_{ia}^{DP} \delta_{pq} - \frac{1}{2}wF_{ia}^{PV} \delta_{pq}.$$

The $\mathbf{D}$ matrix is diagonal with the following elements:

$$D_{ai,ai} = 2, \quad D_{pi,pi} = 2-w, \quad D_{ap,ap} = w.$$

The $\mathbf{V}$ matrix elements for TDCPHF calculations are

$$V_{ai} = 2\langle a|\hat{V}|i\rangle, \quad V_{pi} = (2-w)\langle p|\hat{V}|i\rangle, \quad V_{ap} = w\langle a|\hat{V}|p\rangle.$$

Manuscript received 21 June, accepted 16 August 1982