

THE INCORPORATION OF QUADRATIC CONVERGENCE  
INTO OPEN-SHELL SELF-CONSISTENT FIELD EQUATIONS \*

WILLIAM J. HUNT \*\*, WILLIAM A. GODDARD III \*\*\*  
and THOM. H. DUNNING Jr. †

Arthur Amos Noyes Laboratory of Chemical Physics ††,  
California Institute of Technology, Pasadena, California 91109, USA

Received 4 May 1970

The quadratically convergent approach to solving the correct SCF equations for general open-shell systems (with orthogonal orbitals, i.e., multiconfiguration-SCF or Hartree-Fock) is derived and used to discuss other less complicated approaches. Representative calculations on He  $2^1S(1s)(2s)$  and  $H_2O 2^1A_1 [(1a_1)^2(2a_1)^2(1b_1)^2(1b_2)^2(3a_1)(4a_1)]$  are reported.

## 1. INTRODUCTION

In a previous paper [1] we discussed the proper form of the coupling operators in self-consistent field (SCF) equations. We found that the equations have the form

$$\bar{H}_k \phi_k = [H_k - \sum_j^M |j\rangle\langle j| H_j] \phi_k = 0, \quad k = 1, 2, \dots, M, \quad (1)$$

where  $M$  is the number of occupied orthogonal orbitals. Eq. (1) was derived by allowing general changes in  $\phi_k$  and the orthogonality required changes in the other orbitals,  $\{\phi_j\}$ , and then requiring that the first-order change in the energy be zero. The resulting eq. (1) must be satisfied for the exact SCF solutions. In the previous paper we considered solving (1) by the homogeneous approach in which  $\bar{H}_k$  is evaluated for a set of trial functions  $\{\phi_j^0\}$  and the eigenfunction of  $\bar{H}_k$ ,  $\phi_k = \phi_k^0 + \delta\phi_k$ , is taken as a new trial function for the next iteration. We have not found this homogeneous procedure to be satisfactory. From these calculations it was clear that corrections in the  $\bar{H}_k$  due to the changes in the other orbitals  $\{\delta\phi_j\}$  should be included for satisfactory convergence †††. Such a generalized Newton-

Raphson approach of expanding the usual SCF equations in terms of all first-order changes,  $\{\delta\phi_j\}$ , and then solving the resulting equations for the optimum new  $\{\delta\phi_j\}$  had already been suggested by Hinze and Roothaan [2] †. A similar approach had also been used successfully for GI, SOGI, and PGI calculations where it led to quadratic convergence [4].

For a case of  $P$  basis functions and  $M$  orbitals the derivation in refs. [2, 3] leads to a set of  $PM$  coupled equations to be solved for  $M$  vectors  $|\delta i\rangle$ , each having  $P$  components. The vectors  $|\delta i\rangle$  correct the trial functions  $|i^0\rangle$  to produce the optimum SCF vectors  $|i\rangle$ . However the SCF vectors,  $|i\rangle$ , satisfy  $(M^2 + M)/2$  orthonormality relations and thus the  $PM$  equations above cannot be independent. In ref. [1] we derived the SCF equations (1) *without* introducing additional unknown quantities, i.e., Lagrange multipliers. Thus in addition to the orthonormality relations there are  $MP - (M^2 + M)/2$  conditions from application of the variational principle. The straightforward application of (1) gives  $PM$  conditions

††† Mixing of occupied and virtual orbitals may be described as a one-electron variation; it is natural to treat it by the homogeneous approach. Mixing among occupied orbitals requires a two-electron variation. Thus the homogeneous approach seems less appropriate intuitively.

† Huzinaga [3] gives a generalization of the derivation in ref. [2]. A generalization of our derivation would introduce terms like  $\langle i|h|j\rangle$  and  $\langle ij|g|kl\rangle$  in ref. [2]. This would not make the final equations significantly more difficult.

\* Supported in part by a grant (GP-15423) from the National Science Foundation.

\*\* NDEA Fellow.

\*\*\* Alfred P. Sloan Fellow.

† Present address: Battelle Memorial Institute, Columbus, Ohio.

†† Contribution No. 4038.

$$\langle \mu | \overline{H}_k | \phi_k \rangle = 0, \quad \mu = 1, P. \quad (2)$$

Thus some of the relations (2) must be redundant, and we must rearrange (2) to remove these conditions.

In order to do this we transform the basis set  $\{\chi_\mu\}$  so that the new functions are the SCF orbitals,

$$\{\phi_\mu; \mu = 1, 2, \dots, P\} \quad (3)$$

( $1 \leq \mu \leq M$  occupied,  $M < \mu \leq P$  unoccupied). Then (2) becomes

$$\langle \mu | H_k - H_\mu Q_{\mu k} | k \rangle = 0, \quad k = 1, M; \quad \mu = 1, P,$$

where  $Q_{\mu k} = 0$  if  $\mu > M$  and  $Q_{\mu k} = 1$  if  $\mu \leq M$ . (If  $\mu \leq M$  and  $\mu$  and  $k$  belong to the same shell, then  $Q_{\mu k} = 0$  so that  $\phi_k$  and  $\phi_\mu$  are the usual canonical SCF orbitals.) Then the equations

$$\langle \mu | H_k - H_\mu Q_{\mu k} | k \rangle = 0, \quad k = 1, M; \quad \mu > k, \quad (4a)$$

with the relations

$$\langle i | j \rangle = \delta_{ij}, \quad j = 1, M; \quad i > j \quad (4b)$$

completely specify the SCF orbitals<sup>\*†</sup>. We note that the other equations  $\mu \leq k$  are redundant.

If the eqs. (2) are used as in ref. [2] to derive Newton-Raphson equations, the  $|\delta i\rangle$  vectors are not completely determined; the orthonormality equations are necessary to determine the corrections completely. In ref. [2] terms proportional to the Newton-Raphson equations derived from (4b) were added to make the supermatrix,  $\mathbf{G}$ , symmetric. Adding these terms is essential to the solution of the corresponding MC-SCF equations.

Here we use an alternative approach: the corrections  $|\delta i\rangle$  are expanded so that they automatically satisfy (4b) and the Newton-Raphson procedure only involves the variational conditions (4a). The resulting equations are simpler than those of ref. [2], while more general.

## 2. DERIVATION OF THE EQUATIONS

We now expand the occupied and unoccupied Hartree-Fock orbitals in terms of  $P$  orthonormal trial functions (occupied and unoccupied),  $\{\phi_i^0\}$ , and correction coefficients,  $\Delta_{\nu i}$ ,

\* A similar discussion of integrals such as (4a) as SCF conditions for doublets and triplets is given by Lefebvre [5].

† For  $\mu$  and  $k$  in the same shell we delete  $H_\mu$  in (4) to fix  $\phi_k$  and  $\phi_\mu$  as the canonical SCF orbitals. This leaves the energy and the SCF wavefunction unchanged.

$$\phi_i = \phi_i^0 + \sum_{\nu > i}^P \phi_\nu^0 \Delta_{\nu i} - \sum_{\nu < i} \phi_\nu^0 \Delta_{\nu i}^*; \quad i = 1, \dots, P. \quad (5)$$

Thus the orbitals  $\{\phi_i\}$  satisfy (4b) (through first-order in  $\Delta_{\nu i}$ ). We now solve the variational conditions (4a) for the correction coefficients

$$\{\Delta_{\nu i}; \nu > i, i = 1, M\}. \quad (6)$$

This produces the set of equations

$$\begin{aligned} & \left\{ \sum_{\nu > \mu} \Delta_{\nu \mu}^* \langle \nu | H_k - H_\mu | k \rangle - \sum_{\nu < \mu} \Delta_{\mu \nu} \langle \nu | H_k - H_\mu | k \rangle \right. \\ & + \sum_{\nu > k} \Delta_{\nu k} \langle \mu | H_k - H_\mu | \nu \rangle - \sum_{\nu < k} \Delta_{k \nu}^* \langle \mu | H_k - H_\mu | \nu \rangle \\ & \left. + \langle \mu | \Delta H_k - \Delta H_\mu | k \rangle = -\langle \mu | H_k - H_\mu | k \rangle \right\}; \\ & \nu > i; \quad i = 1, M, \end{aligned} \quad (7)$$

where now  $\mu, k$ , etc., denote  $\phi_\mu^0, \phi_k^0$ , etc., and terms of order  $\Delta^2$  and higher have been omitted.

Next we consider the term  $\langle \mu | \Delta H_k - \Delta H_\mu | k \rangle$  in (7). If the total energy can be written as

$$\begin{aligned} E = & \left\{ 2 \sum_{i=1}^M f_i \langle i | h | i \rangle + 2 \sum_{i>j=1}^M (a_{ij} J_{ij} + b_{ij} K_{ij}) \right. \\ & \left. + \sum_{i=1}^M (a_{ii} J_{ii} + b_{ii} K_{ii}) \right\}, \end{aligned} \quad (8)$$

then  $H_i$  has the form

$$H_i = [f_i h + \sum_j (a_{ij} J_j + b_{ij} K_j)] \quad (9)$$

and  $H_k - H_\mu$  has the form

$$\begin{aligned} H_k - H_\mu = & (f_k - f_\mu) h \\ & + \sum_j [(a_{jk} - a_{j\mu}) J_j + (b_{jk} - b_{j\mu}) K_j]. \end{aligned} \quad (10)$$

Thus using (5) and (10) we obtain

$$\langle \mu | \Delta H_k - \Delta H_\mu | k \rangle = \sum_{\nu > j=1}^P (R_{\mu k, \nu j}^1 \Delta_{\nu j} + R_{\mu k, \nu j}^2 \Delta_{\nu j}^*), \quad (11)$$

where

$$\begin{aligned} R_{\mu k, \nu j}^1 = & (a_{jk} - a_{j\mu} - a_{\nu k} + a_{\nu \mu}) \langle \mu j | g | k \nu \rangle \\ & + (b_{jk} - b_{j\mu} - b_{\nu k} + b_{\nu \mu}) \langle \mu j | g | \nu k \rangle. \end{aligned}$$

$$R_{\mu k, \nu j}^2 \equiv (a_{jk} - a_{j\mu} - a_{\nu k} + a_{\nu\mu}) \langle \mu\nu | g | kj \rangle \\ + (b_{jk} - b_{j\mu} - b_{\nu k} + b_{\nu\mu}) \langle \mu\nu | g | jk \rangle \quad (12)$$

and

$$\langle \mu j | g | k\nu \rangle \equiv \int d^3x_1 \phi_\mu^*(1) \phi_k(1) \int d^3x_2 \frac{1}{r_{12}} \phi_j^*(2) \phi_\nu(2).$$

(Note in (12) that  $a_{ij}$  and  $b_{ij}$  are zero if either  $i > M$  or  $j > M$ .) Combining (11) and (7) we obtain

$$\sum_{\nu > j=1}^P (B_{\mu k, \nu j}^1 \Delta_{\nu j} + B_{\mu k, \nu j}^2 \Delta_{\nu j}^*) = -X_{\mu k}, \quad (13)$$

where

$$B_{\mu k, \nu j}^1 \equiv R_{\mu k, \nu j}^1 \\ + \delta_{jk} \langle \mu | H_k - H_\mu | \nu \rangle - \delta_{\mu\nu} \langle j | H_k - H_\mu | k \rangle, \\ B_{\mu k, \nu j}^2 \equiv R_{\mu k, \nu j}^2 \\ + \delta_{\mu j} \langle \nu | H_k - H_\mu | k \rangle - \delta_{\nu k} \langle \mu | H_k - H_\mu | j \rangle, \\ X_{\mu k} \equiv -\langle \mu | H_k - H_\mu | k \rangle. \quad (14)$$

So far we have allowed the orbitals and correction coefficients  $\{\Delta_{\nu j}\}$  to be complex; usually however, the orbitals and correction coefficients can be taken as real, in which case (13) becomes

$$B\Delta = -X, \quad (15)$$

where

$$B = B^1 + B^2. \quad (16)$$

### 3. COMPARISON WITH PREVIOUS APPROACHES

In order to compare the equations derived above with those from previous treatments, we will consider the mixing of two occupied orbitals,  $\phi_1$  and  $\phi_2$ . We will assume that  $\delta\phi_1$  and  $\delta\phi_2$  are small so that

$$\phi_1 = \phi_1^0 + \lambda \phi_2^0, \quad \phi_2 = -\lambda \phi_1^0 + \phi_2^0. \quad (17)$$

From (14) and (15) we have

$$\lambda = \Delta_{21} = \frac{-X_{21}}{B_{21, 21}^1 + B_{21, 21}^2} \\ = \frac{-\langle 2 | H_1 - H_2 | 1 \rangle}{\langle 2 | H_1 - H_2 | 2 \rangle - \langle 1 | H_1 - H_2 | 1 \rangle + R_{21, 21}}, \quad (18)$$

where  $R \equiv R^1 + R^2$  is given by (12) and depends on the  $a$  and  $b$  terms and hence the state under consideration. For He  $2^1S$  we have  $a_{ii} = 0$ ,  $b_{ii} = 0$ ,  $a_{21} = b_{21} = 1$ , hence

$$R_{21, 21} = -2J_{12} - 6K_{12}. \quad (19)$$

Thus

$$\lambda = \langle 2 | J_2 - J_1 | 1 \rangle / [2J_{12} + 4K_{12} - J_{11} - J_{22}], \quad (20)$$

which is the  $\lambda$  which minimizes (8) (through second-order).

From ref. [1] we might consider the operator

$$\bar{H}_1 = H_1 - |2\rangle \langle 2| H_2 \quad (21)$$

and solve for the new (17) by solving the eigenfunctions of (21). The result is

$$\lambda = - \frac{\langle 2 | H_1 - H_2 | 1 \rangle}{\langle 2 | H_1 - H_2 | 2 \rangle - \langle 1 | H_1 | 1 \rangle}, \quad (22)$$

which in comparison to (18) has a term

$$R_{21, 21} + \langle 1 | H_2 | 1 \rangle = \langle 1 | h | 1 \rangle + 2J_{11} - 2J_{12} - 6K_{12}$$

(for He  $2^1S$ ) missing from the denominator.

Levy [6] suggested that

$$\bar{H}_1 = H_1 - |2\rangle \langle 2| G_2^1 - G_2^1 |2\rangle \langle 2|, \quad (23a)$$

$$G_2^1 = \mu H_1 + (1 - \mu) H_2 \quad (23b)$$

(where  $\mu$  is any real number and the  $f_i$  from (9) has been omitted) could be used to solve for the SCF orbitals. In this case

$$\lambda = \frac{-(1 - \mu) \langle 2 | H_1 - H_2 | 1 \rangle}{(1 - 2\mu) \langle 2 | H_1 | 2 \rangle - 2(1 - \mu) \langle 2 | H_2 | 2 \rangle - \langle 1 | H_1 | 1 \rangle}. \quad (24)$$

Since the numerator is proportional to  $\langle 2 | H_1 - H_2 | 1 \rangle$ ,  $\lambda = 0$  in (24) implies that (4a) is satisfied (assuming  $\mu \neq 1$ ). However

$$\bar{H}_2 = H_2 - |1\rangle \langle 1| G_2^1 - G_2^1 |1\rangle \langle 1|$$

leads to an expression for  $\lambda$  which is equivalent

to (24) only when  $\lambda = 0$ . Thus for functions which do not satisfy (4a), we obtain two different mixing criteria and the method of orthogonalization affects the new orbitals to first order in  $\lambda$ .

Equations of the form (23) with  $\mu = 1$  always have  $\lambda = 0$ , leading to no mixing of occupied orbitals [7-11].

Roothaan's two hamiltonian approach also yields two different values of  $\lambda$ . The single hamiltonian formulation [12] predicts a third value of  $\lambda$ . However  $\lambda = 0$  for any of these cases implies that (4) is also satisfied and the solutions are the optimum ones.

Application of this analysis to mixing of a closed-shell orbital with unoccupied orbitals shows that the  $R_{21, 21} = -2K_{21}$  term is omitted and that the self-term  $J_1 - K_1$  should not be present in the denominator when the usual homogeneous approach is used.

We would expect that slower convergence or even divergence should result when these latter expressions are used in place of (18).

#### 4. SIMPLIFICATIONS

The quadratically convergent equations (18) couple together first-order correlation terms involving all orbitals. Solution of these equations is feasible but more complicated than the solution of the usual homogeneous equations. A simpler approach retaining the essential information for good convergence (even if not quadratic) would be useful.

The usual homogeneous approach (OCBSE) [7] works well for mixing of occupied and virtual orbitals. This suggests solving separately for the mixing among occupied orbitals. If couplings between the  $\Delta_{\nu j}$  coefficients are neglected, then (18) may be used to obtain each mixing coefficient  $\nu \leq M$ . The two processes could then each be iterated to convergence alternately until there is no change in either, but in this case convergence is rather slow [13]. However solving both sets of equations at each iteration produces satisfactory convergence. The additional equations (18) require only a small extra computing effort (over that needed for OCBSE) for many cases (for example open-shell doublets and singlets and MC-SCF wavefunctions composed of double occupied orbitals)\*.

\* For example, for a doublet state we form  $J_o$ ,  $K_o$ , and  $H^{CORE} = h + 2J_c - K_c$ , where c and o refer to closed and open shells. This involves the same computational effort as OSBCE [7].

#### 5. EXAMPLES

As an illustration of the importance of including the optimal coupling of occupied orbitals, we will discuss the case of He  $2^1S$  state (configuration  $(1s)(2s)$ ). Birss and Fraga reported a small basis set calculation using their open-shell equations (which do *not* incorporate optimal coupling of occupied orbitals). We started from their wavefunction and optimized the occupied orbitals *without* allowing mixing of the occupied orbitals (i.e., doing just OCBSE) [7]. The result of this is also given in table 1, where we see that this leads to Birss and Fraga's solution. We then applied the equations in section 1 to obtain the results in the third row of table 1. We see that indeed Birss and Fraga had very much the wrong coupling of the occupied orbitals and that their equations did not provide for optimal combinations of the occupied orbitals. Allowing the proper optimization not only greatly lowers the energy (0.0315 h) but it also improves the quality of the wavefunction (as shown by the improvement in the virial ratio,  $V/2E$ ). For comparison we also show in table 2 the results from a calculation using a larger basis set. Of course the Hartree-Fock wavefunction does not give an upper bound on the energy of the  $2^1S$  state (the GI wavefunction has no orthogonality condition and does give an upper bound [16] as observed in table 2).

Previously [4] we reported an OCBSE calculation on the  $2^1A_1[(1a_1)^2(2a_1)^2(1b_1)^2(1b_2)^2(3a_1)(4a_1)]$  state of  $H_2O$ . We have also carried out fully converged HF calculations on this state as reported in table 3.

From these calculations we see that not allowing for optimal mixing of the orbitals can lead to poor results. In OCBSE the final results depend on the trial vectors and hence on the method of choosing the trial vectors. In the  $H_2O$  excited state calculations the excited orbital was chosen as an improved virtual orbital [17] (IVO) from the ground state calculation, which we see was a fairly good choice.

#### 6. SUMMARY

We have derived equations for MC-SCF orbitals which are quadratically convergent. Redundant equations are eliminated while the orthogonality of the solutions is maintained through first order in the correction coefficients. These equations show that the usual methods do not treat mixing of occupied orbitals correctly. We have proposed a simple method, complementary

Table 1  
Comparison of SCF solutions of  $2^1S$  He without and with optimal mixing a)

	$\phi_{1s}$			$\phi_{2s}$		
	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$
Fraga and Birss [9]	1.0627	-0.0611	-0.2208	-0.0378	0.1067	0.9330
OCBSE	1.06276	-0.06112	-0.22080	-0.03781	0.10660	0.93305
FCHF b)	1.01734	0.00066	-0.10154	-0.01537	-0.18574	1.12186

a) All calculations used the Birss and Fraga basis set of three Slater functions:  $\zeta_{1s} = 2.0$ ,  $\zeta_{1s} = 0.8$ , and  $\zeta_{2s} = 0.575$ .  
b) Fully converged HF.

Table 2  
Comparison of energies and properties for  $2^1S$  He using the various wavefunctions

Method	Basis set	Energy	$\epsilon_{1s}$	$\epsilon_{2s}$	$V/2E^c$
Fraga and Birss [9]	A a)	-2.1376	-1.5906	-0.2336	--
OCBSE	A	-2.137589	-1.5906	-0.2336	1.0327
FCHF d)	A	-2.169162	-1.7294	-0.1843	0.9972
FCHF d)	B b)	-2.169845	-1.7154	-0.1869	1.0003
GI b)	B	-2.141180	-1.7404	-0.1412	1.0080
Exper. [15]		-2.1460			

a) See footnote a), table 1.  
b) The basis set consists of five Slater functions:  $\zeta_{1s} = 1.4642$ ,  $\zeta_{2s} = 2.9430$ ,  $\zeta_{2s} = 1.8698$ ,  $\zeta_{2s} = 0.6114$ , and  $\zeta_{2s} = 1.5746$  (ref. [14]).  
c) The virial ratio, this should be 1.0 for an optimally scaled wave.  
d) Fully converged HF.

Table 3  
Comparison of OCBSE and fully converged HF functions for some states of  $H_2O$  and  $H_2O^+$

	State	
	$2^1A_1 H_2O$	$1^2A_1 H_2O^+$
Energy		
OCBSE	-75.6992	-75.5348
FCHF a)	-75.7050	-75.5357
$\epsilon_{4a_1}$		
OCBSE	- 0.1653	- 0.1833
FCHF a)	- 0.1769	- 0.1832

a) FCHF denotes fully converged HF.

to the OCBSE method, which requires little extra effort, is of great generality, and has good convergence properties. We also note that some treatments predict two different values for the mixing of two occupied orbitals and may lead to difficulties.

## REFERENCES

- [1] W. A. Goddard III, T. H. Dunning Jr. and W. J. Hunt, Chem. Phys. Letters 4 (1969) 231.
- [2] J. Hinze and C. C. J. Roothaan, Progr. Theoret. Phys. (Kyoto) 40 (1967) 37.
- [3] S. Huzinaga, Progr. Theoret. Phys. (Kyoto) 41 (1969) 307.
- [4] W. A. Goddard III, R. C. Ladner and D. L. Huestis, to be published.
- [5] R. Lefebvre, in: Modern quantum chemistry, Vol. 1, ed. O. Sinanoğlu (Academic Press, New York, 1965) p. 125.
- [6] B. Levy, J. Chem. Phys. 48 (1968) 1994.
- [7] W. J. Hunt, T. H. Dunning Jr. and W. A. Goddard III, Chem. Phys. Letters 3 (1969) 606.
- [8] G. Das and A. C. Wahl, J. Chem. Phys. 44 (1966) 87.
- [9] S. Fraga and F. W. Birss, J. Chem. Phys. 40 (1964) 3203.
- [10] F. W. Birss and S. Fraga, J. Chem. Phys. 38 (1963) 2552.
- [11] A. Veillard and E. Clementi, Theoret. Chim. Acta 7 (1967) 133.
- [12] C. C. J. Roothaan, Rev. Mod. Phys. 32 (1960) 179.
- [13] G. Das, J. Chem. Phys. 46 (1967) 1568.
- [14] S. L. Guberman and W. A. Goddard III, unpublished.
- [15] J. C. Slater, Quantum theory of atomic structure (McGraw-Hill, New York, 1960) Vol. 1.
- [16] W. A. Goddard III, Phys. Rev. 176 (1968) 106.
- [17] W. J. Hunt and W. A. Goddard III, Chem. Phys. Letters 3 (1969) 414.