Non Self-Consistent Field Theory

II. A Modification of Roothaans Method

T. K. Lim and M. A. Whitehead

Quantum Chemistry Laboratory*, Chemistry Department, McGill University, Montreal 2, Quebec, Canada

Roothaan's SCF method [2] is reformulated so that two non-SCF methods are developed to solve the eigenequation. The results from these methods can be used as starting eigenfunctions for Roothaan's SCF method.

La méthode SCF de Roothaan [2] est reformulée en développant deux méthodes non-SCF pour la solution de l'équation aux valeurs propres. Les résultats de ces deux méthodes peuvent être utilisées comme fonctions de départ pour la méthode SCF de Roothaan.

Die „self-consistent field" Methode von Roothaan [2] wird neu formuliert: dabei werden zwei Nicht-SCF Methoden zur Lösung der Eigenwertgleichung entwickelt. Deren Resultate können als Anfangsfunktionen bei Rechnungen mit der SCF-Methode von Roothaan benutzt werden.

I. Introduction

In a previous paper [1] three non-SCF methods were proposed, which required optimizing all the molecular orbitals, the occupied MO's C^o and the virtual unoccupied MO's C^u. These methods all required one orthonormalization, two diagonalizations and two sorting procedures.

In this paper, based on the same idea as [1], Roothaan's SCF method [2] is reformulated to give an R^o-independent Hamiltonian matrix F^K so that a direct non-SCF solution is possible. The result of the present methods can be used as the initial guess for Roothaan's SCF method, so that the difficulty of guessing the input coefficient matrix can be avoided. Also, one of the new methods requires only one orthonormalization of the basis set, one diagonalization of the new F^K, and one sorting procedure for the eigenvalues and their corresponding eigenvectors, and is thus more efficient than the earlier methods [1]. However, this method is limited to special cases [3].

II. Theory

The one-electron spatial molecular orbitals ϕ in the LCAO approximation can be expressed by

$$\phi = \mathscr{S}\, C \tag{1}$$

---

* The title Theoretical Chemistry has been transferred to the Division of Physical Chemistry.

where $\mathscr{S} = (\mathscr{S}_1\,\mathscr{S}_2\ldots\mathscr{S}_m)$ is a vector of m non-orthogonal basis atomic orbitals, $\boldsymbol{\phi} = (\phi_1\phi_2\ldots\phi_m)$ is a vector of m MO's, and

$$\boldsymbol{C} = [\boldsymbol{C}^o \mid \boldsymbol{C}^u]\,. \tag{2}$$

$\boldsymbol{C}^o$ is a $m \times n$ coefficient matrix corresponding to the n occupied MO's $\phi_1, \phi_2, \ldots, \phi_n$, and $\boldsymbol{C}^u$ is a $m \times (m-n)$ coefficient matrix corresponding to the $(m-n)$ virtual MO's $\phi_{n+1}, \ldots, \phi_m$. Eq. (1) can be rewritten as

$$\boldsymbol{\phi} = \mathscr{S}'\,\boldsymbol{A}^{-1}\,\boldsymbol{A}\,\boldsymbol{C}' = \mathscr{S}'\,\boldsymbol{C}' \tag{1a}$$

using the orthonormalization process previously described [1]; the prime designates orthonormalized quantities, and $\boldsymbol{A}$ is a non-singular matrix [1].

The m orbital energy sum, ε, can be expressed as

$$\varepsilon = \mathrm{Tr}(\boldsymbol{R}\boldsymbol{F}^G) \tag{3}$$

where [1]

$$\boldsymbol{R} = \boldsymbol{C}\,\boldsymbol{C}^\dagger = \boldsymbol{C}^o\,\boldsymbol{C}^{o\dagger} + \boldsymbol{C}^u\,\boldsymbol{C}^{u\dagger} = \boldsymbol{R}^o + \boldsymbol{R}^u \tag{4}$$

$$\boldsymbol{F}^G = \boldsymbol{H} + \boldsymbol{G}(\boldsymbol{R}^o) \tag{5}$$

and

$$G_{\mu\nu} = \sum_{\lambda,\sigma=1}^{m} R_{\lambda\sigma}^o\,[2\,\langle\mu\nu\mid\sigma\lambda\rangle - \langle\mu\lambda\mid\sigma\nu\rangle]\,.$$

Notice that $\boldsymbol{C}'$ is a unitary matrix, hence,

$$\boldsymbol{C}'\,\boldsymbol{C}'^\dagger = \boldsymbol{C}'^\dagger\,\boldsymbol{C}' = 1_m\,,$$

and

$$\mathrm{Tr}(\boldsymbol{T}\,\boldsymbol{G}) = \mathrm{Tr}(\boldsymbol{R}^o\,\boldsymbol{K}) \tag{6}$$

where $\boldsymbol{T}$ and $\boldsymbol{K}$ are two new matrices. [Eqs. (19) and (20) of [1]]. Thus,

$$\begin{aligned}
\varepsilon &= \mathrm{Tr}(\boldsymbol{R}\,\boldsymbol{H} + \boldsymbol{C}^\dagger\,\boldsymbol{G}\,\boldsymbol{C})\\
&= \mathrm{Tr}(\boldsymbol{R}\,\boldsymbol{H} + \boldsymbol{C}'^\dagger\,\boldsymbol{A}^\dagger\,\boldsymbol{G}\,\boldsymbol{A}\,\boldsymbol{C}')\\
&= \mathrm{Tr}(\boldsymbol{R}\boldsymbol{H} + \boldsymbol{A}^\dagger\,\boldsymbol{G}\,\boldsymbol{A}\,\boldsymbol{C}'\,\boldsymbol{C}'^\dagger)\\
&= \mathrm{Tr}(\boldsymbol{R}\,\boldsymbol{H} + \boldsymbol{T}\,\boldsymbol{G}) \qquad\qquad (7)\\
&= \mathrm{Tr}(\boldsymbol{R}\,\boldsymbol{H} + \boldsymbol{R}^o\,\boldsymbol{K})\\
&= \mathrm{Tr}[(\boldsymbol{R}^o + \boldsymbol{R}^u)\,\boldsymbol{H} + \boldsymbol{R}^o\,\boldsymbol{K}]\\
&= \mathrm{Tr}(\boldsymbol{R}^o\,\boldsymbol{F}^K + \boldsymbol{R}^u\,\boldsymbol{H})\\
&= \mathrm{Tr}(\boldsymbol{C}^{o\dagger}\,\boldsymbol{F}^K\,\boldsymbol{C}^o) + \mathrm{Tr}(\boldsymbol{C}^{u\dagger}\,\boldsymbol{H}\,\boldsymbol{C}^u)
\end{aligned}$$

where

$$\boldsymbol{F}^K = \boldsymbol{H} + \boldsymbol{K}(\boldsymbol{T})\,. \tag{8}$$

The application of the variation method to $\boldsymbol{C}^o$, keeping $\boldsymbol{C}^u$ constant, with $\boldsymbol{C}^o$ orthonormalized

$$\boldsymbol{C}^{o\dagger}\,\boldsymbol{S}\,\boldsymbol{C}^o = 1_n \tag{9}$$

gives

$$\delta\varepsilon = \mathrm{Tr}[(\delta\boldsymbol{C}^{o\dagger})\,\boldsymbol{F}^K\,\boldsymbol{C}^o + \boldsymbol{C}^{o\dagger}\,\boldsymbol{F}^K(\delta\boldsymbol{C}^o)] = \mathrm{Tr}[(\delta\boldsymbol{C}^{o\dagger})\,\boldsymbol{F}^K\,\boldsymbol{C}^o + (\delta\widetilde{\boldsymbol{C}}^o)\,\boldsymbol{F}^{K*}\,\boldsymbol{C}^{o*}] = 0 \tag{10}$$

and

$$\delta(C^{o\dagger} \, S \, C^o) = (\delta C^{o\dagger}) \, S \, C^o + C^{o\dagger} \, S \, (\delta C^o) = (\delta C^{o\dagger}) \, S \, C^o + (\delta \widetilde{C^o}) \, S^* \, C^{o*} = 0 \;. \quad (11)$$

Multiplying Eq. (11) by $-\varepsilon^o$, a multiplier matrix, and adding it to Eq. (10) gives,

$$\delta\varepsilon' = \mathrm{Tr}[(\delta C^{o\dagger}) \, (F^K \, C^o - S \, C^o \, \varepsilon^o) + (\delta \, \widetilde{C^o}) \, (F^{K*} \, C^{o*} - S^* \, C^{o*} \, \varepsilon^o)] = 0 \;.$$

The superscripts $*, \sim, \dagger$ denote the complex conjugate, transpose, and hermitian respectively. Hence for any variation of $\delta C^{o\dagger}$ and δC^o,

$$F^K \, C^o = S \, C^o \, \varepsilon^o \qquad\qquad (12)$$

and

$$F^{K*} \, C^{o*} = S^* \, C^{o*} \, \varepsilon^o \;. \qquad\qquad (13)$$

Taking the complex conjugate of Eq. (13) and comparing it with Eq. (12), gives

$$\varepsilon^{o*} = \varepsilon^o$$

i.e., ε^o is hermitian, and thus it follows that Eqs. (12) and (13) are equivalent.

Consequently, eigenequation (12) can be solved directly without an iterative procedure because F^K is a R^o-independent Hamiltonian, depending only on T, which can be obtained by orthonormalizing the basis set before the calculation starts.

III. Methods of Calculation

It has been shown that Roothaan's eigenequation [2]

$$F^G \, C^o_R = S \, C^o_R \, \varepsilon^o_R \qquad\qquad (14)$$

can be replaced by eigeneq. (12) if both the virtual and occupied one electron molecular orbitals are considered; in Eq. (14)

$$F^G = H + G(R^o) \qquad\qquad (15)$$

and the subscript R signified Roothaan's method. When the eigenvector C^o has been obtained from Eq. (12), (F^K does not represent the true Hamiltonian) the *true* eigenvector and eigenvalue C^o_T and ε^o_T can be obtained by either of the following methods.

Method I: The Equivalent Method

It is assumed that

$$C^o_T \simeq C^o \qquad\qquad (16)$$

which is similar to the assumption that,

$$C^o_T \simeq C^o_R$$

in Roothaan's method, where C^o_R is obtained from minimizing the ground state energy [Eq. (14)]. Therefore,

$$R^o_T \simeq C^o \, C^{o\dagger} \qquad\qquad (17)$$

$$\varepsilon^o_T \simeq C^{o+} \, F^G \, C^o \qquad\qquad (18)$$

and the ground state energy E_e is

$$E_e = 2\mathrm{Tr}(\boldsymbol{R_T^0}\, \boldsymbol{H}) + \mathrm{Tr}[\boldsymbol{R_T^0}\, \boldsymbol{G}(\boldsymbol{R_T^0})]\,. \tag{19}$$

Thus, after solving $\boldsymbol{C^o}$ from Eq. (12), $\boldsymbol{R_T^o}$ is approximated from Eq. (17), from which $\boldsymbol{F^G}$ can be constructed, and then ε_T^o calculated using Eq. (18).

Method II: The Modification Method

In this method the $\boldsymbol{C^o}$ obtained from Eq. (12) is regarded as the first approximation to $\boldsymbol{C_T^o}$. Then $\boldsymbol{R^o}$ and $\boldsymbol{F^G}$ are constructed directly so that the $\boldsymbol{C_T^o}$ and ε_T^o can be calculated from the eigeneq. (20).

$$\boldsymbol{F^G}\, \boldsymbol{C_T^o} = \boldsymbol{S}\, \boldsymbol{C_T^o}\, \varepsilon_T^o\,. \tag{20}$$

Here it is *assumed* that the eigenvector calculated from Eq. (12), if modified to satisfy the condition of the minimization of ground state energy, would be the true eigenvector $\boldsymbol{C_T^o}$.

The solution of Eq. (20) is equivalent to solving

$$\boldsymbol{F^{G'}}\, \boldsymbol{C_T^{o}}' = \boldsymbol{C_T^{o}}'\, \varepsilon_T^o \tag{21}$$

where

$$\boldsymbol{F^{G'}} = \boldsymbol{S^{-1/2}}\, \boldsymbol{F^G}\, \boldsymbol{S^{-1/2}} \tag{22}$$

and

$$\boldsymbol{C_T^o} = \boldsymbol{S^{-1/2}}\, \boldsymbol{C_T^{o}}' \tag{23}$$

where $\boldsymbol{F^{G'}}$ is diagonalized to produce $\boldsymbol{C_T^{o}}'$ and ε_T^o, so that $\boldsymbol{C_T^o}$ can be calculated from Eq. (23).

Since Eq. (20) is exactly the same as the eigenequation of the SCF method, it would be expected that continuous iteration of Eq. (20) would converge to the SCF result.

The validity of this method is discussed in Section V.

IV. Illustrative Examples

In order to compare the results with those in [1], cis- and trans-butadiene are again used as examples. The geometry of these compounds was given and the SCF calculations were described in [1], so only the results are given for comparison.

The method is again tested for the NNDO (no neglect of differential overlap) and the KNDO ($\boldsymbol{K}$ matrix neglect of differential overlap) cases.

In the non-SCF calculation, the $\boldsymbol{T}$ and $\boldsymbol{K}$ matrices are the same as defined in [1] for both NNDO and KNDO cases.

The other quantities in Tabs. 1 and 2 such as E_e, E_N, E_n, ε_i, $\varDelta E$, f etc., all have the same meaning as in [1]. MRT stands for the results from the Modification of Roothaan's Theory.

Table 1. *Comparison of MRT and SCF Results of Trans-butadiene*

	NNDO			KNDO			Obs.
	MRT-(I)	MRT-(II)	SCF	MRT-(I)	MRT-(II)	SCF	
E_e	-87.8379	-89.3646	-89.3970	-107.2719	-107.9260	-107.9260	
$E_N = E_e + E_n$	-83.9670	-85.4946	-85.5270	-103.4019	-104.0560	-104.0560	
ε_1	-10.6343	-11.0876	-10.3781	-21.6264	-21.8025	-21.5472	
$\varepsilon_2 = -I$	-7.5975	-7.9934	-7.0342	-14.6918	-14.8411	-14.5143	-9.0
$\varepsilon_3 = -A$	-2.5201	-2.0668	-1.3951	-1.6877	-1.5117	-1.0317	
ε_4	0.4854	0.8813	2.0442	9.2546	9.4039	10.0086	
R_{ab}	1.3796	1.3697	1.3715	1.3746	1.3729	1.3729	1.337
R_{bc}	1.4788	1.4860	1.4727	1.4648	1.4650	1.4650	1.483
$f(N \to V_{1a})$	0.2579	0.2217	0.3194	0.7375	0.7564	0.7652	0.53
$f(N \to V_{23})$	0.7413	1.2249	1.3157	3.0149	3.1944	3.2290	
$\Delta E(N \to V_{23})$	4.0186	6.0015	5.9375	13.4768	13.7148	13.8672	6.0
$\Delta E(N \to V_{24})$	3.8181	6.3889	6.3683	20.1114	21.5153	21.7941	7.2
E_N (cis) $-$ E_N (trans)	-0.1181	0.1513	0.1133	0.3338	0.1620	0.1627	~ 0.1

Table 2. *Comparison of MRT and SCF Results of Cis-butadiene*

	NNDO			KNDO			SCF
	MRT-(I)	MRT-(II)	SCF	MRT-(I)	MRT-(II)	SCF	
E_e	-88.9260	-90.1833	-90.2537	-108.5757	-109.0580	-109.0587	
$E_N = E_e + E_n$	-84.0860	-85.3433	-85.4137	-103.7357	-104.2180	-104.2187	
ε_1	-10.5769	-10.9778	-10.4318	-21.8013	-21.9315	-21.7349	
$\varepsilon_2 = -I$	-7.2179	-7.6057	-6.9879	-14.6741	-14.7931	-14.5799	
$\varepsilon_3 = -A$	-2.1646	-1.7638	-1.3176	-1.6276	-1.4974	-1.1915	
ε_4	0.7889	1.1766	2.1211	10.8271	9.9460	10.4103	
R_{ab}	1.3773	1.3695	1.3709	1.3733	1.3718	1.3719	
R_{bc}	1.4773	1.4906	1.4757	1.4683	1.4695	1.4693	
$f(N \to V_{1a})$	0.0002	0.0001	0.0020	0.0090	0.0131	0.0133	
$f(N \to V_{23})$	0.3939	0.5617	0.6259	1.4788	1.5478	1.5605	
$\Delta E(N \to V_{23})$	4.1074	5.6652	5.6689	13.2016	13.3557	13.4600	
$\Delta E(N \to V_{24})$	4.0285	6.4353	6.4393	20.7712	21.9506	22.1704	

V. Discussion

In order to discuss the foundation and significance of the non-SCF theory, it is necessary to understand the relation between the Slater determinantal wave function, the Hartree-Fock wave function, and the non-SCF wave function. The ground state Slater determinantal wave function $^1\Psi_0$ of a molecule with $2n$ electrons (assumed to be a closed shell) can be expressed as

$$^1\Psi_0 = \frac{1}{\sqrt{2n}\,!}\det \mid \phi_1(1)\ \alpha(1)\ \phi_1(2)\ \beta(2)\ldots\phi_n(2n-1)\ \alpha(2n-1)\ \phi_n(2n)\ \beta(2n)\mid. \quad (24)$$

If one of the electrons, $\phi_i(k)$, is removed from the molecule, then the wave function $^2\Psi_i$ is of the form

$$^2\Psi_i = \frac{1}{\sqrt{(2n-1)!}}\det \mid \phi_1(1)\ \alpha(1)\ \phi_1(2)\ \beta(2)\ldots\phi_{i-1}(2i-3)\ \alpha(2i-3)$$

$$\phi_{i-1}(2i-2)\ \beta(2i-2)\ \phi_i(2i) \left\{\begin{array}{l} \alpha(2i) \\ \beta(2i) \end{array}\right\}.$$

$$\phi_{i+1}(2i+1)\ \alpha(2i+1)\ \phi_{i+1}(2i+2)\ \beta(2i+2)\ldots$$

$$\ldots\phi_n(2n-1)\ \alpha(2n-1)\ \phi_n(2n)\ \beta(2n)\mid. \quad (25)$$

The ground state energy E_e and the orbital energy ε_i are then

$$E_e = \int {}^1\Psi_0^*\ H\ {}^1\Psi_0\ d\tau \quad (26)$$

and

$$\varepsilon_i = E_e - \int {}^2\Psi_i^*\ H\ {}^2\Psi_i\ d\tau = C_i^\dagger\ F\ C_i \quad (27)$$

where C_i is the ith column of C. It is understood that both the Hartree-Fock-Roothaan method and the non-SCF method are based on the assumption that Eqs. (24) and (25) are the best approximation to the true wave function, so that the problem becomes one of calculating good ϕ_i's. If all the semi-empirical methods and the Hückel method are regarded as special cases of the Hartree-Fock-Roothaan method, then the Hartree-Fock-Roothaan method is the only method so far known that can be used for quantum mechanical molecular calculation of the ϕ_i's, and all results of quantum mechanical calculations must approximate or converge to the Hartree-Fock-Roothaan result. (Extensions such as the extended Hartree-Fock method etc., are not considered).

In fact, the Hartree-Fock-Roothaan equation which is the *condition* of ground state energy minimization, is only one of the possible equations that can be used to calculate the ϕ_i's. Although it cannot be disputed that the minimization of the ground state energy does produce *good* ϕ_i's, nevertheless any other theoretically justifiable methods that could produce *good* ϕ_i's, as evidenced by agreement with experimental results, must also be accepted even if their results are not the same as the Hartree-Fock-Roothaan method. Thus it is important to note that the criteria of whether the ϕ_i's which are calculated from any method are correct or not depends on the comparison of *all* the physical properties about which information can be obtained from the ϕ_i's and experimental results. The ground state energy is only one of these physical properties.

With this idea in mind, a non-SCF theory was developed to give good ϕ_i's by optimizing all the orbital energies, as shown in Eq. (27), rather than by minimizing the ground state energy as shown in Eq. (26). From the theoretical background it can be expected that the SCF method will produce a more accurate ground state energy, while the non-SCF method will produce more accurate orbital energies. From the calculated results it can be seen that the SCF method which removes all the doubly counted electronic repulsions cannot predict the orbital energies very well, while the non-SCF method which counts the electronic

repulsion twice seems unable to predict the ground state energy as well as the SCF method. Therefore, is it possible to effect a compromise between these two extremes ? The *modification method* does so: first, the ϕ_i's are calculated by optimizing all the orbital energies [Eq. (27)], and then modified so as to satisfy the condition of the minimization of ground state energy [Eq. (26)]. (i.e., to satisfy the Hartree-Fock equations or Roothaan's eigenequation in the LCAO form).

It could also be expected that the use of the non-SCF results as the initial guess in the SCF method, would result in rapid convergence to the SCF results. The iterations do converge to the SCF results, but from our calculations it seems that the improvement in efficiency is not so great as expected, because in the first iteration the orbital energies suddenly drop considerably, so that several iterations are required to adjust it back to the initial value. Nevertheless, it is expected that other cases with larger Hamiltonian matrices will give the increased efficiency expected.

Although one additional diagonalization time can be saved in method I as compared to method II, method I suffers because,

1. F^K has more electronic repulsion terms than F^G, which causes a slightly higher ground state energy, and

2. the matrix product $C^{o\dagger} F^G C^o$ is only approximately diagonal due to the fact that F does not commute with F^K, and C^o is calculated from F^K.

Method II takes C^o as the first approximation, so that the true eigenvector C_T^o can be obtained from the assumption that C^o must also satisfy Roothaan's Eq. (14). This eliminates the disadvantages mentioned above and gives nearly the same results as the SCF method. The results of the present method lie between those of the SCF method and those of the non-SCF methods of [1], but are closer to those of SCF method.

References

1. LIM, T. K., and M. A. WHITEHEAD: Theoret. chim. Acta **7**, 48 (1967).
2. ROOTHAAN, C. C. J.: Revs. Mod. Phys. **23**, 69 (1951).
3. LIM, T. K., and M. A. WHITEHEAD: Unpublished results.

Prof. M. A. WHITEHEAD
Quantum Chemistry Laboratory,
Chemistry Department
McGill University
Montréal 2, Quebec, Canada