Theoret. Chim. Acta (Berl.) 46, 325–329 (1977)

THEORETICA CHIMICA ACTA

Short Communications

On the "Level-shifting" Method for Converging Hartree–Fock Wave Functions

Prem Kishore Mehrotra*

Department of Chemistry, Cornell University, Ithaca, New York 14853, USA

A new formulation of "level-shifting" method is presented, which provides a more economical procedure for computations.

Key words: SCF convergence and level-shifting

In a recent study, Saunders and Hillier [1] have presented a "level-shifting" method for converging Hartree–Fock wave functions when solving the Roothaan's self-consistent-field (SCF) equations [2] for single determinantal configurational wave functions. The purpose of the present paper is to comment on the "level-shifting" procedure as formulated in the papers of Saunders and Hillier [1], and then to suggest an alternative procedure which requires less computer-storage and is more efficient for computations.

We consider here only the Roothaan's closed-shell procedure, but the above method can also be extended to open-shell systems using either of the restricted-Hartree–Fock (RHF) [3] or unrestricted-Hartree–Fock (UHF) [4] versions. We assume that the calculations are to be performed in the basis of n-linearly independent real atomic orbitals (AO's). Moreover, with no further loss of generality; the AO's are assumed to constitute an orthonormal set. In a typical cycle of Roothaan's iterative procedure [2]; say for a $2m_1$ electron-system, we construct the m_1-orthonormal trial doubly occupied molecular orbitals (DOMO's), and $m_2 = n - m_1$ virtual molecular orbitals (VMO's) as linear combination of n-AO's; such that all n-MO's form an orthonormal set. From the trial DOMO's the Fock matrix over the AO basis ($\boldsymbol{F}^{ao}$) is constructed, and $\boldsymbol{F}^{ao}$ is diagonalized[1] to yield a new set of MO coefficient matrix $\boldsymbol{C}$;

$$\boldsymbol{F}^{ao}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{\Lambda} \tag{1}$$

---

* Correspondences after Sept. 1, 1977, should be addressed to Chemistry Department, Hunter College of the City University of New York, 695 Park Avenue, New York, N.Y. 10021.

[1] In a diagonalization, it is implicit here that the eigenvalues (and the corresponding eigenvectors) have been arranged according to the descending order in energy.

the diagonal matrix "Λ" produces the eigenvalues. From the new n-orthonormal MO's; m_1-DOMO's are chosen corresponding to the m_1-vectors of the lowest eigenvalues, and the whole procedure is repeated until the convergence is achieved to a pre-specified threshhold.

We are at liberty to perform the diagonalization of Eq. (1) in the basis of trial MO's, i.e., for the trial MO coefficient matrix C^{i-1} (where the superscript denotes the iteration cycle number); Eq. (1) can be transformed to:

$$(\boldsymbol{C}^{i-1})^{+}\boldsymbol{F}^{\mathrm{ao}}\boldsymbol{C}^{i-1}(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i}=(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i}\boldsymbol{\Lambda} \tag{2}$$

Equation (2) can further be written as:

$$\boldsymbol{F}^{\mathrm{mo}}\boldsymbol{C}'=\boldsymbol{C}'\boldsymbol{\Lambda} \tag{3}$$

where

$$\boldsymbol{F}^{\mathrm{mo}}=(\boldsymbol{C}^{i-1})^{+}\boldsymbol{F}^{\mathrm{ao}}\boldsymbol{C}^{i-1} \tag{4}$$

and

$$\boldsymbol{C}'=(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i} \tag{5}$$

In Eqs. (2–5), the matrix $\boldsymbol{C}^i$ is the MO coefficient matrix at iteration cycle i. Eq. (3) is equivalent to Eq. (2). Now the $\boldsymbol{C}'$ matrix yields the MO coefficient matrix in the basis of trial MO coefficient matrix C^{i-1}, instead of AO basis as in Eq. (1). The new MO coefficient matrix in the AO basis can easily be obtained from Eq. (5) by the following transformation:

$$\boldsymbol{C}^{i}=\boldsymbol{C}^{i-1}\boldsymbol{C}' \tag{6}$$

The Roothaan's iterative procedure need not always converge, it can oscillate or diverge. Several aspects of the convergence problem are analysed by Saunders and Hillier [1], and a "level-shifting" procedure is presented to ensure the convergence. The key steps of this method are as follows:

(1) From $\boldsymbol{F}^{\mathrm{ao}}$ compute $\boldsymbol{F}^{\mathrm{mo}}$ according to Eq. (4).

(2) To $\boldsymbol{F}^{\mathrm{mo}}$, a pre-selected level-shift matrix "Δ" of order $n\times n$ is added. This matrix is of special form: it is a diagonal matrix, moreover the first "m_1" (a number reminiscent of number of the DOMO's) diagonal elements are zero, and remaining "m_2" diagonal elements are all equal to a single positive level-shift parameter "x". In a matrix notation "Δ" can be represented as follows:

$$\begin{array}{cc} & m_1\times m_1 \qquad\qquad \\ \Delta= & \left(\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & \begin{matrix} x & & & \\ & x & 0 & \\ & & \ddots & \\ & 0 & & x \end{matrix} \end{array}\right) \\ & \qquad\qquad m_2\times m_2 \end{array}$$

(3) The resulting level-shift matrix is then diagonalized,

$$(\boldsymbol{F}^{\mathrm{mo}}+\boldsymbol{\Delta})\boldsymbol{C}'=\boldsymbol{C}'\boldsymbol{\Lambda} \tag{7}$$

to yield a new set of MO coefficient matrix in the basis of trial coefficient matrix $\boldsymbol{C}^{i-1}$. The MO coefficient matrix $\boldsymbol{C}^{i}$ in the basis of AO's is then traced back from Eq. (6). The steps 1, 2 and 3 are repeated until the SCF convergence is achieved. The judicious choice of the level-shift parameter "x" at different cycles ensures the convergence on most occasions.

The above formalism of the "level-shifting" procedure is however not the most efficient one, when it is documented in the computer programs for performing MO calculations. For closed-shell molecules, the additional space of three triangles ($n(n+1)/2$ elements) is required. Of this, a $n\times n$ array is used for storing the MO coefficient matrix $\boldsymbol{C}^{i-1}$, which is needed for transforming $\boldsymbol{F}^{\mathrm{ao}}$ to $\boldsymbol{F}^{\mathrm{mo}}$ according to Eq. (4), and retrieval of $\boldsymbol{C}^{i}$ in Eq. (6). The remaining space is used for storing intermediate results during the multiplication of three matrices in the Eq. (4). This extra space will be increased by a factor of two for the UHF MO calculations [4]. The constraints of the storage of computer often force the use of auxiliary disc-storages slowing down the speed of the calculations. Saunders and Hillier point out that the transformation of $\boldsymbol{F}^{\mathrm{ao}}$ to $\boldsymbol{F}^{\mathrm{mo}}$ prior to the diagonalization has obvious numerical advantages if a Jacobi procedure is used for the diagonalization, because $\boldsymbol{F}^{\mathrm{mo}}$ will be of pseudo-diagonal form. Our experience, however, tells us that often, especially dealing with large matrices of size 40–80, the step 1 in the "level-shifting" procedure where three matrices are multiplied is several orders more time consuming than a simple diagonalization, if the Givens' procedure [5] is followed for diagonalization. Thus, the above formulation of the "level-shifting" method not only requires the extra storage, but as well impairs the speed of the calculations.

Equation (7), however, can be transformed into another form which is more suitable for the numerical computations. On substituting Eqs. (4) and (5) in Eq. (7), Eq. (7) transforms to:

$$\begin{gathered}(\boldsymbol{F}^{\mathrm{mo}}+\boldsymbol{\Delta})\boldsymbol{C}'=\boldsymbol{C}'\boldsymbol{\Lambda}\\ ((\boldsymbol{C}^{i-1})^{+}\boldsymbol{F}^{\mathrm{ao}}\boldsymbol{C}^{i-1}+\boldsymbol{\Delta})(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i}=(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i}\boldsymbol{\Lambda}\end{gathered} \tag{8}$$

$$(\boldsymbol{F}^{\mathrm{ao}}+\boldsymbol{\Delta}')\boldsymbol{C}^{i}=\boldsymbol{C}^{i}\boldsymbol{\Lambda} \tag{9}$$

where

$$\boldsymbol{\Delta}'=\boldsymbol{C}^{i-1}\boldsymbol{\Delta}(\boldsymbol{C}^{i-1})^{+} \tag{10}$$

In the derivation of Eq. (9) from (10), the orthonormality property of the MO's has been used,

$$(\boldsymbol{C}^{i-1})^{+}\boldsymbol{C}^{i-1}=\boldsymbol{C}^{i-1}(\boldsymbol{C}^{i-1})^{+}=\mathbf{1} \tag{11}$$

In Eq. (9), Roothaan's iterative procedure [2] has been recaptured in its conventional form. The present formalism of the "level-shifting" procedure avoids the time consuming step of transforming $\boldsymbol{F}^{\mathrm{ao}}$ into $\boldsymbol{F}^{\mathrm{mo}}$, as in Eq. (4). Moreover, retrieval of $\boldsymbol{C}^i$ according to Eq. (6) is no more needed, because the new MO coefficient matrix is already in the basis of AO's. The extra step performed is, of course, the transformation of Δ into Δ' according to Eq. (10), the multiplication of three matrices again. But due to the very special form of Δ, the transformation of Eq. (10) is extremely efficient, does not require storage of intermediate results. The only additional storage now needed is extra space for storing $n(n+1)/2$ elements of Δ' (Δ' is symmetric, hence only one triangle need be stored). We, no more, need to store $\boldsymbol{C}^{i-1}$. Thus, the above formulation of the "level-shifting" method needs only the additional storage of one triangle instead of three triangles, as in the formalism of Saunders and Hillier. It avoids the time consuming transformation of $\boldsymbol{F}^{\mathrm{ao}}$ to $\boldsymbol{F}^{\mathrm{mo}}$. The savings in time and core is much more pronounced in case of the UHF MO calculations. This gain in computational efficiency is however not without some loss in flexibility of the original formulation. For example, the damping procedure of Hartree [6] can also be incorporated along with the level-shifting in the formalism of Saunders and Hillier, in the form of a "damp factor", when $\boldsymbol{F}^{\mathrm{ao}}$ is transformed to $\boldsymbol{F}^{\mathrm{mo}}$. This cannot be done in the present method. Usually, the value of the "damp factor" is taken as unity, which tantamounts to no damping at all [1, 3]. However, this damping option can be of importance especially when studying excited states. The only way to include this feature in the present method is to follow the original procedure of Hartree [6]. Another advantage of transforming $\boldsymbol{F}^{\mathrm{ao}}$ to the trial MO basis is that the block of the Fock operator connecting the occupied with virtual orbitals is available, which can be used to provide a measure of SCF convergence. This criterion of convergence in certain circumstances is advantageous over the orthodox procedure where the difference in either the total energy or the density matrix at two successive cycles is used to measure the accuracy of the trial wave function. For example, a highly level-shifted procedure may converge very slowly, such that the changes in the wavefunctions are very small from one cycle to the next, even when the wave function is rather inaccurate. This criterion of convergence is sacrificed in the present method.

The practical utility of the computational efficiency gained in the present formalism at the expense of some loss in flexibility of the original method may not be so much felt in the *ab initio* SCF procedure where the saving in computation time achieved this way is only a meagre fraction of the total computation time of the calculation. However, this saving in time is sizeable in the semiempirical calculation of type CNDO, INDO etc., where most of the computation time is spent in the SCF part of the calculation.

Acknowledgements. This work was carried out when the author was a research associate at Cornell University. He expresses his thanks to Professor Roald Hoffmann for the use of facilities of his department. The constructive criticism of the manuscript by a referee, especially pointing out the shortcomings of the present formalism is gratefully acknowledged.

References

1. Saunders, V. R., Hillier, I. H.: Intern. J. Quantum Chem. **7**, 699 (1973); Hillier, I. H., Saunders, V. R.: Proc. Roy. Soc. **A320**, 161 (1970)
2. Roothaan, C. C. J.: Rev. Mod. Phys. **23**, 69 (1951); Roothaan, C. C. J.: Rev. Mod. Phys. **32**, 179 (1960)
3. Guest, M. F., Saunders, V. R.: Mol. Phys. **28**, 819 (1974)
4. Pople, J. A., Nesbet, R. K.: J. Chem. Phys. **22**, 571 (1954)
5. An excellent presentation of the Givens procedure is found in Ortega, J. M.'s article in "Mathematics for digital computers", Ralsten, A., Wilf, H. S. Eds. Vol. 2, page 94. New York: J. Wiley 1960
6. Hartree, D. R.: The calculation of atomic structures. New York: J. Wiley 1957

Received April 5, 1977/June 10, 1977