

Orthonormality-constrained orbital optimisations in MC-SCF theory

DEBASHIS MUKHERJEE

Department of Physical Chemistry, Indian Association for the Cultivation of Science, Calcutta 700 032

MS received 8 October 1977

Abstract. A simple algorithm for the direct solution of the MC-SCF orbital equations of McWeeny is proposed and its performance is compared with those of two gradient optimisation methods currently in use. This method has been found to possess good convergence property and is suited to treat a variety of multi-configuration wave functions. Some possible modes of useful applications of the method are suggested.

Keywords. Atoms and molecules; MC-SCF method.

1. Introduction

The multi-configuration SCF (MC-SCF) theory has received considerable interest in recent years in generating good-quality wave-functions from a relatively short CI-expansion. The MC-SCF wave-function is a fully optimised CI-type function of the form $\psi = \sum_I C_I \psi_I$, such that the energy $E = \langle \psi | H | \psi \rangle$ for normalised ψ is stationary with respect to variations in both the CI-coefficients C_I , and the set of orbitals ϕ_A appearing in ψ , subject to suitable constraints (Frenkel 1934; Slater 1953; Löwdin 1953). In a customary procedure adopted for molecules, the orbitals are expanded in terms of a fixed set of basis-functions (B-F), and the variation of orbitals in ψ is induced by variation of the linearly combining B-F (LCBF) coefficients. The energy E , in this case, is given by

$$E = \sum_{\substack{A, B \\ p, q}} T_{Ap}^\dagger \langle p | h | q \rangle T_{qB} P_{1BA} + \frac{1}{2} \sum_{\substack{A, B, C, D \\ p, q, r, s}} T_{Ap}^\dagger T_{qB} \langle pq | g | rs \rangle T_{Cr}^\dagger T_{sD} P_{2CD, AB} \quad (1)$$

where T is the matrix of the LCBF coefficients, and the capital and small suffixes stand for the orbital and basis-function indices respectively. The quantities $\langle p | h | q \rangle$ and $\langle pq | g | rs \rangle$ are the one and two electron integrals of the Hamiltonian in terms of the B-F's, and P_1 and P_2 are the one and two electron density matrices. Variation

of E with respect to the CI coefficients C_I lead to the familiar CI secular equations:

$$\sum_I \langle \psi_I | H | \psi_J \rangle C_I = EC_I \quad (2)$$

and the corresponding orbital equation coming from the variation of the LCBF's takes the form

$$\mathbf{B} = \mathbf{S}\mathbf{T}\epsilon \quad (3a)$$

$$\text{where } \mathbf{B} = \mathbf{H}\mathbf{T}\mathbf{P}_1 + \mathbf{Z} \quad (3b)$$

$$\text{subject to } \mathbf{T}^\dagger \mathbf{S}\mathbf{T} = \mathbf{1}. \quad (4)$$

In (3b), h stands for the matrix of the integrals $\langle p | h | q \rangle$, and \mathbf{Z} is defined as

$$Z_{pA} = \sum_{\substack{B, C, D \\ q, r, s}} T_{qB} \langle pq | g | rs \rangle T_{Cr}^\dagger T_{sD} P_{z_{CD, AB}} \quad (5)$$

ϵ is the Lagrange's multiplier matrix, ensuring the orthonormality of the orbitals appearing in ψ , and is hermitian.

The general MC-SCF orbital eq. (3) for an MC-SCF function was first given by McWeeny (1955, 1968), and since then rederived by others (Gilbert 1965; Das and Wahl 1967; Adams 1967). However, except for some special selection of configurations (see e.g. Das and Wahl 1966, Hinze and Roothaan 1967; Clementi and Veillard 1967; McWeeny 1975), a general solution of the orbital eq. (3) to determine the optimised orbitals has been considered awkward and has never been attempted. McWeeny, for example, advocates a direct minimisation of the energy expression (1) with respect to LCBF matrix T using gradient methods, rather than solving eq. (3) straightway (Mukherjee and McWeeny 1970; Claxton *et al* 1970; Kuprievich and Shramko 1972; Levy 1970). The purpose of the present paper is to suggest a simple method of solution of eq. (3) for a general MC-SCF function. The method will hereafter be called A. For an assessment of its performance, we shall compare the convergence properties with those of some gradient minimisation methods currently in use, which maintain orthonormality of orbitals during the variations. These latter methods will be briefly reviewed under the headings, methods B and C.

2. Method A

For a self-consistent solution of the eqs (2) and (3), subject to the constraints (4), we must omit from the energy expression (1) all the 'virtual' orbitals, i.e., orbitals that do not actually enter the CI expansion of ψ . If there are m BF's and n ($n \leq m$) orbitals in ψ , then T is an ($m \times n$) matrix of LCBF's, and ϵ is an ($n \times n$) hermitian matrix, ensuring orthonormality of n 'occupied' orbitals. The restriction as stated above is essential for the working of the method, otherwise eq. (3a) will be indeterminate.

We may write eq. (3a) as

$$S^{-1/2} B = S^{1/2} T \epsilon. \quad (6)$$

Left multiplying (6) with its hermitian conjugate, and using the hermiticity of ϵ , we have

$$B^\dagger S^{-1} B = \epsilon^2 \quad (7a)$$

$$\text{or} \quad (B^\dagger S^{-1} B)^{1/2} = \epsilon. \quad (7b)$$

Hence, using (7b), we may eliminate the Lagrange's multiplier matrix ϵ from (3a), and may write it as

$$B = ST (B^\dagger S^{-1} B)^{1/2} \quad (7c)$$

$$\text{or} \quad T = S^{-1} B (B^\dagger S^{-1} B)^{-1/2}. \quad (8)$$

Equation (8) indicates that, for a non-singular matrix ϵ , the orbital eq. (3a) may be cast in a form which allows an SCF-like iteration method for determining the optimised orbitals.

For the total minimisation of energy for a general MC-SCF function, we have tried two schemes (A1 and A2). In A1, for a fixed set of density matrices P_1 and P_2 , the matrix $S^{-1} B (B^\dagger S^{-1} B)^{-1/2}$ is calculated with the current LCBF matrix T , and is taken as the updated matrix of T . With this new T , a fresh CI calculation is done to get updated P_1 and P_2 , and the process is repeated till self-consistency. In algorithm A2, we repeatedly make use of the relation.

$$T^i \rightarrow T^{i+1} = S^{-1} B^i (B^{i\dagger} S^{-1} B^i)^{-1/2} \quad (9)$$

for a fixed set of P_1 and P_2 till self-consistency is reached, after which the updated P_1 and P_2 are found out from a fresh CI.

3. Method B

Here, we write the updated LCBF matrix in the form

$$T \rightarrow TU \quad (10)$$

where U is a unitary matrix, with

$$U = [1 + K] [(1 + K)^\dagger (1 + K)]^{-1/2}. \quad (11)$$

K is a square-matrix. For a full variation, it is *essential* here to include in T the LCBF coefficients of *all* the m orbitals that can be built up from m BF's*. Taking the $(m \times m)$ matrix K as our variables, the first order change in energy is given by

$$\delta E = \text{Tr} [(B^\dagger T) - (T^\dagger B)] K. \quad (12)$$

*This is unlike that in A. In A there is an implied inversion of the Lagrange's multiplier matrix ϵ , and the inclusion of virtual orbitals would have led to a null block in ϵ in the virtual space of the orbital set.

We now choose K , according to the steepest descent theory, along the negative of the gradient direction:

$$K = -\lambda [(T^\dagger B) - (B^\dagger T)]. \quad (13)$$

For the determination of the optimum step-length, we use an empirical quadratic interpolation formula analogous to that used by Mukherjee and McWeeny (1970). The method B is essentially similar to the method (2) of Mukherjee and McWeeny except for two minor variations: (i) we have chosen U as in (11), mainly to make use of the computer programme for the calculation of square root inverse of a matrix needed for all the three methods discussed and, (2) in the estimate of the second order term, we calculate energy with respect to an orthonormal set $T_{\text{est}} = TU_{\text{est}}$, using eq. (13) for K with a $\lambda = \lambda_{\text{est}}$. Thus for the estimation of the second order term in λ , we are using the actual energy surface. Mukherjee and McWeeny (1970) used orbitals orthonormal only to the first order for the same purpose, and it has been our experience that in unfavourable cases this leads to a negative value of the second order term.

4. Method C

In this method, we would like to vary E directly in terms of the orbital coefficients T , maintaining the orthonormality constraint. In the spirit of the techniques employed by Fletcher (1970), Kari and Sutcliffe (1970, 1973), we write T in terms of a new set of variables:

$$T = Y(Y^\dagger SY)^{-1/2}. \quad (14)$$

T , written this way, automatically satisfies the constraints (4). For an $(m \times m)$ matrix T , Y is also of the same dimension. Defining a matrix M to be

$$M = (Y^\dagger SY)^{-1/2}. \quad (15)$$

We have the first order variation of M as

$$\delta M = -\frac{1}{2} M^2 \delta Y^\dagger SY + Y^\dagger \delta Y M, \quad (16)$$

(See e.g. Mukherjee 1974; Kari *et al* 1973).

In this derivation, we have assumed that derivation of M from unit matrix is rather small, so that considerable simplification is achieved. Kari and Sutcliffe (1973) have discussed the rigorous derivation leading to a more complicated expression and have also indicated the mode of simplification leading to (16). The first order variation of energy as a function of Y then takes the form (Mukherjee 1974):

$$\delta E = \text{Tr} [\delta Y^\dagger V] \quad (17a)$$

with

$$V = 2(BM - \frac{1}{2}STT^\dagger BM - \frac{1}{2}STB^\dagger TM). \quad (17b)$$

Again, the change δY will be taken along the negative of the gradient direction, with a step-length estimated as in method B.

5. Applications to conjugated systems

The two variants of the method A, and the methods B and C have been used to calculate the ground and the first excited singlet and triplet states of three conjugated hydrocarbons: (1) dimethylene cyclobutene, (2) fulvene and (3) trans-trans-hexatriene in the π -electron approximation. The choice of the semi-empirical parameters are the same as in our previous communications (Mukherjee 1974, 1975). Under ZDO, the matrix S reduces to a unit matrix.

For the ground state calculation, the CI function includes the HF and all the closed-shell type doubly and quadruply excited configurations. The total number of determinants is nineteen. For the excited singlet and triplet states, only the two determinants obtained by exciting an electron from the highest occupied orbital in the H-F configuration to the lowest excited orbital are taken. The CI-coefficients for the excited states are determined only by the spin-coupling scheme and not by the orbitals chosen, so that the energy has to be minimised with respect to LCBF's only. Thus, for these states methods A1 and A2 are identical.

6. Results and discussion

The results of calculation for the compounds are summarised in table 1. For these and other compounds studied by us, method A are in general faster convergent than

Table 1. Ground and excited state π -electron energies in eV

Compounds States	Dimethylene cyclobutene	Fulvene	Trans-trans- hexatriene
1. Ground State	-156.7796	-158.1366	-149.0119
2. Excited singlet state	-150.8078	-154.9722	-143.4923
3. Excited triplet state	-153.7889	-155.7229	-146.3777

Table 2. Convergence of the different methods for dimethylene cyclobutene

No. of Iterations	Gain in energy over the first iteration in eV			
	Methods A		Method B	Method C
	A1	A2		
0	0.0000	0.0000	0.0000	0.0000
1	0.0327	0.0639	0.0070	0.0070
2	0.0589	0.0926	0.0154	0.0156
3	0.0746	0.1083	0.0256	0.0260
4	0.0843	0.1088	0.0379	0.0381
5	0.0912	Converged	0.0526	0.0528
No. of times the energy and the gradient is evaluated to reach convergence	11+	7+	16	14

⁺In methods A, only the energy has to be calculated. The starting orbitals are the HF orbitals for the ground state.

the gradient optimisation methods B and C. The trend is found to be more marked for the ground state. This is illustrated in table 2, where the total number of times the energy (and where necessary the gradient) has to be evaluated for reaching global convergence is explicitly shown for the ground state energy calculation of dimethylene cyclobutene. It has been found that method A2 is always superior to A1. It is to be expected, because A2 arrives at a local minimum with respect to the orbitals for fixed sets P_1 and P_2 , while A1 seeks for only an approximate minimum. In A2, usually two, and at most three cycles of the iteration step (9) are found to be necessary to reach the local minimum with respect to orbitals. The increased time incurred in step (9) is more than compensated for by much faster rate of convergence.

It is a currently held opinion, that, between the SCF and steepest descent methods, the SCF procedure is, in general, superior for closed and simple open-shell system—when it converges of course. Our results seem to indicate that an analogous situation may hold good for MC-SCF theories also.*

A useful feature of method A is the easy implementability and smooth performance for a variety of wave-functions chosen. The technique may even be used to find out optimised orbitals for closed and open-shell SCF theories, instead of Roothaan technique, and a detailed comparative study of their relative performance would be interesting. Furthermore, in methods A, the Lagrange's multiplier matrix ϵ need not be diagonal, so that determination of well-localised orbitals starting from a chosen set of localised bond-orbitals may be practicable. Preliminary studies on the determination of localised HF orbitals for the same compounds led to highly localised orbitals, and more detailed investigation is under way. The method seems well suited to optimise the orbitals for genuinely open-shell single configurations for which SCF-theory is cumbersome to develop, and thus will be helpful to study excited state geometry, excited state molecular potential and excited state basicity calculations.

Acknowledgements

The Author expresses his sincere thanks to Dr N G Mukherjee, for helpful correspondence and to Mr S Adnan and Dr S Bhattacharya for illuminating discussions.

References

- Adams W H 1967 *Phys. Rev.* **156** 109
- Adnan S, Bhattacharya S and Mukherjee D (to be published)
- Claxton T A and Weiner B 1970 *Trans. Faraday Soc.* **66** 2113
- Clementi E and Veillard A 1967 *Theoret. Chim. Acta* **7** 133
- Das G and Wahl A C 1966 *J. Chem. Phys.* **44** 87
- Das G and Wahl A C 1967 *J. Chem. Phys.* **47** 2934
- Fletcher R 1970 *Mol. Phys.* **19** 55
- Frenkel J 1934 in *Wave Mechanics—Advanced General Theory* (Oxford: Clarendon Press)
- Gilbert T L 1965 *J. Chem. Phys.* **43** 5248
- Hinze J and Roothaan C C J 1967 *Prog. Theoret. Phys.* **10** 37

*Recently we have finished a convergence study of ground and excited state MC-SCF wave-functions of some N-heterocyclics (Adnan *et al*, to be published) which strengthens our belief about the efficacy of the above procedure.

- Kari R and Sutcliffe B 1970 *Chem. Phys. Lett.* **7** 149
Kari R and Sutcliffe B 1973 *Int. J. Quantum Chem.* **7** 459
Kuprievich V A and Shramko O V 1972 *Int. J. Quantum Chem.* **6** 327
Levy B 1970 *Int. J. Quantum Chem.* **4** 297
Lowdin P O 1953 *Phys. Rev.* **97** 1474
McWeeny R 1955 *Proc. R. Soc. London* **A232** 114
McWeeny R 1968 *Faraday Soc. Symp.* **2** 7
McWeeny R 1975 *Chem. Phys. Lett.* **35** 13
Mukherjee N G and McWeeny R 1970 *Int. J. Quantum Chem.* **4** 97
Mukherjee D 1974 *Int. J. Quantum Chem.* **8** 247
Mukherjee D 1975 *Int. J. Quantum Chem.* **9** 943
Slater J C 1953 *Phys. Rev.* **91** 528