

Direct Minimization of the Energy Functional in the LCAO-MO Density Matrix Formalism

IV. All-Pair Excitation Wavefunctions in Multiconfigurational Self-Consistent Field

Piercarlo Fantucci

Istituto di Chimica Generale dell'Università, Centro C.N.R., Via Venezian, 21, I-20133, Milano, Italy

Stefano Polezzo

Centro C.N.R. "Relazioni fra struttura e reattività chimica", Via Golgi, 19, I-20133, Milano, Italy

A previously proposed method of energy minimization is developed for MC SCF wavefunctions formed by all-pair excitations for a closed-shell system. The orbital coefficients are optimized by a gradient approach using a suitable orthogonal transformation of the atomic basis, while optimum CI coefficients are determined solving the usual secular problem for the lowest eigenvalue, after each optimization of the orbitals. Applications to LiH and NH₃ molecules show that the method is numerically well stable, and is capable of accounting for a large part of the correlation energy giving results which compare well with those of the conventional CI method.

Key words: Direct minimization of energy – PEMCSCF formalism

1. Introduction

In this paper the previously proposed method of direct energy minimization by a double iteration procedure [1a, b, c] is developed for a MC SCF expansion consisting of a single closed-shell determinant together with all double excitation functions (single determinants) obtained by promoting a pair of electrons from occupied to unoccupied orbitals, the so-called Paired Excitation MC SCF method (PEMCSCF) of Ref. [2].

The PEMCSCF method has received a great deal of attention, being a restricted but effective way to obtain molecular correlation energies [3a, b, 4 and Refs. therein]. Essentially two approaches have been developed for optimum orbital determination, either by solving iteratively a series of coupled pseudo-eigenvalue problems (involving Lagrangian multipliers), or by directly minimizing the energy functional, e.g. by a steepest descent technique or two-by-two rotations [3]. In particular, recently a direct variational approach has been given in which the

orbitals are obtained from a single Fock-like pseudo-eigenvalue equation for all the orbitals [2].

In our formulation the LCAO coefficient optimization is achieved by a direct minimization of the energy acting through a suitable orthogonal transformation, while the optimization of the CI expansion coefficients is obtained by solving the usual secular equations for the lowest eigenvector after each orbital optimization. In Sect. 2 we give the derivation of the energy together with its gradient and stationary conditions; in Sect. 3 some connections with other works are pointed out; in Sect. 4 the minimization steps are described; in Sect. 5 numerical applications to LiH and NH₃ molecules are presented to test the performance of the procedure.

2. Energy Gradient and Stationary Conditions. Optimum Orbitals

Let us consider a system with a non-degenerate closed-shell ground state represented by a multi-configuration wavefunction of the form

$$\Psi = b_0 \Psi_0 + \sum_I \sum_V^{m_0, m_v} \Psi_{I,V} b_{VI}, \quad (1)$$

in which Ψ_0 is a basic determinant of doubly-occupied MO's and the $\Psi_{I,V}$ are similar single determinants obtained by exciting a pair of electrons from an occupied orbital I of the m_0 orbitals doubly-occupied in Ψ_0 ($\{I\}$ set) to a virtual orbital V of m_v orbitals unoccupied in Ψ_0 ($\{V\}$ set). The whole set $\{K\}$ of orbitals $\{I\}$ and $\{V\}$ is assumed orthonormal and the CI expansion coefficients b_0 and b_{VI} fulfill the orthonormality condition

$$b_0^2 + \text{tr}(\mathbf{b}\hat{\mathbf{b}}) = 1, \quad (2)$$

where \mathbf{b} is the $m_v \times m_0$ matrix of elements b_{VI} .

In the following I, I', \dots denote the orbitals of the $\{I\}$ set, V, V', \dots the orbitals of the $\{V\}$ set, and K, L, \dots the orbitals of both sets. All quantities are taken real, for convenience.

Going directly into the LCAO expansion of the MO $\{K\}$ in terms of an orthonormal atomic set of m basis functions $\{\chi\}$ and retaining the notation of Ref. [1c] we write

$$|I\rangle = |\chi\rangle \mathbf{C} \quad \text{and} \quad |V\rangle = |\chi\rangle \mathbf{D} \quad (\text{with } \langle \chi | \chi \rangle = \mathbf{1}_m). \quad (3)$$

Moreover, following Ref. [2, 5c] we define the $m \times m$ density matrix for the K th MO, $|K\rangle$,

$$\mathbf{R}^K = \mathbf{T}_K \hat{\mathbf{T}}_K, \quad (= |K\rangle \langle K|) \quad (4)$$

and note that \mathbf{R}^K is the AO representation of the projection operator onto the one-dimensional subspace spanned by the K th MO, \mathbf{T}_K being the corresponding single column of the matrix \mathbf{C} or \mathbf{D} , which collects the LCAO coefficients. The following constraints must be satisfied by the density matrices \mathbf{R}^K (idempotency, exclusiveness and identity resolution)

$$\mathbf{R}^K \mathbf{R}^L = \delta_{KL} \mathbf{R}^K \quad (\text{tr } \mathbf{R}^K = 1; \sum_K \mathbf{R}^K = \mathbf{1}_m; \text{ all } K, L). \quad (5)$$

Again the following matrices associated with a matrix \mathbf{R} are defined as usual [1c] in the AO basis r, s, t, u, \dots

$$[\mathbf{G}(\mathbf{R})]_{rs} = \sum_{t,u}^m R_{tu} (2\langle ru | st \rangle - \langle ru | ts \rangle)$$

$$[\mathbf{K}(\mathbf{R})]_{rs} = \sum_{t,u}^m R_{tu} \langle ru | ts \rangle. \quad (r, s = 1, \dots, m) \quad (6)$$

Then the matrix elements of the Hamiltonian \mathcal{H} in terms of the matrices \mathbf{R}^k (AO basis) turn out to have the following expressions

$$\langle \Psi_0 | \mathcal{H} | \Psi_0 \rangle \equiv E_0 = \text{tr}(\mathbf{R}\mathbf{f}) + \text{tr}(\mathbf{R}\mathbf{h})$$

$$\langle \Psi_0 | \mathcal{H} | \Psi_{I,V} \rangle = \text{tr}[\mathbf{R}^I \mathbf{K}(\mathbf{R}^V)]$$

$$\langle \Psi_{I,V} | \mathcal{H} | \Psi_{I',V'} \rangle = \delta_{II'} \delta_{VV'} E_0 + \delta_{II'} \text{tr}[\mathbf{R}^V \mathbf{K}(\mathbf{R}^{V'})]$$

$$+ \delta_{VV'} \text{tr}[\mathbf{R}^I \mathbf{K}(\mathbf{R}^{I'})] - 2\delta_{II'} \delta_{VV'} \text{tr}[\mathbf{R}^V \mathbf{G}(\mathbf{R}^{I'})]$$

$$+ 2\delta_{II'} \delta_{VV'} \text{tr}[(\mathbf{R}^V - \mathbf{R}^{I'})\mathbf{h}], \quad (7)$$

where

$$\mathbf{h} \equiv \mathbf{h}(\mathbf{R}) = \mathbf{f} + \mathbf{G}(\mathbf{R}) \quad \text{and} \quad \mathbf{R} = \sum_I^{m_0} \mathbf{R}^I.$$

Introducing the two symmetric matrices

$$\mathbf{A} = \tilde{\mathbf{b}}\mathbf{b} \quad \text{and} \quad \mathbf{B} = \mathbf{b}\tilde{\mathbf{b}}, \quad (8)$$

the energy expectation value E corresponding to the wavefunction (1) is found to be

$$E = E_0 + 2 \sum_I^{m_0} \sum_V^{m_0} \{ b_0 b_{VI} \text{tr}[\mathbf{R}^V \mathbf{K}(\mathbf{R}^I)] - b_{VI}^2 \text{tr}[\mathbf{R}^V \mathbf{G}(\mathbf{R}^I)] \}$$

$$+ \sum_I^{m_0} \sum_{I'=I}^{m_0} (2 - \delta_{II'}) A_{II'} \text{tr}[\mathbf{R}^I \mathbf{K}(\mathbf{R}^{I'})]$$

$$+ \sum_V^{m_0} \sum_{V'=V}^{m_0} (2 - \delta_{VV'}) B_{VV'} \text{tr}[\mathbf{R}^V \mathbf{K}(\mathbf{R}^{V'})]$$

$$+ 2 \sum_V^{m_0} B_{VV} \text{tr}(\mathbf{R}^V \mathbf{h}) - 2 \sum_I^{m_0} A_{II} \text{tr}(\mathbf{R}^I \mathbf{h}). \quad (9)$$

In the MC SCF theory the two sets of coefficients, CI expansion coefficients $\{b_0, b_{VI}\}$ and LCAO MO coefficients $\{T_{rK}\}$, are varied in turn. The variation in E with respect to the CI coefficients is simply a linear variation problem involving the computation of the lowest eigenvector of the CI matrix of elements (7) and will not be discussed further. Instead, the non-linear problem of variation with respect to the orbital coefficients T_K , or the density matrices \mathbf{R}^k , will be considered now.

If an orthogonal transformation \mathbf{U} is performed on the full atomic basis $|\chi\rangle$, namely if $|\chi\rangle \rightarrow |\chi\rangle \mathbf{U}$, the density matrices \mathbf{R}^k are transformed into

$$\bar{\mathbf{R}}^k = \mathbf{U} \mathbf{R}^k \tilde{\mathbf{U}}, \quad (10)$$

and the energy \bar{E} is again given by the expression (9) but with \bar{R}^K in place of R^K . Of course the constraints (5) hold true also for the transformed matrices \bar{R}^K .

As in earlier works [1], the orthogonal matrix we take is

$$U(X) = -\mathbf{1}_m + 2P^{-1}, \quad \text{with } P = \mathbf{1}_m + X - \hat{X},$$

where X is an arbitrary $m \times m$ matrix and the elements of $X - \hat{X}$ may be regarded as independent unconstrained variables. Then, letting $U \rightarrow U + \delta U$, the first-order change $\delta \bar{E}$ of the energy \bar{E} with respect to X is easily worked out (the method follows closely one developed elsewhere [1]) and after some straightforward algebra the gradient matrix G_x for variation in X is found to be

$$G_x = -8\bar{P}^{-1}(E_x - \tilde{E}_x)P^{-1}, \quad (11)$$

where the matrix E_x is defined by

$$\begin{aligned} E_x = & \mathbf{h}\bar{R} + \sum_I \sum_V \{b_0 b_{VI} [K(\bar{R}^I)\bar{R}^V + K(\bar{R}^V)\bar{R}^I] - b_{VI}^2 [G(\bar{R}^I)\bar{R}^V + G(\bar{R}^V)\bar{R}^I]\} \\ & + \sum_I \sum_{I'=I} (2 - \delta_{II'}) A_{II'} K(\bar{R}^{I'}) \bar{R}^I \\ & + \sum_V \sum_{V'=V} (2 - \delta_{VV'}) B_{VV'} K(\bar{R}^{V'}) \bar{R}^V \\ & + \sum_V B_{VV} [\mathbf{h}\bar{R}^V + G(\bar{R}^V)\bar{R}] - \sum_I A_{II} [\mathbf{h}\bar{R}^I + G(\bar{R}^I)\bar{R}]. \end{aligned}$$

The condition for a stationary energy, $G_x = \mathbf{0}$, is then equivalent to the condition that E_x has to be a symmetric matrix, i.e.

$$E_x = \tilde{E}_x. \quad (12)$$

These are $m(m-1)/2$ relations just necessary to determine the elements of the arbitrary skew-symmetric matrix $X - \hat{X}$. Eventually the density matrices \bar{R}^K and the orbitals $\bar{T}_K = U T_K$ are obtained, which therefore yield a stationary value of the energy functional (9) and are then the optimum quantities.

Of course the expectation value of any physical quantity can be expressed in terms of the matrices \bar{R}^K : thus, for example, the electronic part of the dipole moment $\mathcal{M} = \sum_i m_i$ is

$$\langle \Psi | \mathcal{M} | \Psi \rangle = 2 \operatorname{tr} (m \bar{P}_1)$$

where

$$\bar{P}_1 = \bar{R} + \sum_V B_{VV} \bar{R}^V - \sum_I A_{II} \bar{R}^I$$

is the one-body density matrix (a factor 2 apart) and m is the matrix of elements $\langle r | m_i | s \rangle$.

It is worth noting that the energy can be obtained as a side result in the gradient computation employing the relation

$$\bar{E} = \text{tr } E_x + \text{tr } (f\bar{P}_1).$$

3. Relation to Other Work

In this section the short line over the matrices R^k is omitted for simplicity. Introducing the effective Hamiltonians [2, 3a] h^I for the $\{I\}$ set orbitals and h^V for the $\{V\}$ set orbitals

$$h^I = (1 - A_{II})f + \sum_{I'} (1 - A_{II} - A_{I'I'})G(R^{I'}) + \sum_{I'} A_{II'}K(R^{I'}) + \sum_V [(B_{VV} - b_{VI}^2)G(R^V) + b_0 b_{VI}K(R^V)], \quad (13)$$

$$h^V = B_{VV} \left[f + \sum_I G(R^I) \right] + \sum_I [b_0 b_{VI}K(R^I) - b_{VI}^2 G(R^I)] + \sum_{V'} B_{VV'}K(R^{V'}),$$

it is a straightforward matter to prove the identity

$$E_x = \sum_K^m h^K R^K \quad (K \text{ runs over all orbitals}), \quad (14)$$

and the stationary energy condition (12) can thus be cast into the form

$$G_{x0} \equiv \sum_K h^K R^K - \sum_K R^K h^K = 0. \quad (15)$$

Now, multiplying (15) by, say, R^K on one side and R^L on the other the following relations are obtained

$$-R^K G_{x0} R^L \equiv R^K (h^K - h^L) R^L = 0 \quad (K, L = 1, \dots, m; K \neq L), \quad (15')$$

which are the necessary and sufficient conditions for the existence of the optimum MO coefficients just in the form given in Ref. [2] (see also [5, 1a]). The conditions (15') can be brought together to define a single effective (Fock-like) Hamiltonian \bar{h}^F [2, 5], whose commutation conditions with each R^K , $[\bar{h}^F, R^K] = 0$, are equivalent to Eq. (15')

$$\bar{h}^F = \sum_K^m R^K d_K R^K - \frac{1}{2} \sum_{K,L}^m c_{KL} R^K G_{x0} R^L + \frac{1}{2} \sum_{K,L}^m c_{KL} R^L G_{x0} R^K, \quad (16)$$

where d_K is an arbitrary symmetric matrix and c_{KL} are arbitrary nonzero numerical coefficients ($c_{KL} \neq c_{LK}$). General comments on \bar{h}^F and its connection with the level-shifting technique [4] are reported in Ref. [5c].

As shown in Ref. [6a, b], the relations (15') express the Hermitian conditions which the Lagrangian multiplier matrix must satisfy. They are the AO representation of the so-called generalized Brillouin theorem for optimum orbitals of Refs. [6a, b] (see also [1b]).

The matrix E_x defined by the identity (14) is the correspondent in the AO representation of the single-form non-Hermitian operator introduced in Ref. [3a], being indeed

$$E_x|I\rangle \equiv h^I|I\rangle \quad \text{and} \quad E_x|V\rangle \equiv h^V|V\rangle.$$

Multiplying the stationary conditions (15) by, say, R^L on the right we obtain, for the PEMCSCF theory, the equation

$$(h^L - \sum_K R^K h^K) R^L = 0 \quad (L=1, \dots, m),$$

which is the correspondent in the AO representation of a particular case (for $\lambda_{ji}=1$, to be exact) of the correct variational equation for optimum manifolds R^L (or orbitals $|L\rangle$, if we had multiplied by $|L\rangle$ instead of R^L) introduced and discussed at length in Refs. [6a, b] in the context of the general SCF theory: this equation is the starting point of the coupling operator method (see also [6c]).

4. Computational Procedure for Energy Minimization

In order to reach a stationary point on the energy surface satisfying the (minimum) conditions (12) for orbitals along with optimum CI expansion coefficients, an iterative process can be set up according to the following steps [1c].

- 1) The LCAO coefficient matrix to start with may be (but not necessarily) ground state SCF solution orbitals: the matrices C , D and $R^K = T_K \hat{T}_K$ are obtained in this way.
- 2) The eigenvalue problem corresponding to the CI matrix of elements (7) is solved for the lowest eigenvector, obtaining the initial CI coefficients $\{b_0, b_{VI}\}$.
- 3) The CI coefficients so found are held fixed and the minimization process for the orbital coefficients is started with $X=0$, the gradient G_x is computed and a new point X' is found along the chosen algorithm. With the new matrix $U' = U(X')$, corresponding to X' , the transformed matrices $\bar{R}'^K = U' R^K \bar{U}'$ are calculated and then employed for the new iteration. In the calculations reported below the variable metric algorithm of Murtagh and Sargent has been applied as described in Ref. [1a].
- 4) When some degree of optimization is reached, the matrices \bar{R}^K (or \bar{C} and \bar{D}) so found are used in step 2 to redetermine the new coefficients $\{b_0, b_{VI}\}$; then step 3 follows.

This sequence forms the third part of a computer program. The first part is the Hehre-Pople Gaussian 70 program [7] for the SCF quantities, the second part contains the integral transformation into the basis of Schmidt orthonormal and symmetry adapted orbitals.

Spatial symmetry is readily taken into account due to the block structure of the matrices R^K , $G(R^K)$ and $K(R^K)$.

5. Test Calculations

In this section we present our results for SCF, CI (pair excitations) and MC SCF calculations relative to the energies and the dipole moments of LiH and NH₃ molecules (see Table 1). The comparison with some results of other authors is also reported. All calculations have been done with STO-6G basis expansion [7].

Table 1. MCSCF results for LiH and NH₃ molecules

<i>Total energies (hartree)</i>								
	E_{SCF}	E_{MC}	$E_{\text{SCF}} - E_{\text{CI}}$			$E_{\text{SCF}} - E_{\text{MC}}$		
			<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>c</i>
LiH	-7.9854	-8.0318	0.0367	0.024	—	0.0464	0.031	0.0442
NH ₃	-56.1483	-56.2030	0.0393	0.031	0.0253	0.0547	0.034	0.0297

<i>Dipole Moments (debye)</i>							
	SCF		CI		MCSCF		Exp.
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	
LiH	6.005	5.63	5.956	5.56	5.906	5.11	5.88
NH ₃	2.275	2.06	2.263	2.04	2.239	2.03	1.47

^a present work. ^b Ref. [3]. ^c Ref. [2].

For LiH molecule we assume a bond distance of 3.015 bohr and the same basis set of the Ref. [2]. The ground state SCF energy is near to the Hartree-Fock limit (-7.987313 hartree [8]) and the final MC SCF energy accounts for the 54 per cent of the correlation energy (the exact value being -8.0703 hartree [9]) while our CI calculation accounts for the 42 per cent. Moreover, the MC SCF calculation changes the dipole moment in the proper way.

For the NH₃ molecule the experimental geometry has been assumed ($R_{\text{N-H}} = 1.91$ bohr, $\text{H}\hat{\text{N}}\text{H} = 106.7^\circ$). For the nitrogen atom we use the double- ζ basis of Clementi [10] while for the hydrogen atom the basis set includes the $1s$ and $2s$ orbitals already used in the LiH calculation. The SCF energy is far from the Hartree-Fock limit (-56.225 hartree [11]) and the MC calculation is not able to reach this limit. Thus, the gain in energy may not be considered of correlation nature. The dipole moment, even if far from the experimental value, is corrected by the MC calculation in a right sense.

A positive feature of the method, already noted in our previous papers, is its good numerical stability and its convergence guarantee. In order to give a rough idea of the convergence rate of the method we refer to the LiH calculation. Assuming as convergence criterion a gradient modulus smaller than 0.00015, the energy minimum has been reached after seven cycles, each including the determination of the secular determinant lowest root followed by about thirty iterations for orbital optimization. The values of the parameter α [1] were chosen in the range

0.040–0.050 and no attempt was made of optimizing these values to improve the convergence rate. Each cycle required about 2.5 minutes (CPU time only) on a UNIVAC 1106 computer, but we point out that the present version of the computing program may be largely improved.

References

1. a) Polezzo, S.: *Theoret. Chim. Acta (Berl.)* **38**, 211 (1975)
b) Polezzo, S.: *Theoret. Chim. Acta (Berl.)* **40**, 245 (1975)
c) Fantucci, P., Polezzo, S., Stabilini, M. P.: *Theoret. Chim. Acta (Berl.)* **41**, 311 (1976)
2. Cook, D. B.: *Mol. Phys.* **30**, 733 (1975)
3. a) Golebiewski, A., Nowak-Broclawik, E.: *Mol. Phys.* **26**, 989 (1973)
b) Golebiewski, A., Nowak-Broclawik, E.: *Mol. Phys.* **28**, 1283 (1974)
4. Wood, M. H., Veillard, A.: *Mol. Phys.* **26**, 595 (1973)
5. a) McWeeny, R., Sutcliffe, B. T.: *Methods of molecular quantum mechanics*. London: Academic Press 1969
b) McWeeny, R.: *Mol. Phys.* **28**, 1273 (1974)
c) McWeeny, R.: *Chem. Phys. Letters* **35**, 13 (1975)
6. a) Hirao, K., Nakatsuji, H.: *J. Chem. Phys.* **59**, 1457 (1973)
b) Hirao, K.: *J. Chem. Phys.* **60**, 3215 (1974)
c) Morikawa, T.: *Chem. Phys. Letters* **32**, 521 (1975)
7. QCPE Program N. 236 (Indiana University)
8. Cade, P. E., Huo, W. M.: *J. Chem. Phys.* **47**, 614 (1967)
9. Mukherjee, M. G., McWeeny, R.: *Intern. J. Quantum Chem.* **4**, 97 (1970)
10. Clementi, E.: *J. Chem. Phys.* **40**, 1944 (1964)
11. Rank, A., Allen, L. C., Clementi, E.: *J. Chem. Phys.* **52**, 4133 (1970)

Received January 17, 1977/April 4, 1977