Theoret. chim. Acta (Berl.) 4, 30—36 (1966)

Università di Padova, Istituto di Chimica Fisica, Italia

Two By Two Rotations of Orbitals for Minimizing the Energy of LCAO-MO Wave Functions

By

Massimo Rossi

A new method is studied for finding the molecular orbitals which minimize the energy of an LCAO-MO wavefunction. The method makes use of successive rotations of pairs of orbitals. It can be applied to multi-determinant as well as to single-determinant wavefunctions. Criteria are found to minimize excited states.

Une nouvelle méthode est étudiée ici pour trouver les orbitales moléculaires qui minimisent l'énergie d'une fonction d'onde LCAO-MO. La méthode utilise des rotations successives de couples d'orbitales. La méthode peut minimiser aussi des fonctions d'onde multi-déterminantales et les états excités.

Es wird eine neue Methode zur Bestimmung der M.O.s nach dem Kriterium minimaler Gesamtenergie beschrieben. Dabei werden jeweils zwei M.O.s sukzessive transformiert. Das Verfahren bleibt auch dann anwendbar, wenn für die Zustandsfunktion eine Linearkombination von HS-Determinanten angesetzt wird. Zum Schluß werden noch einige Kriterien für die Bestimmung von angeregten Zuständen angegeben.

Introduction

The problem of minimizing the energy of an LCAO-MO single determinant wavefunction by variation of the coefficients was solved in 1951 [*7*] for the first time for a closed shell ground state. The coefficients can be found solving by numerical methods the Roothaan equations, which are an application to LCAO-MO wave functions of the Hartree-Fock equations [*4*]. The so-called "unrestricted" open shell wave functions [*1*, *5*] can be treated in a similar way, considering separately electrons with opposed spin.

In 1960 [*6*] the way of extending the method to singly and doubly filled open shell wave functions was found. The possibility of operating on density matrices instead of on molecular orbitals was studied also [*3*].

In Roothaan equations the m. o. are eigenfunctions of an operator F depending itself on the m. o. From a starting set of m. o. Π^0 a sequence $\Pi^1, \Pi^2, \ldots$ can be obtained where the Π^{i+1} are eigenvectors of an F built with Π^i. If the sequence converges to a limit Π^* this set is the solution of the Roothaan equations. Energy not always is lowered passing from Π^i to Π^{i+1}: convergence generally can be reached only if Π^0 has been choosen near enough to the unknown Π^*.

The most important feature of the new method we shall expose here is that it minimizes directly the energy, without using the condition $dE = 0$ (that is without passing through Roothaan equations), so that convergence is always reached, independently from the choice of Π^0.

The Method

I. Let us consider a set of n atomic orbitals* $\chi_1, \chi_2, \ldots \chi_n$ and a set Π^0 of n orthonormal molecular orbitals where $\pi_i^0 = \sum_i c_{ij}^0 \chi_j$. In m. o. theory the wave function is represented by an antisymmetrized product of m. o. (with spin part), that we shall write briefly:

$$D = \left| \pi_1^0 \bar{\pi}_2^0 \ldots \pi_m^0 \right|$$

(where the vertical bars indicate antisymmetrisation and normalisation, the bars over the π_i^0 indicate spin α and no bar indicate spin β and position p in the product indicates electron p).

The wavefunction may also be a linear combination of two or more Slater determinants (triplet or singlet excited states, projected "unrestricted" m. o. wavefunctions and so on).

In all cases the energy depends on coefficients c_{ij}^0 which are the elements of a matrix C^0. This matrix can be now modified by a rotation of rows 'a' and 'b', corresponding to the rotation of the orbitals π_a^0 and π_b^0:

$$\begin{aligned} \pi_a^1 &= \cos\varphi\ \pi_a^0 + \sin\varphi\ \pi_b^0 \\ \pi_b^1 &= -\sin\varphi\ \pi_a^0 + \cos\varphi\ \pi_b^0 \,. \end{aligned}$$

Using this new matrix C^1 the energy is a function of φ and by numerical or, when possible, by analitycal methods we can find the point corresponding to the lowest energy, $\varphi_{\min}$. In the new set of orbitals defined by $\varphi_{\min}$ a new pair of orbitals may be choosen and rotated to reach the bottom of the energy well and so on, for all possible pairs. The whole cycle will be repeated many times as long as convergence is obtained, that is when further rotation of orbitals cannot lower the energy again**.

Rotating orbitals has an useful feature: replacing some orbitals in a Slater determinant with a linear combination of other orbitals, a row of the determinant becomes a combination of rows and the determinant a combination of determinants.

For instance

$$D = \left| (u\pi_1 + v\pi_2)\,\pi_3 \ldots \pi_m \right| = u \left| \pi_1 \pi_3 \ldots \pi_m \right| + v \left| \pi_2 \pi_3 \ldots \pi_m \right| \qquad (1)$$

or (remembering that a determinant with two equal rows is zero and that the exchange of rows is equivalent to a change of sign)

$$D = \left| (u\pi_1 + v\pi_2)\,(u\bar{\pi}_1 + v\bar{\pi}_2)\,(-v\pi_1 + u\pi_2)\,(-v\bar{\pi}_1 + u\bar{\pi}_2) \right| = \left| \pi_1 \bar{\pi}_1 \pi_2 \bar{\pi}_2 \right| . \qquad (2)$$

In this way, in most cases, we may expand the wave function after two or more rotations as a linear combination of wave functions in the starting orbitals. This fact will be useful in dealing with excited states, but also to understand some important features of the ground state and give help in calculations.

---

* We shall speak always of atomic orbitals and m. o. but all we shall say is also valid, of course, for each set of linearly independent analytical functions χ_i and for each set of orthonormal functions π_j obtained by a linear combination of the χ_i.

** If the χ_i are orthonormal (if they are not so, they can always be orthonormalised), instead of rotating rows of C^0 it is possible to rotate columns without breaking orthonormality of the π_i. Of course the rotation of columns does not correspond to rotation of orbitals and this is the advantage which induces us to prefere the first way of minimizing the energy.

II. It can be shown that the set of final molecular orbitals Π^* found with the method of rotations, satisfies also the condition $dE = 0$. Indeed, let us express the energy (instead of as a function of the coefficients) as a function of the parameters $\varphi_1, \varphi_2, \ldots, \varphi_{\binom{n}{2}}$, defining the successive rotations of each of the $\binom{n}{2}$ pairs of orbitals. If this successive rotations are defined starting from the set Π^*, we have that in the point $\varphi_1 = \varphi_2 = \ldots = \varphi_{\binom{n}{2}} = 0$ the partial derivatives $\frac{\partial E\,(\varphi_i)}{\partial \varphi_i} = 0$ $(i = 1, 2, \ldots, \binom{n}{2})$.

III. Ground states. A closed shell ground state is represented by a single determinant, where m molecular orbitals are doubly filled with electrons with opposed spin: $\Phi_G = |\,\pi_1\,\overline{\pi}_1\,\pi_2\,\overline{\pi}_2 \ldots \pi_m\,\overline{\pi}_m\,|$.

A rotation of two doubly filled orbitals does not change the determinant (see the example 2 in section *II*), so that the only rotations to do are those of a filled π_a and an empty π_b m. o.

Expanding in the old orbitals we have:

$$\Phi_G\,(\varphi) = \cos^2 \varphi\, \Phi_G\,(0) + \sqrt{2} \sin \varphi \cos \varphi\, \Phi_E\,(0) + \sin^2 \varphi\, \Phi_D\,(0)$$

where $\Phi_D\,(0)$ is the diexcited configuration with two electrons lifted from π_a to π_b and $\Phi_E\,(0)$ is the corresponding monoexcited singlet configuration.

As we see the rotation of orbitals is not equivalent to configuration interaction of Φ_G and Φ_E (as it would happen with π_a singly filled) but an indirect relation exists, because the C. I. matrix element between Φ_G and Φ_E is zero in $\varphi_{\min}$ (see Appendix I). With the final set Π^* all matrix elements with excited states are zero. This is not surprising because, from what was said in section *II* follows that the Brillouin theorem (7) must be also valid here (in fact the m. o. we find are also solution of the Hartree-Fock equations).

In Appendix I some useful relations among the functions $G\,(\varphi)$, $D\,(\varphi)$, $E\,(\varphi)$, $GD\,(\varphi)$, $DE\,(\varphi)$ are listed (see App. I for the meaning of these notations). These relations allow alternative ways for finding the minimum point $\varphi_{\min}$.

When the energy depends on products of order four of sin φ and cos φ like in closed shell ground states (see App. I) a numerical method is necessary to find $\varphi_{\min}$. On the contrary an analytical solution is possible when this dependence is of order two (as happens rotating a singly filled with an empty m. o.).

IV. Symmetry. The existence of symmetry lowers the number of rotations to do. For instance the fact that, in closed shell and in "unrestricted" m. o. ground states, the matrix elements between the ground state and the monoexcited state are zero if π_a and π_b have different symmetry means that a rotation cannot lower the energy and $\varphi_{\min} = 0$.

V. An example. Calculations on an IBM electronic computer were done with our method. In the table an example (benzyl radical) is given. An "unrestricted" m. o. wave function (1, 5) was used for π electrons of the radical. The functions χ_i are here the seven p_z atomic orbitals on carbon atoms. Linearly combining them, two sets of orthonormal m. o. Π^α (for electrons with spin α) and Π^β (for electrons with spin β) can be obtained. The seven molecular orbitals obtained by the Hückel method can be used as starting set $\Pi^{\alpha 0}$ and $\Pi^{\beta 0}$. Hückel m.o.s are symmetric (S) or antisymmetric (A) with respect to reflection in the symmetry plane of the molecule. If the m.o. are filled as in the figure, the only rotations to do are:

$1-5$; $1-7$; $3-5$; $3-7$; $2-6$; $4-6$, in the set Π^α and $1'-5'$; $1'-7'$; $2'-4'$; $2'-6'$; $3'-5'$; $3'-7'$, in the set Π^β (see Fig. 1).

Satisfying convergence in spin density and energy is reached after few cycles, as the table shows. In the last column results are reported which have been obtained solving the Hartree-Fock equations (using of course the same values of the integrals over atomic orbitals; these values and the results in the last column where kindly supplied by Dr. Berthier). No significant difference between the two methods is apparent.

VI. We pass now to discuss the limits of validity of the method. Indeed, not always the minimum we find is the absolute minimum of the energy. It may happen (as in Roothaan's method), choosing a different starting point, that a lower value of the energy is reached. But often the way exists of doing a partial control and a correction on the final set Π^*. For instance, in singly filled m.o. wave functions, after having minimized by two by two rotations, it may be that, combining two excited states $\Pi_{a\rightarrow b}$ and $\Pi_{a\rightarrow c}$, we obtain a state $\Phi_{a\rightarrow b'}$ (where b' is a linear combination of b and c) with a lower energy than the ground state. Then the true ground state is $\Phi_{a\rightarrow b'}$, that is b', instead of a, must be filled[*].

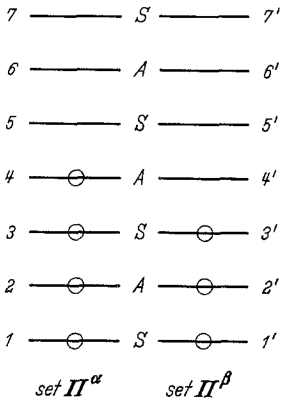

Fig. 1

With this new starting set the entire calculation must be repeated. If 3 is the dimension of our problem, with this correction we may be sure to have found the absolute minimum; in fact, a 3 by 3 rotation among a, b, and c of the set Π^* is equivalent to C. I. plus an exchange of orbitals (see App. II).

Table

The various steps in the convergence process for an unrestricted doublet wave function (benzyl molecule). Comparison with Roothaan method.

Spin density		initial value	1. cycle	2. cycle	3. cycle	4. cycle	5. cycle	Spin density by solving Hartree-Fock equations
on atom	1	+0.570	+0.721	+0.740	+0.741	+0.740	+0.739	+0.742
,,	2	+0.000	−0.118	−0.134	−0.139	−0.141	−0.142	−0.143
,,	3	+0.142	+0.188	+0.204	+0.210	+0.212	+0.213	+0.215
,,	4	+0.000	−0.079	−0.103	−0.110	−0.112	−0.113	−0.115
,,	5	+0.142	+0.173	+0.187	+0.194	+0.197	+0.199	+0.201
Lowering of the energy (eV)		–	−.17511	−.18292	−.18363	−.18373	−.18376	–

In an n dimensional problem ($n > 3$) only n by n rotations would give surely the absolute minimum. They can be done, with the aid of C. I. (analogously as for $n = 3$) only in two cases: when the filled orbitals are 1 or $n-1$. In these

---

[*] If a and b are singly filled and c is empty, and combining $\Phi_{a\rightarrow c}$ with $\Phi_{b\rightarrow c}$ we obtain a state $\Phi_{a'\rightarrow c}$ with lower energy than the ground state, the new ground state is that obtained exchanging a with c and b with a linear combination of a and b (see App. II).

cases the result is not dependent on the starting point. In other cases it is better to start (as in Roothaan method) with a set as near as possible to the presumed true minimum.

VII. Excited states. After having found the set Π^* which minimize the ground state Φ_G we can build with the same Π^* an excited configuration Φ_E, that is a wave function orthogonal to Φ_G.

Now, with the method of rotations, we can minimize the energy of Φ_E by a variation of Π^* without destroying orthogonality to the ground state. In fact, developing the wave function of an excited state Φ_E after two or more successive rotations on the set Π^* as a linear combination of wave functions in Π^* it is possible to see if the ground state Φ_G (Π^*) is mixed or not, that is if Φ_E is orthogonal to Φ_G (Π^*) as is required for an excited state. Generally some types of pairs of orbitals exist, whose rotation do not produce a mixing of the ground state, and so we can minimize Φ_E using only these types of rotations.

The mixing of the ground state depends also on the order of the rotations. If, for instance, rotations (ab), (cd) do not mix the ground state, it may be that they mix it in the inverse order (cd), $ab)$.

In the same way excited states of higher energy may be minimized keeping them orthogonal both to the ground state and to excited states of lower energy.

VIII. Other types of wave functions. The only condition to apply the method of rotations is that the wave function is a function of an unitary matrix. It can be applied of course also to minimize the energy of a wave function which is a linear combination of two or more determinants: this type of wave function is necessary, for instance, if we want to represent a state with spin multiplicity higher than two.

Appendix I

We show that:

a) $dG(\varphi)/d\varphi = -2\sqrt{2}\, GE(\varphi)$
b) $dD(\varphi)/d\varphi = 2\sqrt{2}\, DE(\varphi)$
c) $2\, GD(\varphi) + G(\varphi) + D(\varphi) = \text{const}$
d) $E(\varphi) + G(\varphi) + D(\varphi) = \text{const}$
e) $E(\varphi) - 2\, GD(\varphi) = \text{const}$

where:

$$G = \langle \Phi_G \,|\, H \,|\, \Phi_G \rangle; \qquad EG = \langle \Phi_E \,|\, H \,|\, \Phi_G \rangle; \qquad \text{and so on.}$$

In fact, developing $\Phi_D(\varphi)$ and $\Phi_E(\varphi)$ as $\Phi_G(\varphi)$ in Section III we get:

$$\Phi_D(\varphi) = u^2 \Phi_D - \sqrt{2}\, uv\, \Phi_E + v^2 \Phi_G; \qquad \Phi_E(\varphi) = (u^2 - v^2)\, \Phi_{E'} + \sqrt{2}\, uv\, (\Phi_D - \Phi_G)$$

where

$$u = \cos\varphi; \qquad v = \sin\varphi\,.$$

So we find that:

$$\begin{aligned}
G(\varphi) &= u^4 G + 2\sqrt{2}\, u^3 v\, EG + 2\, u^2 v^2 (E + GD) + 2\sqrt{2}\, uv^3\, ED + v^4 D \\
D(\varphi) &= u^4 D - 2\sqrt{2}\, u^3 v\, ED + 2\, u^2 v^2 (E + GD) - 2\sqrt{2}\, uv^3\, EG + u^4 G \\
E(\varphi) &= (u^4 + v^4)\, E + 2\sqrt{2}\, (u^3 v - uv^3)\, (ED - EG) + 2\, u^2 v^2 (D + G - E - 2\, GD) \\
EG(\varphi) &= u^4\, GE + \sqrt{2}\, u^3 v\, (E + GD - G) + 3\, u^3 v^2 (ED - GE) + \sqrt{2}\, uv^3 (D - E - GD) - v^4\, ED \\
ED(\varphi) &= u^4\, ED + \sqrt{2}\, u^3 v\, (D - E - GD) + 3\, u^2 v^2 (GE - ED) + \sqrt{2}\, uv^3 (E + GD - G) - v^4\, GE \\
GD(\varphi) &= (u^4 + v^4)\, GD + \sqrt{2}\, (u^3 v - uv^3)\, (ED - GE) + u^2 v^2 (G + D - 2\, E)\ ;
\end{aligned}$$

by means of these six expressions we can obtain easily the a) – e).

Appendix II

In part 1. we shall show that a 3 by 3 rotation of two empty orbitals π_2, π_3 and one filled orbital π_1 in a set Π^* which minimizes the energy of the ground state can lower the energy only if a combination exists of the two excited states $\Phi_{1\to 2}$ and $\Phi_{1\to 3}$ with a lower energy than Φ_G and that the new state of lower energy is obtained placing the electron which was in π_1 into a new orbital π_1', linear combination of π_2 and π_3. In part 2 the case of two empty and one filled orbital is examined.

1. 3 by 3 rotation of one filled orbital π_1 and two empty orbitals π_2 and π_3:

$$\pi_i' = \sum_{j=1}^{3} c_{ij}\,\pi_j \; (i = 1, 2, 3)$$

so that

$$\Phi_G' = c_{11}\,\Phi_G + c_{12}\,\Phi_{1\to 2} + c_{13}\,\Phi_{1\to 3}$$

and

$$\langle \Phi_G' \mid H \mid \Phi_G' \rangle = c_{11}^2 \langle \Phi_G \mid H \mid \Phi_G \rangle + c_{12}^2 \langle \Phi_{12} \mid H \mid \Phi_{12} \rangle + c_{13}^2 \langle \Phi_{13} \mid H \mid \Phi_{13} \rangle + \\ + 2\,c_{11}\,c_{12} \langle \Phi_G \mid H \mid \Phi_{12} \rangle + 2\,c_{11}\,c_{13} \langle \Phi_G \mid H \mid \Phi_{13} \rangle + 2\,c_{12}\,c_{13} \langle \Phi_{12} \mid H \mid \Phi_{13} \rangle ,$$

but after having minimized with two by two rotations

$$\langle \Phi_G \mid H \mid \Phi_{12} \rangle = \langle \Phi_G \mid H \mid \Phi_{13} \rangle = 0 \;;$$

thus we may write

$$\langle \Phi_G' \mid H \mid \Phi_G' \rangle = \langle \Phi_G \mid H \mid \Phi_G \rangle + c_{12}^2 \varDelta_{12} + c_{13}^2 \varDelta_{13} + 2\,c_{12}\,c_{13} \langle \Phi_{12} \mid H \mid \Phi_{13} \rangle$$

where

$$\varDelta_{1i} = \langle \Phi_{1i} \mid H \mid \Phi_{1i} \rangle - \langle \Phi_G \mid H \mid \Phi_G \rangle .$$

It must be (otherwise 3 by 3 rotation is uneffective in lowering the energy)

$$\langle \Phi_G' \mid H \mid \Phi_G' \rangle < \langle \Phi_G \mid H \mid \Phi_G \rangle ,$$

that is, the following condition must be satisfied

$$I = (c_{12}^2 \varDelta_{12} + c_{13}^2 \varDelta_{13} + 2\,c_{12}\,c_{13} \langle \Phi_{12} \mid H \mid \Phi_{13} \rangle) < 0 .$$

Let us introduce the two new variables c_{12}', c_{13}'

$$c_{12}' = c_{12}\,(1 - c_{11}^2)^{-1/2} , \qquad c_{13}' = c_{13}\,(1 - c_{11}^2)^{-1/2}$$

so that

$$c_{12}'^{\,2} + c_{13}'^{\,3} = 1$$

and we can write

$$c_{12}' = \cos\varphi , \qquad c_{13}' = \sin\varphi \;;$$

"I" will then have the form

$$I = (1 - c_{11}^2)\, I'$$

where I' is:

$$I' = \cos^2\varphi\, \varDelta_{12} + \sin^2\varphi\, \varDelta_{13} + 2 \sin\varphi \cos\varphi \langle \Phi_{12} \mid H \mid \Phi_{13} \rangle .$$

Introducing the new wave function

$$\Phi_{I'} = \cos\varphi\, \Phi_{12} + \sin\varphi\, \Phi_{13}$$

we see that its energy is

$$\langle \Phi_{I'} \mid H \mid \Phi_{I'} \rangle = I' + G$$

that is I' will be negative only if C. I. between $\Phi_{1\to 2}$ and $\Phi_{1\to 3}$ gives a state with lower energy than the ground state.

If $\varphi_{\min}$ is found with the C. I. it follows that

$$\langle \Phi_G' \mid H \mid \Phi_G' \rangle = \langle \Phi_G \mid H \mid \Phi_G \rangle - (1 - c_{11}^2) \mid I' (\varphi_{\min}) \mid$$

has the lowest value for $c_{11} = 0$;
then

$$c_{11} = 0 ; \qquad c_{12} = \cos(\varphi_{\min}) ; \qquad c_{13} = \sin(\varphi_{\min}) ,$$

that is

$$\pi_1' = \cos(\varphi_{\min})\, \pi_2 + \sin(\varphi_{\min})\, \pi_3$$

(as was said in *VI*).

2. Combination of one empty orbital π_3 and two filled orbitals π_1 and π_2

$$\Phi'_G = |\pi'_1 \pi'_2| = (c_{11} c_{22} - c_{12} c_{21}) \Phi_G + (c_{12} c_{23} - c_{13} c_{21}) \Phi_{23} + (c_{12} c_{23} - c_{13} c_{21}) \Phi_{13}$$

or, remembering that C is an unitary matrix,

$$\Phi'_G = c_{33} \Phi_G - c_{32} \Phi_{23} + c_{31} \Phi_{13} ;$$

then in the same way as before we get

$$\pi'_1 = \pi_3$$
$$\pi'_2 = \cos(\varphi_{\min}) \pi_1 + \sin(\varphi_{\min}) \pi_2$$

where $\varphi_{\min}$ is found with C. I.

The author is indebted with Professor B. PULLMAN and Dr. A. PULLMAN for their suggestions and helpful advises. He aknowledges also Dr. BERTHIER for useful discussions.

This work prepared in part at the Institut de Biologie Physico-Chimique, Université de Paris, was supported by a grant CY-3073 of the US Public Health Service (National Cancer Institute) to that Institute.

References

[1] BERTHIER, G.: J. chim. Physique **51**, 363 (1954).
[2] BRIOULLIN, L.: La méthode du champ self consistent, Actualité Sc. et Ind. Nr. 71 (1933).
[3] MCWEENY, R.: Proc. Roy. Soc. **A 235**, 496 (1956).
[4] HARTREE, D. R.: The calculation of atomic structure. New York: Wiley 1957.
[5] POPLE, J. A., and R. K. NESBET: J. chem. Physics. **22**, 571 (1954).
[6] ROOTHAAN, C. C. J.: Rev. mod. Physics. **32**, 179 (1960).
[7] — Rev. mod. Physics. **23**, 68 (1951).

(Received June 18, 1965)