Theor Chim Acta (1988) 73: 383-391

Theoretica
Chimica Acta

Studies in the paired orbital method*

II. The energy expression and its derivatives with respect to the nonlinear parameters

Ruben Pauncz**

Department of Applied Mathematics, University of Waterloo, Waterloo,
Ontario, N2L 3G1, Canada

(Received May 14, revised July 27/Accepted August 10, 1987)

The energy expression and its derivatives have been derived in the paired orbital (PO) method. Several relations have been derived for the values of the basic functions at the Hartree-Fock limit. These relations permit to draw conclusions about the behavior of the energy expression as a function of the nonlinear parameters around this limit point and it shows that one can expect an improvement in energy as compared to the Hartree-Fock value using the nonlinear parameters as variational parameters.

Key words: Alternant molecular orbital method — Paired orbital method — Electron correlation

1. Introduction

The paired orbital method (PO) [1, 2] is a variant of the different orbitals for different spins (DODS) approach, introduced by Löwdin [3]. It can be considered as a generalization of the alternant molecular orbital (AMO) method [4-6]. In the AMO method one uses the alternancy symmetry [7, 8], the PO method is valid for any system. The general philosophy of the method was presented in a previous paper [2].

---

[*] Dedicated to Professor J. Koutecký on the occasion of his 65th birthday

[**] *Permanent address*: Department of Chemistry, Technion, Israel Institute of Technology, Haifa, Israel 32000

The goal of the present paper is to consider the energy expression and its derivatives with respect to the nonlinear parameters. The derivatives of the energy are needed when we minimize the energy with respect to these parameters using standard optimization techniques which require analytical expressions for the first and second derivatives of the energy.

After an outline of the basis of the PO method the general energy expression is considered. Next the derivatives are presented and the derivatives of the symmetric sums are obtained. The latter play a central part in these expressions. These quantities are evaluated explicitly for the Hartree-Fock limit (single determinant with doubly occupied orbitals). The knowledge of these quantities allows to draw conclusions about the behavior of the energy as a function of the nonlinear parameters around the Hartree-Fock limit (Taylor expansion of the energy).

2. The PO method

The wavefunction is of the DODS form:

$$\Psi = N\mathscr{A}\Phi\mathscr{O}_S\alpha(1)\cdots\alpha(n)\beta(n+1)\cdots\beta(2n), \tag{1}$$

where $\mathscr{A}$ is the antisymmetrizer, N is a normalization constant and $\mathscr{O}_S$ is the spin projection operator. We shall consider the case of even number of electrons ($N=2n$), and the singlet states ($\mathscr{S}=0$). Φ is a spatial (freeon) wavefunction which is a product of orbitals:

$$\Phi = u_1(1)\cdots u_n(n)v_1(n+1)\cdots v_n(2n). \tag{2}$$

The PO method is obtained when u_i and v_i are formed from a set of orthogonal orbitals in the following way:

$$u_i = \cos\theta_i\psi_i + \sin\theta_i\psi_{i'} \quad \text{and} \quad v_i = \cos\theta_i\psi_i - \sin\theta_i\psi_{i'}, \tag{3}$$

where ψ_i ($i=1,\ldots,n$) is a doubly occupied orbital of the conventional single determinant (Hartree-Fock) wavefunction, and $\psi_{i'}$ is a virtual orbital with which it is paired. For the AMO method the pairing was dictated by the alternancy symmetry [7, 8]. For a general system the pairing is obtained [2] by maximizing the sum:

$$C = \sum_{i=1}^{n} (ii\,|\,i'i'). \tag{4}$$

The algorithm for obtaining the paired orbitals is described in [2]. The ψ_i are chosen as localized orbitals, and the $\psi_{i'}$'s occupy the same region of space as their partners. The $\psi_{i'}$'s are also localized.

3. The paired orbital energy expression

In [2] we have presented the energy expression for the PO method. Since we need the derivatives we shall repeat the main points. Let us introduce the overlap integral between the u_i and v_i:

$$\langle u_i\,|\,v_i\rangle = \lambda_i = \cos 2\theta_i. \tag{5}$$

There are three important functions of the λ_i which appear in the energy expression:

$$\Lambda_q = \sum_{k=0}^{n-q} (-1)^k C(S, k) S_k(x_1, \ldots, x_n), \qquad (q = 0, 1, 2), \tag{6}$$

where $x_i = \lambda_i^2$, $C(S, k)$ are the spin projection (Sanibel) coefficients [9]. For the singlet state they are given as follows:

$$C(S, k) = (-1)^k \binom{n}{k} \Big/ (n+1), \tag{7}$$

and $S_k(x_1, \ldots, x_n)$ (abbreviated as S_k) is the kth symmetric sum formed from $x_1, \ldots, x_n$:

$$\begin{aligned} S_0 &= 1, \\ S_1 &= x_1 + \cdots + x_n, \\ S_2 &= \sum_{i<j}^{n} x_i x_j, \\ S_n &= x_1 \cdots x_n. \end{aligned} \tag{8}$$

Λ_0 is the normalization integral:

$$\Lambda_0 = \langle \Psi_0 | \Psi_0 \rangle. \tag{9}$$

Let us divide the energy expression into two parts. The first one corresponds to the one-electron operators in the Hamiltonian (kinetic energy and nuclear-electron attraction), the second one to the two-electron operators (electron-electron repulsion). The one-electron part is:

$$\begin{aligned} E_1 &= E_{1a} + (E_{1b}/\Lambda_0) \\ &= \sum_{i=1}^{n} w_i - \Lambda_0^{-1} \sum_{i=1}^{n} \Delta w_i \lambda_i \frac{\partial}{\partial x_i} (\Lambda_1 + \Lambda_0). \end{aligned} \tag{10}$$

Here

$$w_i = h_i + h_{i'}, \qquad \Delta w_i = h_{i'} - h_i, \tag{11}$$

and

$$h_i = \int \psi_i^* h \psi_i \, dv, \qquad h_{i'} = \int \psi_{i'}^* h \psi_{i'} \, dv, \tag{12}$$

i.e. they are the diagonal elements of the one-electron Hamiltonian over the occupied localized orbitals and their virtual pairs, respectively.

The two-electron part reads as follows:

$$E_2 = (E_{2a} + E_{2b} + E_{2c}) \Lambda_0^{-1}, \tag{13}$$

where

$$E_{2a}=\sum_{i=1}^{n}[A_{ii}\gamma_{ii}+A_{i'i'}\gamma_{i'i'}-B_i\gamma_{ii'}-C_i\delta_{ii'}], \tag{14}$$

$$E_{2b}=\sum_{i<j}^{n}[\Gamma(i,j)-\Delta(i,j)], \tag{15}$$

$$E_{2c}=-\sum_{i<j}^{n}Z(i,j). \tag{16}$$

$\Gamma(i,j)$, $\Delta(i,j)$, and $Z(i,j)$ are defined as follows:

$$\Gamma(i,j)=\Gamma_{ij}\gamma_{ij}+\Gamma_{ij'}\gamma_{ij'}+\Gamma_{i'j}\gamma_{i'j}+\Gamma_{i'j'}\gamma_{i'j'}, \tag{17}$$

$$\Delta(i,j)=\Delta_{ij}\delta_{ij}+\Delta_{ij'}\delta_{ij'}+\Delta_{i'j}\delta_{i'j}+\Delta_{i'j'}\delta_{i'ji'}, \tag{18}$$

$$Z(i,j)=Z_{ij}\zeta_{ij}+Z_{ij'}\zeta_{ij'}. \tag{19}$$

The basic two electron integrals γ_{ij}, δ_{ij}, and ζ_{ij} are respectively the coulomb, exchange and a new kind of integral which is characteristic of the AMO and PO methods:

$$\gamma_{ij}=(ii|jj), \qquad \delta_{ij}=(ij|ji), \qquad \zeta_{ij}=(ij|j'i'). \tag{20}$$

The $\lambda(\theta)$ dependent factors in Eqs. (14-19) are given as follows:

$$A_{ii}=(1/4)(1+\lambda_i)^2\frac{\partial}{\partial x_i}(\Lambda_1+\Lambda_0), \tag{21}$$

$$A_{i'i'}=(1/4)(1-\lambda_i)^2\frac{\partial}{\partial x_i}(\Lambda_1+\Lambda_0), \tag{22}$$

$$B_i=(1/2)(1-\lambda_i^2)\frac{\partial}{\partial x_i}(\Lambda_1-\Lambda_0), \tag{23}$$

$$C_i=(1-\lambda_i^2)\frac{\partial}{\partial x_i}\Lambda_1, \tag{24}$$

$$\Gamma_{ij}=(1+\lambda_i)(1+\lambda_j)(a+[\lambda_i+\lambda_j]b+\lambda_i\lambda_jc), \tag{25}$$

$$\Delta_{ij}=(1+\lambda_i)(1+\lambda_j)(a+[2(\lambda_i+\lambda_j)-1-\lambda_i\lambda_j]b+\lambda_i\lambda_jc)/2, \tag{26}$$

$$Z_{ij}=\sin 2\theta_i \sin 2\theta_j(a+b)-(1/4)\sin 4\theta_i \sin 4\theta_j(b+c), \tag{27}$$

$$Z_{ij'}=\sin 2\theta_i \sin 2\theta_j(a+b)+(1/4)\sin 4\theta_i \sin 4\theta_j(b+c), \tag{28}$$

where

$$a=\frac{\partial^2}{\partial x_i\,\partial x_j}\Lambda_2, \qquad b=\frac{\partial^2}{\partial x_i\,\partial x_j}\Lambda_1, \qquad c=\frac{\partial^2}{\partial x_i\,\partial x_j}\Lambda_0. \tag{29}$$

From the expressions in Eqs. (25-26) one obtains the corresponding factors involving i' and/or j' by reversing the sign of the λ_i and/or λ_j.

4. The derivatives of the symmetric sums

In Sect. 3 we have seen that the symmetric sums and the Λ_q play a central role in the PO energy expression. The key point in forming the derivatives of the energy will be to obtain expressions for the derivatives of the symmetric sums and the basic functions Λ_q's. We shall also need the derivatives for the special case when all the $\lambda_i = 1$.

Consider first the derivatives of S_k:

$$\frac{\partial S_k}{\partial x_i} = S^i_{(k-1)}(x_1, \dots, x_n). \tag{30}$$

The symmetric sum S_k is linear in x_i, the derivative consists of $k-1$ factors formed from $x_1, \dots, x_{(i-1)}, x_{(i+1)}, \dots, x_n$. The same expression is obtained if we omit from the $S_{(k-1)}$ the terms which contain x_i:

$$\frac{\partial S_k}{\partial x_i} = S_{(k-1)} - x_i \frac{\partial}{\partial x_i} S_{(k-1)}. \tag{31}$$

Using Eq. (31) iteratively we arrive at the following result:

$$\frac{\partial S_k}{\partial x_i} = \sum_{r=0}^{k-1} (-1)^r S_{(k-1-r)} x_i^r. \tag{32}$$

The second derivative of S_k with respect to x_i is zero. Next consider the mixed second derivative of S_k:

$$\begin{aligned}\frac{\partial^2}{\partial x_i\,\partial x_j} S_k &= \sum_{r=0}^{k-1} (-1)^r \frac{\partial}{\partial x_j} S_{k-1-r} x_i^r \\ &= \sum_{r=0}^{k-2} \sum_{s=0}^{k-2-r} (-1)^{r+s} S_{k-2-r-s} x_i^r x_j^s.\end{aligned} \tag{33}$$

Denote $r + s = t$ and change the order of summation:

$$\frac{\partial^2}{\partial x_i\,\partial x_j} S_k = \sum_{t=0}^{k-2} (-1)^t S_{k-2-t} \sum_{r=0}^{t} x_i^r x_j^{t-r}. \tag{34}$$

Define:

$$P_t(x_i, x_j) = \sum_{r=0}^{t} x_i^r x_j^{t-r}. \tag{35}$$

Now the final result reads as follows:

$$\frac{\partial^2}{\partial x_i\,\partial x_j} S_k = \sum_{t=0}^{k-2} (-1)^t S_{k-2-t} P_t(x_i, x_j). \tag{36}$$

The result can be generalized. Let us introduce:

$$P_t(x_1, \dots, x_l) = \sum_{j_1} \cdots \sum_{j_l} x_1^{j_1} \cdots x_l^{j_l}, \qquad (j_1 + \cdots + j_l = t). \tag{37}$$

It is easy to see that the lth derivative of S_k is given as follows:

$$\frac{\partial^l}{\partial x_1 \cdots \partial x_l} S_k = \sum_{t=0}^{k-l} (-1)^t S_{k-l-t} P_t(x_1, \ldots, x_l). \tag{38}$$

These are the only expressions needed for the derivatives of the Λ_q's.

5. The values of the derivatives for $\lambda_i = 1$

Consider the special case when all $\theta_i = 0$ (or $\lambda_i = 1$). We shall denote this point in the parameter space by 0. The value of the symmetric sum is given as:

$$(S_k)_0 = \binom{n}{k}. \tag{39}$$

We shall need a number of identities for the sums of the binomial coefficients. In Eqs. (40)-(42) we list those which will be used later:

$$\sum_{k=0}^{n} \binom{k+a}{k} = \binom{n+a+1}{a+1}, \tag{40}$$

$$\sum_{k=0}^{a} (-1)^k \binom{n}{a-k} \binom{k+b}{k} = \binom{n-b-1}{a}, \tag{41}$$

$$\sum_{k=b}^{n-a} \binom{n-k}{a} \binom{k+b}{k} = \binom{n+1}{a+b+1}. \tag{42}$$

We shall show first that the value of P_t at 0 is given as follows:

$$P_t(x_1, \ldots, x_l)_0 = \binom{t+l-1}{l-1}. \tag{43}$$

It is easy to obtain the value of $P_t(x_1, x_2)_0$:

$$P_t(x_1, x_2)_0 = \left(\sum_{r=0}^{t} x_i^r x_j^{t-r} \right)_0 = \binom{t+1}{1} = t+1. \tag{44}$$

We prove the validity of (43) in an inductive way. Equation (44) shows that the relation is valid for $l = 2$. Assume that it is valid up to l, then we show that it is valid to $(l+1)$.

$$P_t(x_1, \ldots, x_l, x_{l+1}) = \sum_{r=0}^{t} P_r(x_1, \ldots, x_l) x_{l+1}^{t-r}. \tag{45}$$

For all $x_i = 1$ we have:

$$P_t(x_1, \ldots, x_{l+1})_0 = \sum_{r=0}^{t} \binom{r-1+l}{l-1} = \binom{t+l}{l}. \tag{46}$$

Here we have used the binomial identity (40). In the next step we show that the lth derivative of S_k is given by the following expression:

$$\left(\frac{\partial^l}{\partial x_1 \cdots \partial x_l} S_k \right)_0 = \binom{n-l}{k-l}. \tag{47}$$

In order to prove the validity of (47), let us use Eqs. (38) and (43):

$$\left(\frac{\partial^l}{\partial x_1 \cdots \partial x_l} S_k\right)_0 = \sum_{t=0}^{k-1} (-1)^t \binom{n}{k-l-t} \binom{t+l-1}{l-1}. \tag{48}$$

Using the binomial identity (41), we obtain (47).

Finally we consider the derivatives of the Λ_q's at 0. We shall prove that it is given by a very simple expression:

$$\left(\frac{\partial^l}{\partial x_1 \cdots \partial x_l} \Lambda_q\right)_0 = \frac{1}{l+1} \binom{l}{q}^{-1}. \tag{49}$$

In the proof we shall start with the definition of the Λ_q (Eq. (6)) and use Eq. (47):

$$\left(\frac{\partial^l}{\partial x_1 \cdots \partial x_l} \Lambda_q\right)_0 = \frac{1}{n+1} \sum_{k=0}^{n-q} \binom{n}{k}^{-1} \binom{n-l}{k-l}. \tag{50}$$

After some simple algebraic manipulations we obtain:

$$\left(\frac{\partial^l}{\partial x_1 \cdots \partial x_l} \Lambda_q\right)_0 = \frac{(n-l)!(l-q)!q!}{(n+1)!} \sum_{k=0}^{n-q} \binom{k}{l-q} \binom{n-k}{q} = \frac{1}{l+1} \binom{l}{q}^{-1}. \tag{51}$$

In the last step we used the binomial identity (42). It is quite remarkable that the final result is independent of *n*.

6. The derivatives of the energy with respect to the θ_i's

Let us write the energy expression in the following form:

$$E = E_{1a} + \Lambda_0^{-1} E_r = E_{1a} + f(\lambda) + g(\theta), \tag{53}$$

where

$$f(\lambda) = \Lambda_0^{-1}(E_{1b} + E_{2a} + E_{2b}) = \Lambda_0^{-1} E_{la}, \qquad g(\theta) = \Lambda_0^{-1} E_{2c}. \tag{54}$$

We shall see in the following that it is important to take the derivatives with respect to the θ_i's in spite of the fact that the dominant part of the energy (E_{la}/Λ_0) can be expressed as a functions of the λ_i's. The reason is that (E_{2c}/Λ_0) is not analytic in the λ_i's around the point 0.

As a first step evaluate the derivatives of Λ_0:

$$\frac{\partial \Lambda_0}{\partial \theta_i} = -2 \sin 2\theta_i \frac{\partial \Lambda_0}{\partial \lambda_i} = -2 \sin 4\theta_i \frac{\partial \Lambda_0}{\partial x_i}. \tag{55}$$

From Eq. (55) we obtain the second derivatives:

$$\frac{\partial^2 \Lambda_0}{\partial \theta_i \, \partial \theta_j} = 4 \sin 4\theta_i \sin 4\theta_j \frac{\partial^2 \Lambda_0}{\partial x_i \, \partial x_j}, \tag{56}$$

and

$$\frac{\partial^2 \Lambda_0}{\partial \theta_i^2} = -8 \cos 4\theta_i \frac{\partial \Lambda_0}{\partial x_i}. \tag{57}$$

Observe that $\partial^2 \Lambda_0 / \partial x_i^2 = 0$.

At the point 0 the values of these derivatives are given as follows: from Eq. (51) one obtains:

$$\left(\frac{\partial \Lambda_0}{\partial x_i}\right)_0 = 1/2. \tag{58}$$

Substituting these values into Eqs. (55)-(57) we have:

$$\left(\frac{\partial \Lambda_0}{\partial \theta_i}\right)_0 = 0, \qquad \left(\frac{\partial^2 \Lambda_0}{\partial \theta_i\, \partial \theta_j}\right)_0 = 0, \qquad \left(\frac{\partial^2 \Lambda_0}{\partial \theta_i^2}\right)_0 = -4. \tag{59}$$

The general expression for the derivatives of the energy with respect to θ_i's is obtained from Eq. (53) as follows:

$$\frac{\partial E}{\partial \theta_i} = \Lambda_0^{-1} \frac{\partial E_r}{\partial \theta_i} - \Lambda_0^{-2} E_r \frac{\partial \Lambda_0}{\partial \theta_i}. \tag{60}$$

Using Eq. (54) one obtains:

$$\frac{\partial E}{\partial \theta_i} = \Lambda_0^{-1}\left[-2 \sin 2\theta_i \frac{\partial E_{Ia}}{\partial \lambda_i} + \frac{\partial E_{2c}}{\partial \theta_i}\right] + \Lambda_0^{-2} 2 \sin 4\theta_i \frac{\partial \Lambda_0}{\partial x_i} E_r. \tag{61}$$

In order to investigate the behavior of the energy around the point 0, let us consider separately $\partial E_{2c}/\partial \theta_i$. It is evident from Eq. (61) that the other terms vanish at this point. The essential part in $\partial E_{2c}/\partial \theta_i$ is the expression:

$$\begin{aligned}\frac{\partial Z_{ij}}{\partial \theta_k} = \frac{\partial x_k}{\partial \theta_k}\Bigg[& \sin 2\theta_i \sin 2\theta_j \frac{\partial^3(\Lambda_2 + \Lambda_1)}{\partial x_i\, \partial x_j\, \partial x_k} \\ & - (1/4) \sin 4\theta_i \sin 4\theta_j \frac{\partial^3(\Lambda_0 + \Lambda_1)}{\partial x_i\, \partial x_j\, \partial x_k}\Bigg], \qquad k \neq i, j.\end{aligned} \tag{62}$$

The corresponding expression for $Z_{ij'}$ is obtained by changing the sign of the second term. It is easy to see that for all $\theta_i = 0$ both expressions vanish. The same is true for $\partial Z_{ik}/\partial \theta_k$ and $\partial Z_{ik'}/\partial \theta_k$. The final conclusion is that 0 is an extremum point for the energy as a function of the θ_i's.

The nature of the extremum can be investigated considering the second derivatives. From Eq. (60) one obtains:

$$\frac{\partial^2 E}{\partial \theta_i^2} = \Lambda_0^{-1} \frac{\partial^2 E_r}{\partial \theta_i^2} - 2\Lambda_0^{-2} \frac{\partial \Lambda_0}{\partial \theta_i}\left(\frac{\partial E_r}{\partial \theta_i} - E_r \Lambda_0^{-1} \frac{\partial \Lambda_0}{\partial \theta_i}\right) - \Lambda_0^{-2} E_r \frac{\partial^2 \Lambda_0}{\partial \theta_i^2}. \tag{63}$$

Using Eq. (59) and the fact that $(\Lambda_0)_0 = 1$ one obtains:

$$\left(\frac{\partial^2 E}{\partial \theta_i^2}\right)_0 = \left(\frac{\partial^2 E_r}{\partial \theta_i^2}\right)_0 + 4(E_r)_0. \tag{64}$$

Using the explicit form of E_r and the results obtained in Sects. 4-5 one obtains finally a remarkably simple end result for the second derivatives at the point 0:

$$\left(\frac{\partial^2 E}{\partial \theta_i^2}\right)_0 = -4\delta_{ii'}, \qquad \left(\frac{\partial^2 E}{\partial \theta_i\, \partial \theta_j}\right)_0 = -4\zeta_{ij'}. \tag{65}$$

The essential point is that all the elements of the Hessian matrix are negatives, and thus, if we vary the θ_i's around the Hartree-Fock limit we shall certainly be able to make improvement in the energy.

7. Discussion

The essential results of the paper are contained in Sects. 4 and 5. By using these results one can obtain the derivatives of the energy expression with respect to the nonlinear parameters (θ_i's). It was especially interesting to obtain the value of the first and second derivatives at the Hartree-Fock limit.

Acknowledgments. A large part of this research was performed during the author's stay at the Department of Chemistry, University of California, Santa Barbara. The author would like to express his gratitude to Profs. Bernard Kirtman and W. E. Palke for their kind hospitality and for many helpful discussions. It is a pleasure to thank Professors J. Čížek and E. R. Vrscay at this department for their hospitality and for many helpful discussions. The author is grateful to Ms. F. Vinette for her help and advice in the preparation of the manuscript.

References

1. Pauncz R, Chen MB, Parr RG (1982) Proc Natl Acad Sci USA 79:705
2. Pauncz R, Kirtman B, Palke W E (1987) Int J Quantum Chem, accepted
3. Löwdin PO (1955) Phys Rev 97:1474, 1490, 1509
4. Löwdin PO (1954) Symposium on Molecular Physics at Nikko, Japan, Maruzen Co, Tokyo: 13
5. Pauncz R (1967) The alternant molecular orbital method. Saunders, Philadelphia
6. Mayer I (1980) Adv Quantum Chem 12:189
7. Coulson CA, Rushbroke GS (1940) Proc Cambridge Phil Soc 36:193
8. Koutecký J, Paldus J, Čížek J (1985) J Chem Phys 83:1722
9. Manne R (1966) Theor Chim Acta 6:116