Theoret. Chim. Acta (Berl.) 49, 67–76 (1978)

THEORETICA CHIMICA ACTA

A Steepest-Descent Method for the Calculation of Localized Orbitals and Pseudoorbitals

Hermann Stoll, Gerhard Wagenblast and Heinzwerner Preuss

Institut für Theoretische Chemie der Universität Stuttgart, Pfaffenwaldring 55, D-7000 Stuttgart 80, Federal Republic of Germany

A method for direct calculation of localized non-orthogonal orbitals, which has been proposed by the authors recently, is extended to cases where the overlap between different subsystems is very large. This is achieved by using a steepest-descent procedure. In addition, a computationally simple treatment of correlation effects is introduced into the method by means of the density functional formalism. Results of the method are given for e.g. LiH, CH_4, Ne_2, CO, $(FH)_2$.

Key words: Localized orbitals – Steepest-descent method – Density functional

1. Introduction

Localized orbitals are of particular importance for the reduction of the computational effort in *ab initio* HF calculations.

This is due to the following reasons:

Firstly, localized orbitals determined in calculations on small molecules are transferable to large molecule clusters, at least as an appropriate starting point for the SCF iteration [1].

Secondly, molecular symmetry can be taken into account in a rather simple way when using localized orbitals; this is most easily seen for solids with translational symmetry: the localized Wannier functions of a single elementary cell contain all the information on the wave function of the solid [2].

Thirdly, for non-orthogonal localized orbitals a corresponding partitioning of the basis set seems to be possible with relatively small loss in the accuracy of the HF ground state energy [3].

0040-5744/78/0049/0067/$02.00

Finally, the attachment of the localized orbitals to subsystems permits a classification of the interactions into strong intra-group interactions and interactions between orbitals from different subsystems, which should be weaker in comparison.

The latter can be treated in an approximate way, as has been shown in [4].

The direct calculation of localized orbitals was initiated by the work of Adams and Gilbert [5, 6] and has been undertaken by several authors since then [3, 4, 7–17]. The use of a localization operator A, suggested by Adams and Gilbert, leads to two difficulties, however:

Firstly, in many cases A is chosen in such a way that the HF iteration yields localized or delocalized orbitals depending on the initial choice of the orbitals [3, 7, 8]; this problem does not appear if truncated localized basis sets are used for the subsystems, but in this case the degree of localization becomes basis-set dependent.

Secondly, for very strong overlap between different subsystems, serious convergence problems may arise; these problems can be overcome, naturally, by orthogonalizing or semiorthogonalizing basis functions belonging to different localization centres, but in this case the computational effort increases rapidly. In this paper a steepest-descent method is described which avoids the above-mentioned difficulties.

The treatment of correlation has not been considered so far in connection with the direct determination of localized orbitals. The reduction of the computational effort for the HF calculation makes sense only if correlation is included in a rather simple way. A CI calculation [18] or a perturbational treatment as in the PCILO method [19] would involve complicated matrix element evaluation on the *ab initio* level. In this paper we propose an alternative method based on the density functional formalism [20] to evaluate the correlation energy directly.

2. The Method

We give here a brief outline of our method; only those points which differ from our first paper in this series [4] are described in detail.

The procedure for the direct determination of localized HF orbitals can be broken down into the following six steps:

1) A correspondence is established between basis functions, occupied HF orbitals and subsystems. Different subsystems need not necessarily be attributed to different localization centres: our treatment includes the construction of non-orthogonal (pseudo-) orbitals belonging to the same atom. One then makes an initial guess for the localized orbitals; the number of orbital coefficients $C_{p\alpha}$ is relatively small, the coefficients being non-vanishing only if p and α belong to the same subsystem. After that, the orbitals are symmetrically orthogonalized within each subsystem.

2) The Fock matrix is evaluated in blockdiagonal form; matrix elements between different subsystems need not be calculated. For the interaction of a given subsystem with the others a Hartree-type approximation is introduced (exchange terms are neglected; for the Coulomb terms a superposition of least-squares fitted charge densities is used). The computational effort in this stage is given by $\max(n^3N, nN^2)$.

(n: maximum number of basis functions for a single subsystem,
N: total number of basis functions)

For large clusters $N \gg n$ holds, and the computational effort increases only as N^2.

3) For non-orthogonal orbitals the total energy can be expressed by

$$E = \sum_{\alpha,\beta} (F_{\alpha\beta} + h_{\alpha\beta}) S^{-1}_{\beta\alpha} \tag{1}$$

(α, β: orbital indices)

A steepest-descent method employing the energy gradient with respect to the elements of the density matrix has been proposed by McWeeny [21]; the use of the energy gradient with regard to the LCAO coefficients has been discussed by Fletcher [22, 23]. The latter approach seems to be more appropriate in connection with our method than McWeeny's because the orbital-coefficient matrix has block structure which is not true for the density matrix

$$\rho = \sum_{\alpha,\beta} |\alpha\rangle S^{-1}_{\alpha\beta} \langle\beta| \tag{2}$$

For the energy gradient the non-orthogonality of the localized orbitals has to be taken into account; we get:

$$\frac{\partial E}{\partial C_{r\gamma}} = 4(B_{r\gamma} - A_{r\beta} B_{\beta\gamma}) \tag{3}$$

with $A_{r\gamma} = S_{r\alpha} S^{-1}_{\alpha\gamma}$

$B_{r\gamma} = F_{r\alpha} S^{-1}_{\alpha\gamma}$

(α, β, γ: orbital indices, r: basis function index)

The second derivative of E with respect to $C_{r\gamma}$ and $C_{s\gamma}$ (neglecting $\partial F_{pq}/\partial C_{r\gamma}$) yields

$$\begin{aligned}\frac{\partial^2 E}{\partial C_{r\gamma}\partial C_{s\gamma}} &= 4S^{-1}_{\gamma\gamma}(F_{rs} - A_{r\alpha}F_{\alpha s} - F_{r\alpha}A_{\alpha s} + A_{r\alpha}F_{\alpha\beta}A_{\beta s}) \\ &\quad - 4(S^{-1}_{\gamma\alpha}B_{\alpha\gamma})(S_{rs} - A_{r\beta}S_{\beta s}) \\ &\quad - A_{r\gamma}\frac{\partial E}{\partial C_{s\gamma}} - A_{s\gamma}\cdot\frac{\partial E}{\partial C_{r\gamma}}\end{aligned} \tag{4}$$

In the first approximation we insert $S^{-1}_{\alpha\beta} \approx \delta_{\alpha\beta}$ in (1), (3) and (4); all expressions can then be evaluated in terms of the integrals calculated in step 2.

4) In order to improve successively on this first approximation, two-body (and, if required, three- and four-body) corrections are determined. To do that, steps 2 and 3 are repeated, combining pairs (tripels, quadrupels) of subsystems to new extended subsystems. For these new systems correction terms of the following form are added to the inverse overlap matrix:

$$S_{\alpha\beta}^{-1}-\tfrac{1}{2}\sum_{\gamma}\left(S_{\alpha\gamma}S_{\gamma\beta}^{-1}+S_{\alpha\gamma}^{-1}S_{\gamma\beta}\right)\Delta_{\alpha,\beta} \tag{5}$$

$$\Delta_{\alpha,\beta}=\begin{cases}1 \text{ if } \alpha,\ \beta \text{ belong to the same subsystem}\\ 0 \text{ otherwise.}\end{cases}$$

The terms (5) are constructed in such a way that corrections with vanishing trace in the density matrix result; when taken together these corrections build up the exact HF density matrix (2). We have shown in [4] that it is in most cases sufficient to include two-body corrections between neighbouring subsystems in (1).

5) Using the corrected expressions for the energy gradient (3) and the second derivative $\partial^2 E/\partial C_{r\gamma}\partial C_{s\gamma}$ (4), a new set of improved orbitals is generated by solving for each orbital γ

$$\sum_r \Delta C_{r\gamma}\frac{\partial^2 E}{\partial C_{r\gamma}\partial C_{s\gamma}}=-\frac{\partial E}{\partial C_{s\gamma}} \tag{6}$$

Steps 2 to 4 are then repeated until self-consistency is achieved.

6) After the SCF iteration has converged the charge density is used to calculate the correlation energy.

Hohenberg and Kohn [20] have shown that for the ground state of each symmetry type the correlation energy can be evaluated exactly if the exact charge density, the exchange-correlation functional and the HF energy are known. We approximate the exact charge density by the HF density and the correlation functional by the corresponding expression for the electron gas:

$$\Delta E_c=\int \rho\varepsilon_c(\rho)\,d\tau, \tag{7}$$

$\varepsilon_c(\rho)$: correlation energy per unit charge of a homogeneous electron gas with charge density ρ.

For $\varepsilon_c(\rho)$ we employ an analytical fit proposed by Gunnarsson and Lundqvist [25]. Our treatment deviates from that of Gunnarsson and Harris [25, 26] in that we do not approximate both exchange and correlation by the local density functional.

For the evaluation of (7) we break the charge density ρ down to partial densities ρ_i with regard to the subsystems i and integrate using spherical coordinates for each localization centre:

$$\begin{aligned}\Delta E_c &= \sum_i \int \rho_i \varepsilon_c(\rho)\, d\tau \\ &= \sum_i \int dr_i r_i^2 \int \rho_i \varepsilon_c(\rho)\, d\Omega_i \qquad (8)\end{aligned}$$

For the angular integration we average $\varepsilon_c(\rho)$ analytically

$$\rho \rightarrow \bar{\rho} = \frac{\int \rho_i \rho \, d\Omega_i}{\int \rho_i \, d\Omega_i} \qquad (9)$$

Then a numerical integration over r_i follows. The computational effort for the calculation of (7) is proportional to N^2.

3. Applications

The method described above should be useful for large molecule clusters where the computational effort for *ab initio* HF calculations is prohibitive. In order to assess the merits and shortcomings of the method, however, we compare results for some small molecules with results from conventional *ab initio* HF-Roothaan calculations (Tables 1–7).

1) LiH: We have introduced two subsystems here, one for each orbital. To describe the polarization of the 1*s*(H)-orbital, the *p*- and the most diffuse *s*-function on Li have been attributed to subsystem 2. Table 1 shows some points of the potential curve. Our results (E) differ from the conventional *ab initio* results (E_{HF}) by ~10^{-3} a.u. This is due to the restriction which the partitioning of the basis set imposes on the variational freedom.

Table 1. Results for LiH
(All values are given in a.u.)

Basis set: Li: 5*s*/1*p*
H: 5*s*

Definition of the subsystems:

Subsystem	Orbital	Basis functions
1	1*s* (Li)	Li: $s_1, \ldots, s_5$
2	1*s* (H)	H: $s_1, \ldots, s_5$ Li: s_5, p

Total energy E as a function of the LiH-distance r:

r	E	E_{HF}
2.92	−7.96913	−7.97039
3.02	−7.96984	−7.97103
3.22	−7.96893	−7.97005

Table 2. Results for CH_4
(All values are given in a.u.)

Basis set: C: $6s/2sp$ (hybrid)
H: $5s$

Definition of the subsystems:

Subsystem	Orbital	Basis functions
1	$1s$ (C)	C: $s_1, \ldots, s_6$
2...5	σ (CH)	H: $s_1, \ldots, s_5$ C: sp_1, sp_2

Total energy E and equilibrium CH distance r:

	2-body-approx.	3-body-approx.	HF
E	-39.6867	-39.6148	-39.7445
r	1.184	1.181	1.176

Breakdown of the 2- and 3-body corrections

$C-\sigma_{CH}$:	0.325	$\sigma_{CH}-\sigma_{CH}$:	0.277
$\sigma_{CH}-C-\sigma_{CH}$:	0.009	$\sigma_{CH}-\sigma_{CH}-\sigma_{CH}$:	0.005

2) CH_4: Here, too, we have attributed each orbital to a different subsystem. The deviation in the total energy is larger ($\sim 10^{-1}$ a.u.) due to the fact that the overlap between the orbitals is very strong. Nevertheless, the 3-body corrections are small in comparison to the 2-body terms, indicating a rapid convergence of the expansion for the total energy.

3) Ne: All subsystems belong to a common localization centre. In this special case our method can be described as an *ab initio* pseudopotential method. Contrary to the usual pseudopotential procedures, which consider only two subsystems (valence and core), the number of subsystems here is only restricted by the number of orbitals (5 for Ne). The error in the total energy

Table 3. Results for Ne
(All values are given in a.u.)

Basis set: $7s/3p$

Definition of the subsystems:

Subsystem	Orbital	Basis functions
1	$1s$	$s_1, \ldots, s_5$
2	$2s$	$s_6, s_7/s_5, \ldots, s_7$
3...5	$2p$	$p_1, \ldots, p_3$

Total energy:

2-body-approx.	3-body-approx.	HF
$-128.2920/-128.3567$	$-128.1695/-128.2534$	-128.2834

Breakdown of 2- and 3-body corrections:

$1s-2s$	5.9615	$2s-2p$	-0.3308
$1s-2p$	0.0073	$2p_i-2p_j$	-0.0031
$1s-2s-2p$	0.0408		

Table 4. Results for Ne_2
(All values are given in a.u.)

Basis set, subsystems as in Table 3.

Total energy E as function of the Ne–Ne distance r:

r	E (2-body-approx.)	E_{HF}
2	−255.6349	−255.4004
3	−256.5227	−256.4816
4	−256.5799	−256.5641
5	−256.5839	−256.5669
6	−256.5840	−256.5669

Breakdown of 2-body corrections ($\geqslant 0.1$) for $r=3$.

$1s_A/2s_A$	0.9211	$1s_A/2p_{xB}$	−0.2565	$2s_A/2p_{xB}$	0.3890
$2s_A/2p_{xA}$	−0.1622	$2s_A/2s_B$	−0.2196	$2p_{xA}/2p_{xB}$	0.4682

is $\sim 10^{-1}$ a.u., if the 2s-pseudoorbital is built up from two basis functions; if three functions are used, the error decreases to $\sim 10^{-2}$ a.u.

4) Ne_2: The values which we obtain for the repulsive curve are in reasonable agreement with the HF values, although the overlap for $r=2$ and $r=3$ bohr is very strong (this is indicated by the large two-body corrections A/B between different atoms and the large change in the intraatomic two-body corrections A/A).

5) FH: Here one of the subsystems includes two orbitals. The two-body corrections give the total energy with an accuracy of $\sim 10^{-2}$ a.u.

6) $(FH)_2$: The difference between our values and the HF results is rather large here both for the binding energy and the equilibrium F–F distance. These deviations are due to the basis superposition effect, which is not negligible

Table 5. Results for FH
(All values are given in a.u.)

Basis set: F: 4s/2p
H: 2s

Definition of the subsystems:

Subsystem	Orbital	Basis functions
1	1s(F), 2s(F)	F: $s_1,\ldots,s_4$
2	σ(FH)	H: s_1, s_2 F: s_4, p_1, p_2
3, 4	$P_{y,z}$(F)	F: p_1, p_2

Total energy E and equilibrium FH distance r:

	2-body-approx.	HF
E	−96.8159	−96.8373
r	1.855	1.845

Table 6. Results for $(FH)_2$
(All values are given in a.u.)

Basis set as in Table 5.

Geometry: linear, F–H. . .F–H, $d_{FH}=1.733$
each FH monomer defines a subsystem.

Binding energy E_B and equilibrium FH-distance r:

	2-body-approx.	HF
E_B	0.0057	0.0103
r	5.8	5.1

2-body correction at $r=5$: 0.0214

for the HF calculation with the relatively small $4s/2p$ basis set. As our method explicitly avoids basis superposition, it is not surprising that our values are in good agreement with the results of [27] which are near the HF limit ($E_B=0.0055$ a.u., $r=5.5$ b).

7) CO: Our total energy value differs from the HF result by $\sim 10^{-1}$ a.u.; it could be improved probably by adding the s_2-functions to the σ(CO) basis set.

In order to test the correlation density functional derived from the electron gas formula [25], we have calculated correlation energies for atoms (Table 8) and the correlation contribution to the binding energy of diatomic molecules (Table 9). Table 8 shows (in agreement with results of Tong [24]) that the electron gas expression overestimates the correlation energy by a factor 2–3.

This is not surprising: The electron gas formula is only expected to be a good approximation if the electron number is large and the electron density nearly

Table 7. Results for CO
(All values are given in a.u.)

Basis set: C: $(7s/3p)/[4s/2p]$
O: $(7s/3p)/[4s/2p]$

Definition of the subsystems:

Subsystem	Orbital	Basis functions
1	1s(C)	C: s_1, s_2
2	1s(O)	O: s_1, s_2
3	σ(CO)	C: s_3, s_4, p_1, p_2 O: s_3, s_4, p_1, p_2
4, 5	π(CO)	C: p_1, p_2 O: p_1, p_2

Total energy E and equilibrium CO-distance r:

	2-body-approx.	HF
E	−112.4917	−112.5643
r	2.14	2.12

Table 8. Correlation energy of atoms. a) Electron gas approximation, b) exact non-relativistic results [28], c) ratio *a*/*b*
(All values are given in a.u.)

	Basis set	a)	b)	c)
H	6*s*	−0.03576	—	—
He	8*s*	−0.13853	−0.0420	3.3
Li	7*s*	−0.19079	−0.0454	4.2
C	7*s*/3*p*	−0.44611	−0.1551	2.9
N	7*s*/3*p*	−0.53732	−0.1861	2.9
O	7*s*/3*p*	−0.66668	−0.2539	2.6
F	10*s*/5*p*	−0.79160	−0.3160	2.5
Na	10*s*/4*p*	−0.99340	−0.386	2.6

homogeneous; for the atoms of Table 8, however, we have few electrons and, at least near the atomic nucleus, the density gradient is very large.

Nevertheless, the electron gas approximation can be useful if we consider correlation energy differences, e.g. between molecules and the constituent atoms: The changes in the electron density are rather small and homogeneous and the electron number remains constant. Table 9 proves these points: The binding energies become considerably better than the HF binding energies in all cases. The correlation contribution to the binding energy is in general too small, however; the results are best for the alkali diatomics Li_2, Na_2, where ~90% of the correlation contribution is recovered; the error becomes larger with increasing number of valence electrons: we obtain for N_2 ~60% and for F_2 ~50% of the correlation contribution. The reason is the following: we calculate the correlation energy using the HF charge density; this density is a valid starting-point only if the HF determinant is the leading term in the CI expansion of the ground-state wavefunction. If the number of near-degenerate states increases, the accuracy of our results is expected to become poorer.

Table 9. Correlation energy of diatomic molecules. a) Binding energy (correlation treated in the electron gas approximation), b) experimental values of the binding energy, c) contribution of correlation (in electron gas approximation) to the binding energy, d) experimental binding energy minus RHF energy.
(All values are given in a.u.)

	Basis set	a)	b)	c)	d)
H_2	5*s*/1*p*	0.1807	0.1745	0.0477	0.0411
Li_2	7*s*/2*p*	0.0361	0.0386	0.0307	0.0323
LiH	7*s*/2*p*, 5*p*	0.0981	0.0926	0.0464	0.0382
N_2	7*s*/3*p*	0.2024	0.3638	0.1038	0.1702
CO	7*s*/3*p*, 7*s*/3*p*	0.2856	0.4132	0.0708	0.1232
FH	10*s*/5*p*, 5*s*/2*p*	0.1843	0.2249	0.0446	0.0661
F_2	7*s*/3*p*	0.0025	0.0617		
Na_2	10*s*/4*p*	0.0253	0.028		

References

1. Christoffersen, R. E.: Advan. Quantum Chem. **6**, 333 (1972)
2. Wannier, G.: Phys. Rev. **52**, 191 (1937)
3. Matsuoka, O.: J. Chem. Phys. **66**, 1245 (1977)
4. Stoll, H., Preuss, H.: Theoret. Chim. Acta (Berl.) **46**, 11 (1977)
5. Adams, W. H.: J. Chem. Phys. **34**, 89 (1961); Phys. Rev. Letters **32**, 1093 (1974)
6. Gilbert, T. L.: in: Molecular orbitals in chemistry, physics and biology, Löwdin, P., Pullman, B. eds. New York: Academic Press 1964
7. Kunz, A. B.: Phys. Stat. Sol. **36**, 301 (1969); **46**, 385 (1971)
8. Kunz, A. B.: J. Phys. **B6**, L47 (1973)
9. Gilbert, T. L., Kunz, A. B.: Phys. Rev. **B10**, 3706 (1974)
10. Niessen, W. v.: J. Chem. Phys. **56**, 4290 (1970)
11. Niessen, W. v.: Theoret. Chim. Acta (Berl.) **33**, 7 (1974)
12. Peters, D.: J. Chem. Phys. **46**, 4427 (1968)
13. Wilhite, D. L., Whitten, J. L.: J. Chem. Phys. **58**, 948 (1973)
14. Anderson, P. W.: Phys. Rev. **181**, 25 (1969)
15. Daudey, J. P.: Chem. Phys. Letters **24**, 574 (1974)
16. Kambara, T.: J. Phys. Soc. Japan **39**, 187 (1975); Progr. Theoret. Phys. **55**, 1335 (1976)
17. Mehler, E. L.: J. Chem. Phys. **67**, 2728 (1977)
18. Meyer, W.: in: Modern theoretical chemistry, Vol. 3, Schaefer III, H. F. ed. New York: Plenum Press 1977
19. Diner, S., Malrieu, J. P., Claverie, P.: Theoret. Chim. Acta (Berl.) **13**, 1 (1969)
20. Hohenberg, P., Kohn, W.: Phys. Rev. **136**, B864 (1964)
21. McWeeny, R.: Proc. Roy. Soc. (London) **A235**, 496 (1956)
22. Fletcher, R.: Mol. Phys. **19**, 55 (1970)
23. Cooper, J. A. R., Sabo, D. W.: Chem. Phys. Letters **45**, 473 (1977)
24. Tong, B. Y., Sham, L. J.: Phys. Rev. **144**, 1 (1966)
25. Gunnarsson, O., Lundqvist, B. I.: Phys. Rev. **B13**, 4274 (1976)
26. Gunnarsson, O., Harris, J., Jones, R. O.: Phys. Rev. **B15**, 3027 (1977)
27. Yarkony, D. R., O'Neil, S. V., Schaefer III, H. F., Baskin, C. P., Bender, C. F.: J. Chem. Phys. **60**, 855 (1974)
28. Veillard, A., Clementi, E.: J. Chem. Phys. **49**, 2415 (1968)

Received January 4, 1978/March 1, 1978