Studies in Electronegativity. II. Comparison of the OCEE and the CNDO Schemes*

BJØRN VOIGT

Department of Physical Chemistry, H. C. Ørsted Institute, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark

Fundamental problems in a comparison of results from the Overlap Corrected Electronegativity Equalization (OCEE) and the CNDO methods are discussed. The proper basis functions for the OCEE scheme are derived and employed in the calculation of Coulomb and exchange integrals for the water molecule. The results are compared with those obtained using the symmetrically orthogonalized (Löwdin-) basis appropriate to the CNDO approximation.

1. INTRODUCTION

The analysis of the principle of electronegativity equalization¹ which was presented in the initial publication in this series² lead to the proposal of the Overlap Corrected Electronegativity Equalization (OCEE) scheme. In this semi-empirical scheme a description of the molecular electronic structure by means of localized molecular orbitals (LMO's) is attempted. The LMO's are constructed from atomic hybrid orbitals (AHO's) with explicit inclusion of the overlaps between the bonded orbitals. Overlap integrals between non-bonded AHO's, on the other hand, are neglected.

In the following publication in the series³ a detailed comparison of results from the OCEE and the CNDO/2⁴ schemes is carried out. Such a comparison presents some fundamental problems which are discussed below.

2. THE PROBLEM OF TRUE BASIS FUNCTIONS

The differences in the treatment of overlaps in the OCEE and CNDO schemes imply dif-

ferences in the true basis functions assumed in the two theories. It is important that these differences be recognized. The direct output of the actual calculations consists of sets of coefficients. The quantities to be compared, however, are derived from the wavefunctions. A detailed comparison of results from the two schemes is thus only possible if the proper basis orbitals of both schemes are known so that the transitions from the coefficients to the wavefunctions can be made.

In the case of the CNDO scheme, it has been demonstrated⁵ that the approximations introduced may be rationalized if it is assumed that the true basis is that obtained by Löwdin's method of symmetrical orthogonalization.⁶ Thus if

$$\Phi = \{\phi_1, \phi_2, \dots, \phi_m\}$$

denotes the row vector of the original, non-orthogonal atomic basis functions with the metric matrix

$$A \equiv \langle \Phi | \Phi \rangle = \begin{bmatrix} 1 & S_{12} & \dots & S_{1m} \\ S_{21} & 1 & \dots & S_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \dots & 1 \end{bmatrix}$$

then the proper CNDO-basis is

$$\Phi_{\text{CNDO}} = \Phi A^{-\frac{1}{2}} \quad (1)$$

Since it is immaterial whether hybridization is performed before or after symmetrical orthogonalization, Φ may in general be assumed to be made up of atomic hybrid orbitals.

In the OCEE theory the various types of overlaps are not treated equally. Accordingly

* Based in part on a thesis presented to the University of Copenhagen in partial fulfillment of the requirements for the lic. scient. degree.

the matrix of bonded overlaps $\mathbf{T}$, the elements of which are included explicitly in the MO's, must be separated from that containing the nonbonded overlaps which are neglected:

$$\mathbf{A} \equiv \mathbf{1} + \mathbf{T} + \mathbf{S}$$

With proper ordering of basis functions, $\mathbf{T}$ will be block diagonal with 2×2 blocks. $\mathbf{S}$ has a similar block structure the elements within the blocks in this case all being zero. Although $\mathbf{T}$ and $\mathbf{S}$ do not, in general, commute, the elements of $\mathbf{TS} - \mathbf{ST}$ are usually small. If all bonded overlaps are equal and all non-bonded overlaps are equal, then the commutator vanishes identically.

The metric matrix for the proper OCEE-basis is given by

$$\langle \Phi_{\text{OCEE}} | \Phi_{\text{OCEE}} \rangle = \mathbf{1} + \mathbf{T}$$

If $\Phi_{\text{OCEE}} \equiv \Phi \mathbf{A}^{-1} \mathbf{A} = \Phi_{\text{CNDO}} \mathbf{A}$ (2)

then the following equation, which should be satisfied by the transformation matrix $\mathbf{A}$, is obtained:

$$\mathbf{A}^\dagger \mathbf{A} = \mathbf{1} + \mathbf{T} \quad (3)$$

Two solutions to this equation, namely

$$\mathbf{A}_1 = [\mathbf{A}^\dagger (\mathbf{1} + \mathbf{T}) \mathbf{A}]^{\frac{1}{2}} \mathbf{A}^{-\frac{1}{2}} \quad (4)$$

and

$$\mathbf{A}_2 = (\mathbf{1} + \mathbf{T})^{\frac{1}{2}}$$

are of special interest in the present context. They are identical when $\mathbf{S}$ and $\mathbf{T}$ commute. By expanding the square roots in series (assuming convergence), and by retaining contributions up to and including terms of second order in the overlaps,

$$\mathbf{A}_1 - \mathbf{A}_2 \cong \frac{1}{2} (\mathbf{ST} - \mathbf{TS}) \quad (5)$$

It may be shown (see the appendix) that the solution $\mathbf{A}_1$ gives functions which resemble the original basis functions as closely as is possible; the true OCEE-basis must therefore be given by this solution. The basis set obtained in this way is referred to as the Non-Bonded Symmetrical Orthogonal (NBSO)-basis. However, in view of the relation (5) and the near-commutation of $\mathbf{S}$ and $\mathbf{T}$ found in practice, the transformation matrix $\mathbf{A}_2$ may to a very good approximation be used instead.

Eqn. (2) now reads:

$$\Phi_{\text{CNDO}} \cong \Phi_{\text{OCEE}} (\mathbf{1} + \mathbf{T})^{-\frac{1}{2}} \quad (6)$$

That is, the CNDO-basis corresponding to the OCEE-hybridization may be obtained approximately by symmetrically orthogonalizing the OCEE-functions.

Because of the block structure of $\mathbf{1} + \mathbf{T}$, the OCEE-LMO's will remain localized when expressed in the CNDO type basis. In particular the OCEE lone pair functions will be identical to the corresponding CNDO type hybrids. Any delocalizations of orbitals found in the CNDO wavefunction therefore represent genuine differences from the OCEE description and are not merely due to differences in basis sets. As a further consequence of (6) comparison of the hybridization schemes of the two calculations is also justifiable insofar as it is possible to extract hybridization informations from the CNDO results (see Ref. 3).

In order to compare the charge distributions the symmetrical orthogonalization (6) must be performed explicitly. Using primes to denote CNDO type functions, an expression for the OCEE-LMO ψ_{AB} is obtained as:

$$\psi_{AB} = C_A \phi_A + C_B \phi_B = C'_A \phi'_A + C'_B \phi'_B$$

where

$$\begin{aligned} \begin{pmatrix} C'_A \\ C'_B \end{pmatrix} &= \frac{1}{2} \left(\sqrt{1 + T_{AB}} \pm \sqrt{1 - T_{AB}} \right) \begin{pmatrix} C_A \\ C_B \end{pmatrix} \\ &+ \frac{1}{2} \left(\sqrt{1 + T_{AB}} \mp \sqrt{1 - T_{AB}} \right) \begin{pmatrix} C_B \\ C_A \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} \begin{pmatrix} \phi'_A \\ \phi'_B \end{pmatrix} &= \frac{1}{2} \left[\frac{1}{\sqrt{1 + T_{AB}}} \pm \frac{1}{\sqrt{1 - T_{AB}}} \right] \begin{pmatrix} \phi_A \\ \phi_B \end{pmatrix} \\ &+ \frac{1}{2} \left[\frac{1}{\sqrt{1 + T_{AB}}} \mp \frac{1}{\sqrt{1 - T_{AB}}} \right] \begin{pmatrix} \phi_B \\ \phi_A \end{pmatrix} \end{aligned}$$

T_{AB} is the overlap integral between the (unprimed) OCEE orbitals.

In the OCEE-scheme overlap charges are divided equally between the orbitals participating. In a CNDO-scheme no overlap contributions are present. Thus the charges ascribed to atom A from two electrons occupying the LMO above are, for each of the schemes,

$$q_A = 2C_A^2 + 2C_A C_B T_{AB}$$

and

$$q'_A = 2C_A'^2 = q_A - \frac{1}{2}(q_A - q_B)(1 - \sqrt{1 - T_{AB}^2}) \quad (7)$$

Although the difference $q'_A - q_A$ is of second order in the overlap it may still be significant since for σ -bonds T_{AB} may be as large as 0.7–0.8. For a sensible comparison with CNDO results, the OCEE charges should therefore be calculated from (7), that is, relative to the CNDO type basis.

3. TWO-ELECTRON INTERACTIONS

The CNDO approximation represents a powerful way of reducing the number of two-electron integrals necessary in molecular calculations. Only the Coulomb interactions between AO's are retained in this approximation. In the OCEE scheme, the application of the Ruedenberg approximation, which is a generalization of the Mulliken approximation to two-electron distributions,⁷ formally makes it possible to retain even the exchange interactions between AO's.^{2,3} However, when calculated using the proper basis set these interactions are likely to be negligible since the orbitals involved are orthogonal.

Before any steps were taken concerning the exchange integrals in the OCEE scheme the properties of the NBSO basis were investigated in more detail. This was achieved by performing an explicit calculation of integrals for H_2O using the proper basis orbitals (as given by (2) and (4) in Section 2). The H_2O molecule was chosen as a test case since its small bond angle leads to relatively large non-bonded overlaps. The resulting integral values, together with those calculated using the non-orthogonal and the completely symmetrically orthogonalized orbitals (Löwdin-orbitals), are listed in Table 1.

Of prime concern are the exchange integrals between non-bonded orbitals. In the two-center case, these are greatly reduced upon orbital orthogonalization using either of the methods. The values obtained from the NBSO-basis, however, are consistently about twice those calculated using the Löwdin-orbitals. That is, the differential overlap between functions on different centers is somewhat smaller in the latter basis.

The one-center exchange integrals show quite different behaviour, the non-bonded symmet-

Table 1. Coulomb and exchange integrals (a.u.) for H_2O .

Integral ^a	Non-orthogonal orbitals	NBSO orbitals ^b	Löwdin orbitals ^b
$[h_1h_1 h_1h_1]$	0.7500	0.7834	0.7956
$[b_1b_1 b_1b_1]$	0.9843	0.9984	1.0511
$[l_1l_1 l_1l_1]$	1.0113	1.0123	1.0146
$[h_1h_1 h_2h_2]$	0.3409	0.3234	0.2972
$[h_1h_1 b_2b_2]$	0.6303	0.6243	0.5406
$[h_1h_1 b_2b_2]$	0.4624	0.4462	0.4255
$[h_1h_1 l_1l_1]$	0.4534	0.4399	0.4110
$[b_1b_1 b_2b_2]$	0.7806	0.7821	0.8592
$[b_1b_1 l_1l_1]$	0.7705	0.7723	0.8089
$[l_1l_1 l_2l_2]$	0.7537	0.7553	0.7503
$[h_1h_2 h_1h_2]$	0.0398	0.0031	0.0016
$[h_1b_1 h_1b_1]$	0.2709	0.2587	0.0142
$[h_1b_2 h_1b_2]$	0.0289	0.0095	0.0045
$[h_1l_1 h_1l_1]$	0.0314	0.0102	0.0061
$[b_1b_2 b_1b_2]$	0.0752	0.0759	0.0909
$[b_1l_1 b_1l_1]$	0.0753	0.0743	0.0826
$[l_1l_2 l_1l_2]$	0.0688	0.0682	0.0660

^a Notation: b and l are bonding resp. lone pair O-hybrids, b points along the OH-axis; h denotes the H 1s-orbital. Orbital exponents are 2.275 and 1.2 respectively. ^b See Section 2.

rical orthogonalization having virtually no effect on the computed values. Complete symmetrical orthogonalization, however, leads to significantly larger integrals, corresponding to larger differential overlap between the functions. The importance of the one-center exchange interactions has of course been realized earlier. This has led to the proposal of the INDO (Intermediate . . .) scheme.⁸

It follows from the above analysis that in the OCEE theory neglect of exchange seems justified only for two-center integrals. Nevertheless, for convenience even the one-center contributions have been discarded in the calculations published in Ref. 3, a rough correction for the consequent error being made *via* the integral parametrization.

The Coulomb integrals have been included in Table 1 since they demonstrate more clearly than the exchange integrals the similarity of the NBSO and the original non-orthogonal basis. Comparison of the results obtained for these integrals reveals the value of explicitly including, as far as possible, the overlap effect in the MO's. Most semi-empirical schemes

published to date conceal all overlap corrections in a hypothetical basis set which is not, in fact, used in actual computations.

Acknowledgement. The author is grateful to Dr. H. Johansen for supplying most of the programs with which the results in Table 1 were obtained. Thanks are also due to Professor J. P. Dahl for his contributions in many stimulating discussions.

APPENDIX

It is to be demonstrated that when eqn. (4) is applied in (2), the resulting functions Φ_{OCEE} resemble the original functions Φ as closely as is possible. That is, that

$$\text{Tr } \mathbf{M} \equiv \text{Tr} \langle \Phi_{\text{OCEE}} - \Phi | \Phi_{\text{OCEE}} - \Phi \rangle =$$

$$\sum_{i=1}^m \int |\phi_i^{\text{OCEE}} - \phi_i|^2 dv$$

takes on its minimum value for this particular solution.

The matrices $\mathbf{A}$ and $\mathbf{1} + \mathbf{T}$ are positive definite. Using the block structure of $\mathbf{T}$ it may be shown that

$$\mathbf{B} \equiv \mathbf{A}^{\frac{1}{2}} (\mathbf{1} + \mathbf{T}) \mathbf{A}^{\frac{1}{2}}$$

is also positive definite (all the square roots are positive). The general solution of eqn. (3) may thus be expressed as

$$\mathbf{A} = \mathbf{U} \mathbf{B}^{\frac{1}{2}} \mathbf{A}^{-\frac{1}{2}} \quad (\text{A.1})$$

where $\mathbf{U}$ is an arbitrary unitary matrix. For $\mathbf{U} = \mathbf{1}$ (A.1) reduces to (4). Eqns. (A.1) and (2) may be used to derive the relation

$$\begin{aligned} \mathbf{M} = & \mathbf{A} + \mathbf{1} + \mathbf{T} - 2\mathbf{B}^{\frac{1}{2}} + \mathbf{B}^{\frac{1}{2}} (\mathbf{1} - \mathbf{U})^{\dagger} (\mathbf{1} - \mathbf{U}) \mathbf{B}^{\frac{1}{2}} + \\ & [\mathbf{B}^{\frac{1}{2}} \mathbf{U}^{\dagger} \mathbf{B}^{\frac{1}{2}} - \mathbf{A}^{-\frac{1}{2}} \mathbf{B}^{\frac{1}{2}} \mathbf{U}^{\dagger} \mathbf{A}^{\frac{1}{2}}] + [\mathbf{B}^{\frac{1}{2}} \mathbf{U} \mathbf{B}^{\frac{1}{2}} - \mathbf{A}^{\frac{1}{2}} \mathbf{U} \mathbf{B}^{\frac{1}{2}} \mathbf{A}^{-\frac{1}{2}}] \end{aligned} \quad (\text{A.2})$$

There will be no contributions to the trace of $\mathbf{M}$ from the last two terms in (A.2), since the trace of a product of matrices is unchanged by a cyclic permutation of the factors. Thus, defining

$$\mathbf{C} \equiv (\mathbf{1} - \mathbf{U}) \mathbf{B}^{\frac{1}{2}}$$

it is found that

$$\text{Tr } \mathbf{M} = \text{Tr} [\mathbf{A} + \mathbf{1} + \mathbf{T} - 2\mathbf{B}^{\frac{1}{2}}] + \text{Tr} [\mathbf{C}^{\dagger} \mathbf{C}].$$

Finally, since

$$\text{Tr} [\mathbf{C}^{\dagger} \mathbf{C}] = \sum_j \sum_i |C_{ij}|^2 \geq 0$$

the trace of $\mathbf{M}$ will be a minimum in the cases where all elements of $\mathbf{C}$ are zero. Such a case arises for $\mathbf{U} = \mathbf{1}$. *Q.E.D.*

REFERENCES

1. Sanderson, R. T. *Science* 114 (1951) 670.
2. Voigt, B. and Dahl, J. P. *Acta Chem. Scand.* 26 (1972) 2923.
3. Voigt, B. *Acta Chem. Scand. A* 28 (1974) 1068.
4. Pople, J. A., Santry, D. P. and Segal, G. A. *J. Chem. Phys.* 43 (1965) S129; Pople, J. A. and Segal, G. A. *Ibid.* 43 (1965) S136; 44 (1966) 3289.
5. Dahl, J. P. *Acta Chem. Scand.* 21 (1967) 1244.
6. Löwdin, P. O. *Sv. Kem. Tidskr.* 67 (1955) 380.
7. Ruedenberg, K. *J. Chem. Phys.* 19 (1951) 1433.
8. Pople, J. A., Beveridge, D. L. and Dobosh, P. A. *J. Chem. Phys.* 47 (1967) 158.

Received June 4, 1974.