

## IS IT POSSIBLE TO CREATE A BETTER SEMIEMPIRICAL MO THEORY ?

Einar Lindholm

Physics Department, The Royal Institute of Technology  
S-100 44 Stockholm, Sweden

(Received in USA 30 March 1988)

**Abstract.** An outline is given of a future semiempirical MO method, intended for use in organic chemistry.

Dewar /1,2/ has pointed out that it seems to be impossible to calculate the total energy of a molecule with sufficient accuracy by use of ab-initio CI methods. For e.g. benzene the error amounts to 1031 kcal/mol /1/. It is true that cancellation of errors takes place when energy differences are considered, but the cancellation is never perfect. Therefore, calculations of properties like excitation energies and electron affinities have appeared to have very large errors when compared to experiments. It would therefore be desirable to have sufficiently accurate semiempirical methods in which the correlation energies have been included.

Even if the methods, developed by Dewar and coworkers, have appeared to be extremely useful, it is of interest to discuss how a future MO method could be developed. We will below discuss some features of such a method.

1. Since correlation energies should be included, the theory cannot start from the Hartree-Fock theory, since the addition afterwards of correlation energies has appeared to be difficult. Instead, the treatment has to start from density functional theory /3,4/. The resulting formalism is then similar to the Hartree-Fock theory, but the correlation energies are now included /5/. This proves in a mathematical way the possibility to create a MO method in which all correlation is included, although the orbitals are one-particle orbitals. Addition afterwards of CI to such a method is therefore unnecessary.

2. Since the method must be accurate, the ZDO approximation cannot be used.

3. The LCAO approximation must, on the other hand, be used. The total energy expression will then contain one-center and two-center parameters. They should be completely separated so that the total energy expression contains terms, which are truly one-center, and other terms, which are truly two-center (for details see App.A.).

4. The one-center parameters are determined first. The general total energy expression for a molecule is applied to a one-atom molecule. This total energy is compared with experimental total energies for atoms. Since the situation of an atom is different in different molecules, it is not sufficient to compare only with the ground state energy of the atom. It is necessary to use also many excited and ionized states of every atom. This will probably require a large number of one-center parameters, but the advantage is that the atom is correctly described even when it is part of a molecule. In this treatment all core

electrons should be included, which can be done with little consumption of computer time. The procedure is possible as shown by Åsbrink /6,7/. He used 2200 atomic configurations and obtained a very good accuracy (for the carbon atom of the order of 0.03 eV).

5. The determination of the two-center (molecular) parameters by study of molecules has to be performed in two steps.

In the first step certain molecular parameters are determined by study of ionization energies, excitation energies and electron affinities, but the remaining parameters are not present in the expressions used here. They are determined in the second step in which molecular total energies are used (for details see App.B.)

6. In the study of ionization energies and electron affinities Koopmans' theorem may not be used. We have assumed above that our method shall include all correlation energy. It is then necessary to calculate the reorganization energy exactly by use of  $\Delta E(\text{SCF})$  calculations /5/.

7. The orbital energies of the occupied orbitals are related to the experimental ionization energies which are well known. Also "ESCA" ionizations can be included.

8. The orbital energies of the unoccupied orbitals are related to experimental electron affinities of molecules. They have up till now been extremely little known due to experimental difficulties. It is, however, absolutely necessary to include also information related to them. If not, the parametrization may go wrong, even if the main part of the method has been correctly parametrized (IIAM/3 is an example). It is therefore valuable that recently some progress has been made concerning experimental electron affinities /8/. It appeared later that also core excitation can be used to study the unoccupied orbitals in a molecule /9/.

Excitation energies are of course well known but concern usually only  $\pi^*$  orbitals.

9. There is no reason to believe that the parameter  $\beta$  (see App.A.) should be directly proportional to the overlap  $S_{\mu\nu}$ .

10. The remaining parameters are determined by comparison of the calculated total energy with the experimental value for many molecules. Also geometries can be used since the correlation energies do not influence the second step (cf./10/). Such work has been performed with great success by Dewar and coworkers. However, the parametrization, which we propose here, should be easier than in Dewar's work, since many parameters have been fixed by the preceding parametrization procedures, described above. It is our belief that if the preceding parametrization has been performed in a correct way, the resulting method describes Nature well and the final parametrization will be easy.

#### Conclusions.

Since the procedure, proposed here, is based upon density functional theory and no parts of the total energy expressions are omitted, a very good accuracy can be expected. The only serious approximation is the introduction of LCAO, which probably cannot be avoided. This error is, however, compensated by the fact that the atomic orbitals, due to the accurate parametrization mentioned in 4. above, are always optimized.

The old statement that it is necessary to have one semiempirical method for each property of a molecule has its ground in unlogical thinking. For example, the use of Koopmans' theorem implies the hope that two unknown errors cancel each other. It is obvious that a method, deduced by use of Koopmans' theorem, cannot be used for properties for which such a cancellation does not take place. In the same way, the approximation  $V \approx \gamma$  (see App.B.) will make the method less accurate for ioni-

zation energies and electron affinities, although other properties may be calculated with good results.

In his work, Michael Dewar has been able to show that it is possible to create very useful semiempirical methods. His thinking has been very inspiring to many of us in our efforts to study molecules.

#### References.

1. M.J.S. Dewar and D.M. Storch, *J.Am.Chem.Soc.* 107, 3898 (1985).
2. M.J.S. Dewar, *J.Phys.Chem.* 89, 2145 (1985).
3. E. Lindholm and S. Lundqvist, *Phys.Scripta* 32, 220 (1985).
4. E. Lindholm and S. Lundqvist, in "Density Matrices and Density Functionals", (eds. V.H. Smith and R. Erdahl), Reidel Publishing Co., Dordrecht (1987), p.507.
5. E. Lindholm and L. Åsbrink, "Molecular Orbitals and their Energies, Studied by the Semiempirical HAM Method", Springer-Verlag, Heidelberg, (1985).
6. L. Åsbrink, *Phys.Scripta* 28, 394 (1983).
7. L. Åsbrink, TRITA-FYS-1001, Physics Department, The Royal Institute of Technology, Stockholm (1980).
8. E. Lindholm and J. Li, *J.Phys.Chem.* 92, 1731 (1988).
9. E. Lindholm ( in preparation ).
10. M. Levy and J.P. Perdew, *J.Chem.Phys.* 84, 4519 (1986).
11. D.E. Parry and J.P. Whitehead, *J.Comp.Chem.* 3, 265 (1982).
12. D.P. Chong, private communication.
13. L. Åsbrink, C. Fridh, E. Lindholm, S. de Bruijn and D.P. Chong, *Phys. Scripta* 22, 475 (1980).

#### Appendix A.

The statement in 3. is perhaps not immediately obvious. We will therefore illustrate it by an example.

We denote the kinetic energy operator in the Hamiltonian as  $\nabla^2$  and denote two atomic orbitals by  $\phi_\mu$  and  $\phi_\nu$ , where  $\mu$  is on atom A and  $\nu$  on atom B, ( $A \neq B$ ). In the total energy expression there is a term

$$\int \phi_\mu^* \nabla^2 \phi_\nu d\tau$$

which is of mixed one-center and two-center type. We can, however, apply Mulliken's approximation and put

$$\int \phi_\mu^* \nabla^2 \phi_\nu d\tau = \frac{1}{2} S_{\mu\nu} \left[ \int \phi_\mu^* \nabla^2 \phi_\mu d\tau + \int \phi_\nu^* \nabla^2 \phi_\nu d\tau \right] + \beta \quad (1)$$

It is known that Mulliken's approximation is very inaccurate, but the addition of the correction term  $\beta$  makes the relation exact. Since  $\beta$  depends upon  $\mu$ ,  $\nu$  and the internuclear distance R, it is of two-center type. The first two terms are of one-center type.

The terms, resulting from the attractions  $Z_A r_{AA}^{-1}$  in the Hamiltonian, can be handled in a similar way / 5, 6 /.

The terms which describe the electron-electron repulsions contain also three and four center terms. It appears, however, that in a complete treatment ( ref.5, page 80, or ref.13, page 479 ) the most important terms of these types cancel each other completely.

The result is that the general total energy expression, deduced from Hartree-Fock or density functional theory, is

$$E = \text{one-center terms} - \quad (2)$$

$$- \sum_{\mu, \nu, \beta} P_{\mu\nu} \beta_{\mu\nu} + \sum_{A>B} \left[ \sum_{\mu}^A N_{\mu} \sum_{\nu}^B N_{\nu} \gamma_{AB} - Z_B \sum_{\mu}^A N_{\mu} V_{AB} \right] + \quad (3)$$

$$+ \sum_{A>B} \left[ Z_A Z_B R_{AB}^{-1} - Z_A \sum_{\nu}^B N_{\nu} V_{BA} \right] \quad (4)$$

+ two small correction terms, one one-center and one two-center.

Here,  $\gamma_{AB}$  denotes the repulsion between one electron around A and another around B,  $V_{AB}$  the repulsion between one positive charge of  $Z_B$  and one electron around A, and  $N_{\mu}$  the total gross population belonging to  $\phi_{\mu}$  on A.

#### Appendix B.

The first step means a study of the molecule in an SCF calculation. The Fock matrix elements are deduced from eqs. (2) and (3) and the parameters in these expressions are determined from comparisons with experiments. This means (besides the  $\beta$  parameters) in principle the difference  $\gamma_{AB} - V_{AB}$ . A calculated value can be used as a first approximation to this difference /11,12/. The "penetration" difficulties are thus avoided.

The second step means a study of the total energy (heat of formation) of a molecule by comparing eqs. (2), (3) and (4) with experiment. The parameters in eq. (4) can now be determined. This means in principle the difference  $R_{AB}^{-1} - V_{BA}$ .