Theoret. Chim. Acta. (Berl.) 52, 319-328 (1979)

Geometry Optimizations with Explicit Inclusion of Electron Correlation

Giinther Lauer, Karl-Wilhelm Schulte, Armin Schweig, and Walter Thiel

Fachbereich Physikalische Chemie der Universität, Auf den Lahnbergen, D-3550 Marburg, Federal Republic of Germany

Automatic techniques for geometry optimization are applied in conjunction with configuration interaction and perturbation treatments of electron correlation. The computational effort and numerical accuracy of the optimizations are discussed, as well as problems with approximate correlation methods concerning the continuity of the potential surface. The optimized geometries of fourteen molecules obtained with different correlation treatments (MNDO SCF MOs) are compared. The configuration interaction results are reproduced satisfactorily by simple perturbation approaches. The largest change of the optimized SCF geometry is found for hydrogen peroxide.

Key words: Configuration interaction - Second-order perturbation theory

1. Introduction

The investigation of potential surfaces plays a central role in theoretical treatments of chemical rate processes. For a given reaction, the local minima corresponding to reactants and products, and the transition state for their interconversion must be located on the surface. A complete geometry optimization for these stationary points is required to obtain meaningful results.

The most efficient of the optimization procedures available [1, 2] make use of the gradient of the potential energy with respect to the geometric variables. The geometries of stable molecules are found by minimizing their energy using a quasi-Newton method such as the Davidson-Fletcher-Powell (DFP) method [3, 4] or the Murtagh-Sargent method [5]. Transition states can be located by minimizing their gradient norm [6] via a generalized least-squares algorithm [2, 7].

Different variants of these optimization procedures have been applied in conjunction with semiempirical SCF wavefunctions [6, 8-10] as well as *ab initio* SCF wavefunctions [11-15], but there seems to be only little experience with correlated wavefunctions [16]. In the present paper, we describe a systematic study of geometry optimizations with explicit inclusion of electron correlation based on semiempirical molecular orbitals (MOs). Since the semiempirical methods used have been parametrized at the SCF level, an improved agreement between calculated and observed geometries cannot be expected. Therefore the present study is primarily concerned with the optimization techniques applied, the performance of different correlation treatments, and the general trends of the numerical results.

2. Computational Methods

Our program package allows SCF calculations to be carried out by the semiempirical MNDO [17], MINDO/3 [18], and CNDO/2 [19] procedures. Electron correlation can be treated explicitly by defining a set of main configurations and including the effects of the singly and doubly excited configurations in the following manner:

- a) CI: Configuration interaction with all configurations generated.
- b) SELCI: Configuration interaction with selected configurations [20].
- c) PERTCI: SELCI followed by a second-order BWEN perturbation treatment for the remaining configurations [20].
- **d)** RSMP, RSEN, BWEN: Second-order perturbation treatment for all configurations $[21]$ (RS Rayleigh-Schrödinger, BW Brillouin-Wigner, MP Møller-Plesset [22], EN Epstein-Nesbet [23, 24]).

Calculations for singlet, doublet, and triplet states are feasible based on closed-shell SCF molecular orbitals [20].

For any combination of these options, the geometry of a stable molecule can be optimized by a modified DFP method [3, 4, 25]; the search routines are essentially the same as in the available MNDO program [26]. The structures of transition states can be located by a modified nonlinear least-squares method [9, 25].

Rather than cover all combinations of options, the following chapters will focus on the DFP optimization of ground-state singlets, using MNDO for the SCF procedure and defining the SCF determinant to be the only main configuration.

3. Optimization Technique

The DFP method [3, 4] is an iterative procedure which corrects the current set of geometrical variables x^k (*kth* iteration) according to:

$$
x^{k+1} = x^k - \alpha^k H^k g^k \tag{1}
$$

where g^k is the gradient vector evaluated at $x^k \cdot H^k$ is a positive-definite matrix which approximates the inverse Hessian matrix and is updated in each iteration by the standard formula [2-4]. α^k is a scalar which is determined in a line search such that the energy $E(x^{k+1})$ is a minimum along the search direction vector $-H^{k}g^{k}$. Starting from an initial point x^0 , with the gradient g^0 and an initial estimate H^0

obtained from a second gradient calculation at $x^0 + \Delta x^0$ [25], the optimization proceeds completely automatically until the convergence criteria chosen are fulfilled.

The DFP method requires the evaluation of the gradient at least once for each iteration. For variationally optimized SCF wave-functions (including MCSCF), the gradient can be calculated analytically [8, 11, 14, 15] or by finite difference keeping the density matrix constant and recalculating only the relevant integrals [9, 12, 13]. With semiempirical MO methods, both approaches yield the gradient very efficiently; the computation time is usually less than a tenth of that for an SCF calculation. Including electron correlation explicitly by one of the methods listed, the energy is no longer stationary with respect to the LCAO coefficients, and each component g_t^k of the gradient vector must be computed by a central difference formula:

$$
g_i^k = \frac{E(\mathbf{x}^k + d_i \mathbf{e}_i) - E(\mathbf{x}^k - d_i \mathbf{e}_i)}{2d_i} \approx \frac{\partial E(\mathbf{x}^k)}{\partial x_i^k}
$$
(2)

where d_i is an increment (see below), and e_i a vector with its *i*th component unity and the other components zero. For N_n variables $2N_n$ full calculations are necessary to obtain the gradient.

In the original DFP method [3, 4], a cubic interpolation formula is used in the line search to determine α^k in Eq. (1) which requires N_{lc} evaluations of the energy and the gradient (average $\overline{N}_{1c} \approx 2{\text -}3$). This procedure is feasible in conjunction with SCF wavefunctions since the gradient is obtained at little cost. For correlated wavefunctions, it is advantageous, however, to carry out a parabolic line search with N_{lp} evaluations of the energy only (average $\bar{N}_{lp} \approx 5$).

The overall computational effort for a geometry optimization with N_i DFP iterations is estimated in the following way: For SCF wavefunctions, $2 + N_i \overline{N}_{te} =$ M_s evaluations of the energy and the gradient are needed, each of them with an average computation time τ_s . For correlated wavefunctions, there are $2(1 + 2N_v)$ $+ N_i(\overline{N}_{1p} + 1 + 2N_v) = M_c$ evaluations of the energy, with an average computation time τ_c . The ratio of total computation times is then approximated by:

$$
\frac{t_c}{t_s} \approx \frac{M_c \tau_c}{M_s \tau_s} \approx \left[N_v \frac{4 + 2N_i}{2 + N_{lc} N_i} + \frac{2 + \overline{N}_{lp} N_i}{2 + \overline{N}_{lc} N_i} \right] \frac{\tau_c}{\tau_s} \approx N_v \frac{\tau_c}{\tau_s}
$$
\n(3)

As indicated, the term in the square brackets is usually of the order N_v , for typical optimization runs. The computational effort for geometry optimizations with inclusion of electron correlation thus increases linearly with the number of variables, contrary to the SCF case.

Table 1 contains some actual computation times for ethane. The increase in computation time for BWEN compared to SCF is mostly due to the higher number M of function evaluations. For the other correlation treatments, there is the additional effect that the time for a single function evaluation is much greater than in the SCF case. As a consequence, the CI optimization of ethane is more expensive than the SCF optimization by about two orders of magnitude (see Table 1). For

larger molecules, geometry optimizations with inclusion of electron correlation thus seem to be feasible only when using a simple perturbation treatment such as BWEN.

Having discussed the basic optimization technique and the computational effort involved, we shall now turn to problems of numerical accuracy.

The main convergence criterion for DFP optimizations requires all components of the gradient vector to be smaller in absolute value than 1 kcal mol⁻¹ \AA ⁻¹ for bond lengths, and 1 kcal mol⁻¹ rad⁻¹ for angles. With these choices, the bond lengths and angles are optimized to an accuracy of at least $\pm 0.001~\text{\AA}$ and $\pm 0.1^{\circ}$, respectively, unless the potential surface is extremely fiat. Table 2 lists the results of MNDO-CI optimizations for formaldehyde starting from five different initial geometries; in this case, the optimized bond lengths are reproducible to ± 0.0001 Å, and the optimized bond angles to $\pm 0.01^{\circ}$.

| Run | Starting geometry ^b | | | Optimized geometry | | | |
|------------------|--------------------------------|---------------|-----------------------------------|-----------------------|-----------------------------|-----------------------------|--------------------------|
| | $R_{\rm co}(\rm A)$ | $R_{CH}(\AA)$ | $\theta_{\text{\tiny HCO}}$ (deg) | $R_{\rm co}(\rm \AA)$ | $R_{\text{CH}}(\text{\AA})$ | θ_{HCO} (deg) | ΔH^e (kcal/mole) |
| \boldsymbol{A} | 1.208 | 1.116 | 121.75 | 1.2284 | 1.1125 | 123.62 | -50.4078 |
| B | 1.300 | 1.116 | 121.75 | 1.2285 | 1.1125 | 123.60 | -50.4078 |
| \mathcal{C} | 1.208 | 1.000 | 110.00 | 1.2284 | 1.1125 | 123.60 | -50.4078 |
| D | 1.150 | 1.150 | 126.00 | 1.2285 | 1.1125 | 123.60 | $-50,4078$ |
| E | 1.216 | 1.106 | 123.52 | 1.2283 | 1.1124 | 123.61 | $-50,4078$ |

Table 2. MNDO CI geometry optimizations for formaldehyde^a

^a The computation times for the five runs differ by a factor of 1.8 at most. The number of function evaluations varies between 37 (E) and 66 (C).

b Starting geometries: A experimental, *B-D* arbitrary, E SCF-optimized.

c Final heat of formation.

In the gradient calculation, see Eq. (2), the standard values for the increments d_i are 0.001 Å for bond lengths, 0.2° for bond angles, and 0.5° for dihedral angles. A gradient of 1 kcal mol⁻¹ Å⁻¹ then corresponds to an energy difference of 10^{-4} eV in the numerator of Eq. (2). In order to obtain meaningful gradients close to the

minimum by finite difference, the energy must be computed with an accuracy which is appreciably better than 10^{-4} eV. Therefore, as the gradient norm decreases in the course of the optimization, the SCF convergence limit is reset from 10^{-4} eV to 10^{-5} eV, and finally to 10^{-6} eV.

Additional problems can arise from the treatment of correlation (see Chapter 2): The SELCI approach is based on the concept that all important configurations should be selected and included in the configuration interaction. For a balanced description of the potential surface, a selection should be carried out at each point. If the space of selected configurations is different for different points, the surface becomes discontinuous, and the automatic optimization procedures described are no longer applicable. In principle, these considerations hold true analogously for the PERTCI treatment.

It is therefore not surprising that straightforward SELCI optimizations often fail to converge due to the selection of different configurations. On the other hand, PERTCI optimizations are usually successful although there are cases when they do not converge. This different behavior of SELCI and PERTCI is related to the fact that, for numerical purposes, the potential surface needs to be continuous only with an accuracy of the order of 10^{-4} - 10^{-5} eV. In PERTCI, a different selection just changes the partitioning of configurations into selected and remaining configurations [20], with very small effects on the energy, whereas in SELCI it changes the dimensionality of the configuration space which influences the energy rather strongly. Hence, PERTCI optimizations can still be numerically successful for slightly varying selections, contrary to SELCI optimizations.

In order to cope with these problems, two options have been added to our program which allows us to keep the space of selected configurations constant either during the whole optimization or during each gradient calculation. It is recommended to use the first option with SELCI, and the second option with PERTCI, which in our test cases removed the problems encountered. When working with a constant space of selected configurations throughout the whole optimization, the initial selection should include all configurations which are relevant on any part of the potential surface searched during the optimization [27-29].

As for the selection procedure, it should also be noted that the two linearly independent singlets generated in the four-electron problem should both be selected if one of them meets the selection criteria. Otherwise the ambiguity in the definition of these singlets [21] causes optimization problems.

The application of EN perturbation theory to systems with degenerate orbitals presents another problem since the energy obtained is not invariant to unitary transformations among the degenerate orbitals [21,30]. In MNDO-BWEN calculations for a set of 15 molecules the average magnitude of the energy changes due to such transformations turns out to be of the order $5 \cdot 10^{-3}$ eV. This is far above the limit of 10^{-4} - 10^{-5} eV needed for numerical stability of the optimization. On the other hand, it is small enough to introduce the following convention: In an EN perturbation treatment, any degenerate MOs are transformed such that they

conform to the symmetry of a point group with nondegenerate irreducible representations only (e.g. $D_{\infty h} \to D_{2h}$, $C_{3v} \to C_s$, etc.). This convention leads to a unique definition of the potential surface, and removes the optimization problems encountered.

4. Results and Discussion

The molecular geometries optimized with explicit inclusion of electron correlation are expected to differ from the optimized SCF geometries. The present chapter deals with these changes and their dependence on the specific correlation treatment used (see Chapt. 2). SELCI, PERTCI and various second-order perturbation approaches (e.g. BWEN) can be regarded as an approximation to a CI treatment including all singly and doubly excited configurations generated. It is therefore particularly interesting to find out how well these approaches can reproduce the CI results.

Fig. 1. MNDO energy decrease ΔE of ethylene as function of the threshold parameter T. The number of selected configurations is given in the top row (total number 125)

Fig. 2. MNDO energy decrease ΔE of propane as function of the threshold parameter T. The number of selected configurations is given in the top row (total number 1365)

Figures 1 and 2 show the MNDO correlation energy of ethylene and propane at their experimental geometries, as a function of the threshold parameter T used in the selection procedure [20]. A configuration is selected if its interaction with the

main configuration, as estimated by second-order BWEN perturbation theory, is larger than T. The case $T = 0$ corresponds to a selection of all configurations (CI). and the case $T = \infty$ to the selection of no configuration (SCF result for SELCI curve, BWEN result for PERTCI curve). It is obvious from the figures that PERTCI reproduces the CI results very closely, whereas SELCI deviates appreciably for larger T values.

Fig. 3. Optimized bond lengths R_{cc} and R_{CH} of ethylene as functions of the threshold parameter T. The number of selected configurations is given in the top row (total number 125).

Figure 3 shows a similar plot for the optimized bond lengths of ethylene (analogous results are obtained for formaldehyde). Again, PERTCI turns out to be superior to SELCI since the PERTCI curve is much smoother and closer to the CI result. The PERTCI bond lengths are rather insensitive to the actual value of the threshold parameters; $T = 0.002$ eV seems to be a good choice both for SELCI and PERTCI. For larger T values, the SELCI results are inferior even to BWEN (see Fig. 3).

Table 3 lists the SCF optimized geometries (MNDO) for a set of fourteen molecules, along with the changes due to explicit inclusion of electron correlation. First of all, it is gratifying that all correlation treatments studied lead to geometry changes in the same direction and of similar magnitude. The PERTCI results match the CI ones almost perfectly, while SELCI shows occasional deviations. As for the perturbation treatments, the geometry changes due to correlation are usually underestimated by RSMP, and overestimated by RSEN. BWEN reproduces the CI values quite well, and is the best of the perturbation treatments in this respect. The geometry changes in Table 3 follow a typical pattern. Single bonds are usually lengthened by about 0.005 Å, and multiple bonds by about 0.020 Å (triple bonds somewhat more than double bonds). Bond angles at carbon are affected only slightly whereas bond angles at nitrogen and oxygen are usually reduced by about 0.5° leading to more compact structures.

The geometry changes insteam the last six columns are positive (negative) it inclusion of correlation increases the SCF value.
the SCF value.
P Bond lengths A—B in Å, bond angles ABC and dihedral angles ϕ in degree.
9 the SCF value.

Bond lengths A —B in A, bond angles ABC and dihedral angles ϕ in degree. $T = 0.002$ eV. Note that for higher T values, the SELCI results deviate more strongly from the CI results. Drastic geometry changes are only found for hydrogen peroxide (except with SELCI, see Table 3). The MNDO SCF minimum corresponds to the trans-conformation, with a dihedral angle of 180°. Upon inclusion of electron correlation, the minimum occurs at a C_2 structure with a dihedral angle of 132.6° (CI), in much better agreement with the experimental value of 119.1° [31]. This improvement should, however, be regarded as partly fortuitous since electron correlation favors the C_2 structure just by 0.15 kcal/mol relative to the trans-structure. This is sufficient to shift the minimum since the MNDO SCF potential curve for the internal rotation of hydrogen peroxide is extremely flat between 130° and 180° . On the other hand, the experimental trans-barrier is 1.I kcal/mol [32]; the calculated MNDO CI trans barrier of 0.1 kcal/mol indicates that only a small part of the barrier is due to correlation effects. This agrees with *ab initio* SCF calculations [33] which find almost no barrier with an sp-basis set, but a barrier close to the experimental one when adding polarization functions to the basis set. Similarly, another recent *ab initio* study [34] concludes that electron correlation does not affect the calculated barrier heights significantly.

A comparison with experimental data is not included in Table 3 since the MNDO method has been parametrized to reproduce experimental geometries at the SCF level [17]. An improved agreement between calculated and observed structures can therefore be expected only if the method is reparametrized using an explicit correlation treatment. Note that a parametrization of this kind has recently been used to develop the LNDO/S method [35], for the calculation of excitation and ionization energies.

5. Conclusions

Automatic geometry optimizations with explicit inclusion of electron correlation can be carried out by the DFP method or a nonlinear least-squares algorithm requiring only relatively small modifications of existing programs. The main drawback of such calculations is the high computational effort involved, compared to SCF geometry optimizations. Therefore, approximate treatments of electron correlation are needed, especially for larger molecules.

In conjunction with the semiempirical MNDO method, the optimized geometries obtained by CI are reproduced very well by PERTCI, and satisfactorily by BWEN, whereas SELCI is clearly inferior to PERTCI and much more expensive than BWEN. Hence, in geometry optimizations, the correlation treatments CI, PERTCI, and BWEN are recommended. For larger systems, only BWEN is expected to be computationally feasible.

Finally it should be pointed out that the optimization techniques described can, in principle, also be applied with correlated *ab initio* wavefunctions. Due to the extensive computational effort in the *ab initio* case, however, gradient-based geometry optimizations with explicit inclusion of electron correlation will probably be restricted to semiempirical calculations in the near future.

Acknowledgement. This work was supported by the Deutsche Forschungsgemeinschaft and the Fond der Chemischen Industrie. The calculations were carried out on the TR 440 computer of the Rechenzentrum der Universität Marburg. One of us (W.T.) thanks the Verband der Chemischen Industrie for a Liebigstipendium.

References

- 1. Dixon, R. N.: Nonlinear optimization. London: Engl. Univ. Press 1972.
- 2. Fletcher, R.: Comp. Phys. Commun. 3, 159 (1972)
- **3. Davidson, W. C.:** Argonne Natl. Lab. Report ANL-5990 Rev. (1959)
- 4. Fletcher, R., Powell, M. J. D.: Comp. J. 6, 163 (1963)
- 5. Murtagh, B. A., Sargent, R. W. H.: Comp. J. 13, 185 (1972)
- 6. McIver, J. W., Komornicki, A.: J. Am. Chem. Soc. 94, 2625 (1972)
- 7. Powell, M. J. D.: Comp. J. 7, 303 (1965)
- 8. McIver, J. W., Komornicki, A.: Chem. Phys. Letters 10, 303 (1971)
- 9. Dewar, M. J. S., Kollmar, H. W., Lo, D. H., Metiu, H., Student, P. J., Weiner, P. K.: Private communication
- 10. Bloemer, W. L., Brunet, B. L.: J. Chem. Phys. 58, 3735 (1973)
- 11. Pulay, P.: Mol. Phys. 17, 197 (1969); 18, 473 (1970)
- 12. Poppinger, D.: Chem. Phys. Letters 34, 332 (1975); 35, 550 (1975)
- 13. Collins, J. B., Schleyer, P. von R., Binkley, J. S., Pople, J. A. : J. Chem. Phys. 64, 5142 (1976)
- 14. Komornicki, A., Ishida, K., Morokuma, K., Ditchfield, R., Conrad, M.: Chem. Phys. Letters 45, 595 (1977)
- 15. Pulay, P. in: Modern theoretical chemistry, Schaefer, H. F., III, ed., Vol. 4, p. 153. New York: Plenum Press 1977
- 16. Dewar, M. J. S., Doubleday, C. E.: J. Am. Chem. Soc. 100, 4935 (1978)
- 17. Dewar, M. J. S., Thiel, W.: J. Am. Chem. Soc. 99, 4899 (1977)
- 18. Bingham, R. C., Dewar, M. J. S., Lo, D. H.: J. Am. Chem. Soc. 97, 1285 (1975)
- 19. Pople, J. A., Segal, G. A.: J. Chem. Phys. 44, 3289 (1966)
- 20. Hase, H.-L., Lauer, G., Schulte, K.-W., Schweig, A.: Theoret. Chim. Acta (Berl.) 48, 47 (1978)
- 21. Ostlund, N. S., Bowen, M. F.: Theoret. Chim. Acta (Berl.) 40, 175 (1976)
- 22. Moiler, C., Plesset, M. S.: Phys. Rev. 46, 618 (1934)
- 23. Epstein, P. S.: Phys. Rev. 28, 695 (1926)
- 24. Nesbet, R. K.: Proc. Roy. Soc. A230, 312, 322 (1955)
- 25. Weiner, P. K.: Ph.D. Thesis, University of Texas, Austin, Texas, 1975
- 26. Thiel, W.: Program No. 353, Quantum Chemistry Program Exchange, University of Indiana, Bloomington, Indiana
- 27. Buenker, R. J., Peyerimhoff, S. D.: Theoret. Chim. Acta (Berl.) 35, 33 (1974)
- 28. Raffenetti, R. C., Hsu, K., Shavitt, I.: Theoret. Chim. Acta (Berl.) 45, 33 (1977)
- 29. Raseev, G.: Chem. Phys. Letters 47, 36 (1977)
- 30. Pople, J. A., Binkley, J. S., Seeger, R.: Intern. J. Quantum Chem. Symp. 10, 1 (1976)
- 31. Khachkuruzov, G. A., Przhevalskii, I. N.: Opt. Spectry. 36, 172 (1974)
- 32. Hunt, R. H., Leacock, R. A., Petters, C. W., Hecht, K. T. : J. Chem. Phys. 42, 1931 (1965)
- 33. Dunning, T. H., Winter, N. W.: J. Chem. Phys. 63, 1847 (1975)
- 34. Rodwell, W. R., Carlsen, N. R., Radom, L.: Chem. Phys. 31, 177 (1978)
- 35. Lauer, G., Schulte, K.-W., Schweig, A.: J. Am. Chem. Soc. 100, 4925 (1978)

Received December 18, 1978~February 28, 1979