Theoret. Chim. Acta (Berl.) 42, 67–75 (1976)

THEORETICA CHIMICA ACTA

Application of SCF Perturbation Theory to Molecular Calculations

David P. Santry*

Lehrstuhl B für Anorganische Chemie, Technische Universität, Pockelstr. 4,
D-3300 Braunschweig, Federal Republic of Germany

A method for solving Roothaan's molecular orbital equations by means of SCF perturbation theory is presented. An estimate of the accuracy of the third order expansion is made for the CNDO/2 approximation from a comparison of the results from direct calculations. It is found that the third order theory is sufficiently accurate for quantitative studies.

Key words: SCF perturbation theory

1. Introduction

A method is presented here for the calculation of the density matrix and electronic energy of a molecule in terms of a localized basis set of bond and lone-pair orbitals. All electrons in a closed-shell molecule are assigned in pairs to molecular orbitals localized in either bond or lone-pair directions. Interactions between these orbitals are then introduced by means of molecular-orbital based perturbation theory.

This approach has already been applied by Coulson, Redei and Stocker [1, 2] to the theoretical study of the electronic structures of certain semi-conductors, and by several other workers [3-5] to the investigation of hydrocarbon molecules.

The above studies are all based on the Hückel approximate molecular orbital method and therefore avoid the problem of making the perturbation self consistent. The method is extended in the present treatment by calculating all inter bond and lone-pair interactions by means of SCF (self consistent field) perturbation theory [6]. The underlying theory is formally very similar to that developed for the calculation of inter-molecular interactions within a molecular aggregate [7, 8]. The desired formalism may be obtained from the aggregate theory by the substitution of localized bond and lone-pair orbitals for the molecular orbitals localized on the molecules of the aggregate.

The relative ease of solution of the Hückel perturbation equations was an attractive feature of the perturbation approach before the wide-spread availability of computers. The extension of the perturbation theory to the SCF level complicates the solution of the equations to the point where some computational aid is necessary. Thus, the only reason remaining for the adoption of the perturbation approach, in the present application, is the added insight into the electronic structure of molecules it affords. This may be illustrated, for example, by a discussion of the effect of electron delocalization on the long-range nuclear spin coupling constants of hydrocarbon molecules [9].

---

* Present address: Department of Chemistry, McMaster University, Hamilton, Ontario, Canada.

2. Theory

The objective is the solution of Roothaans [10] equation,

$$FC = SCE, \tag{1}$$

for a molecule through the calculation of the density matrix,

$$P_{\mu\nu} = 2 \sum_{i}^{\mathrm{OCC}} C_{\mu i} C_{\nu i}, \tag{2}$$

by means of SCF perturbation theory. All of the symbols in the above equations have their usual meanings. The valence basis set chosen for the solution of Eq. (7) consists of a 1s orbital on each proton and four hybrid orbitals on all other atoms. These hybrid orbitals are directed along bond or lone-pair directions. Furthermore, it will be assumed that the basis set is ordered so that orbitals associated with a given bond are adjacent; all lone-pair orbitals are listed together at the end.

The Fock, overlap and density matrices are naturally partitioned into sub-matrices by the ordering of the basis set in this manner. Assuming there are two orbitals associated with each bond and one with each lone-pair, there will be 2 x 2 intra and inter-bond, 2 x 1 inter-bond lone-pair and 1 x 1 intra lone-pair submatrices. Under the CNDO approximation [11], for example, the intrabond Fock sub-matrix, ${}^{RR}F$, for the Rth bond is given by

$${}^{RR}F_{mm} = U_{mm} + \tfrac{1}{2}(1 - {}^{RR}P_{mm})\gamma_{\alpha\alpha} + \sum_{\beta} (Q_\beta - Z_\beta)\gamma_{\alpha\beta} \tag{3}$$

where, assuming two hybrids per bond, $m = 1$ or 2, and the mth orbital is centered on the αth atom. ${}^{RR}P_{mm}$ is the corresponding element from the Rth intra-bond density sub-matrix. $\gamma_{\alpha\alpha}$ is a CNDO average Coulomb integral, Z_β the effective nuclear charge of the βth atom and Q_β the total electron density associated with the same atom.

$$Q_\beta = \sum_{T}{}' \sum_{n}{}' \, {}^{TT}P_{nn} \tag{4}$$

The summations over T and n are restricted to include only those bonds and orbitals, respectively, associated with the βth atom.

The off-diagonal elements of ${}^{RR}F$ are given by

$${}^{RR}F_{12} = {}^{RR}H_{12} - \tfrac{1}{2}{}^{RR}P_{12}\gamma_{\alpha\beta} \tag{5}$$

where the Rth bond connects atoms α and β. ${}^{RR}H_{12}$ is given, under the CNDO/2 approximation [11], by the product of the overlap integral between the bond orbitals and the appropriate scaling factor.

The Fock sub-matrix, ${}^{RS}F$, between bonds R and S is given by

$${}^{RS}F_{mn} = {}^{RS}H_{mn} - \tfrac{1}{2}{}^{RS}P_{mn}\gamma_{\alpha\beta}, \tag{6}$$

where $m, n = 1$ or 2 and orbitals m and n are centered on atoms α and β, respectively. Various possible contributions to ${}^{RS}H$ are discussed and illustrated in Ref. 5. As the overlap matrix is assumed to be diagonal under the CNDO/2 approximation it does not require discussion.

3. Choice of the Zero Order Equations

It is important to realize that a certain degree of freedom exists both in the selection of elements and from F to be included at the zero order and in the incorporation of the remaining elements into the perturbation scheme. All that is required, for the present purpose, is that all zero order off-diagonal sub-matrices be zero.

${}^{RR}F^{(0)}$ could be defined as the operator for a hypothetically isolated bond or lone-pair through the neglect of all contributions, electrostatic under the CNDO/2 approximation, from electrons in other bonds. These could then be re-introduced as a contribution to the first order Fock sub-matrix, ${}^{RR}F^{(1)}$. Alternatively, the inter-bond electrostatic contribution can be retained in ${}^{RR}F^{(0)}$, so that the zero order equation treats the molecule as a collection of mutually polarizing but otherwise independent bonds. There is little to choose between these two alternative starting points in terms of the computational effort required to solve the resultant equations. The advantage of the first choice is that all of the intra-bond and intra lone-pair Fock sub-matrices are independent, so that all zero order equations can be solved separately. The inclusion of the neglected inter bond contributions to the intra-bond Fock submatrices requires the solution of coupled first order equations. The situation is reversed in the case of the second definition. The inclusion of the inter-bond terms at the zero order requires that the resultant equations be solved as a group of coupled equations. On the other hand, under this definition ${}^{RR}F^{(1)}$ is zero and there are no first order equations to be solved. Also, there appears to be little to choose between the two definitions from the point of view of accuracy. Calculations, for ethyl amine, based on the two different choices of perturbation yielded virtually identical results after the second order.

All of the calculations reported here are based on the second definition of ${}^{RR}F^{(0)}$. The zero order Fock submatrices are given by

$$ {}^{RR}F^{(0)}_{mm} = U_{mm} + \tfrac{1}{2}(1 - {}^{RR}P^{(0)}_{mm})\gamma_{\alpha\alpha} + \sum_{\beta}^{\text{all atoms}} (Q^{(0)}_{\beta} - Z_{\beta})\gamma_{\alpha\beta} \tag{7} $$

and

$$ {}^{RR}F^{(0)}_{12} = {}^{RR}H_{12} - \tfrac{1}{2}{}^{RR}P^{(0)}_{12}\gamma_{\alpha\beta} \tag{8} $$

As mentioned above, all of the zero order intra-bond equations have to be solved as coupled equations since each depends on the density sub-matrix, ${}^{TT}P^{(0)}$, through $Q^{(0)}_{\beta}$, of them all. All elements of H so far neglected are included as a first order contribution.

$$ {}^{RR}F^{(1)} = 0 \tag{9} $$

This follows, only under the CNDO approximation, because ${}^{RR}H^{(1)}$ is zero, and the inter-bond contributions to ${}^{RR}F$ have been included at the zero order. The first order inter-bond and lone pair Fock submatrices are given by

$$ {}^{RS}F^{(1)}_{mn} = {}^{RS}H_{mn} - \tfrac{1}{2}{}^{RS}P^{(1)}_{mn}\gamma_{\alpha\beta} \tag{10} $$

For orders, x, greater than 1

$$ {}^{RR}F^{(x)}_{mm} = -\tfrac{1}{2}{}^{RR}P^{(x)}_{mm}\gamma_{\alpha\alpha} + \sum_{\beta} Q^{(x)}_{\beta}\gamma_{\alpha\beta} \tag{11} $$

$$ {}^{RS}F^{(x)}_{mn} = -\tfrac{1}{2}{}^{RS}P^{(x)}_{mn}\gamma_{\alpha\beta} \tag{12} $$

where R can equal S if m does not equal n.

The second and higher orders in the intra-bond and lone pair submatrices are non-zero because there are constant, non iterative, contributions to $P^{(x)}$ arising from products of pairs of lower order contributions to the occupied molecular orbitals [7, 8].

The SCF perturbation equations are solved for the density sub-matrices by the method outlined in Ref. [7]. The full overlap dependent equations may be solved by the method outlined in Ref. [8]. Once the density submatrices have all been calculated, the molecular electronic energy, W, may be calculated from the equation

$$W = \frac{1}{2} \sum_{\mu} \sum_{\nu} P_{\mu\nu}(H_{\mu\nu} + F_{\mu\nu}) \qquad (13)$$

The present approach does not provide a means for calculating either the molecular orbitals or the orbital energies. These can be calculated [12], under certain circumstances, on completion of the main calculation.

Although the above discussion is limited to the use of individual bond orbitals as zero order basis functions, there is no reason why, with obvious extensions to the theory, molecular orbitals associated with molecular fragments, for example, a methyl group, could not be used for the same purpose. The theory in this form may provide a useful means for the study of group interactions within molecules.

4. Application of the Perturbation Method

The perturbation expansion, taken to the third order, has been applied to calculation of the density matrices and molecular energies of propane, ethanol, ethylamine and ethyl fluoride. The results from these calculations are compared here against those from direct calculations to provide an estimate of the accuracy of the third order expansion.

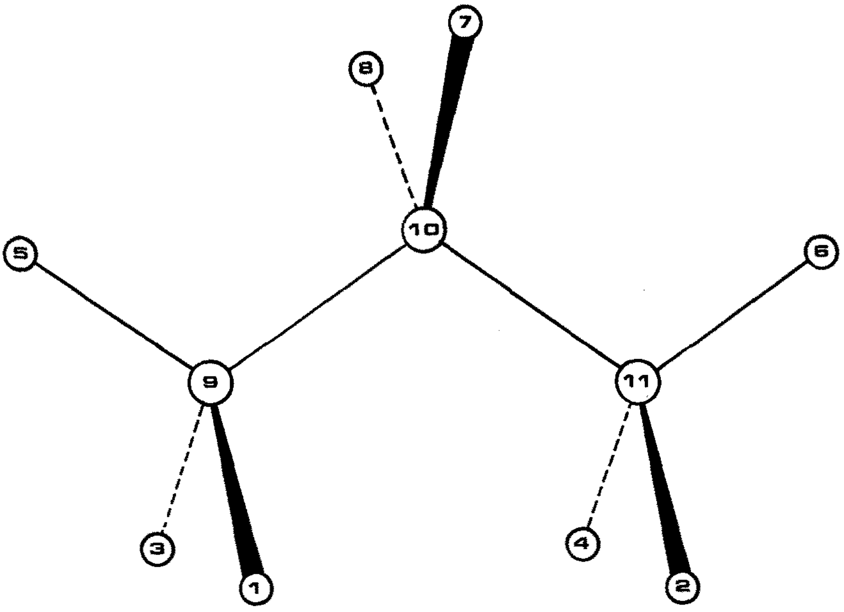

Fig. 1. Molecular conformation and atomic numbering scheme for propane

Table 1. Atomic populations calculated for propane, Fig. 1, directly and by means of SCF perturbation theory. $Q_\beta^{(2)}$, for example, is the second order contribution to the total electron population of atom β. $Q^{(1)}$ is zero under the CNDO approximation and the chosen perturbation

Atom	1	5	7	9	10
$Q^{(0)}$	0.98309	0.983775	0.98527	4.051091	4.02737
$Q^{(2)}$	0.01241	0.01443	0.01634	−0.03669	−0.03782
$Q^{(3)}$	0.00137	0.00523	0.00499	−0.00349	−0.01893
Total	0.99696	1.00344	1.00661	4.01091	3.97063
Direct	0.99736	1.00406	1.00537	4.00804	3.9756

The molecular configuration and atomic numbering scheme selected for propane for the purpose of these calculations is shown in Fig. 1. All C–H and C–C bond lengths are 1.091 and 1.536 Å, respectively, and all bond angles are tetrahedral. The basis set for the perturbation calculation consists of 1*s* orbitals on each proton and four sp^3 hybrids on each carbon. The conventional CNDO basis set was retained for the direct calculations.

The total electron populations for all independent propane atoms are listed by order in Table 1. These data show the perturbation series to be convergent but the level of convergence not to be uniform over the various atoms. Comparison of the atomic populations from the perturbation and direct calculations, also listed in Table 1, shows the error to be the least for atom 1 and the greatest for the bridging carbon, atom 10. The reason for this variation in the convergence of the perturbation expansion can be understood in terms of a discussion given in Ref. [13], and brings to light an important feature of the perturbation approach.

It has been shown [13, 14] that the transfer of electron density between the zero order localized units, here bond and lone-pair orbitals, is a second and higher order effect. Even at the second order the transferred charge is not calculated to a self consistent level, since it originates from the constant [7, 8], non iterative, component of $P^{(2)}$. The first step towards making this feature of the charge distribution self consistent is taken at the third order in the calculation of $P^{(3)}$. Thus even if the elements of the perturbative *Hamiltonian* matrix are roughly equal, the level of convergence of the calculated charge densities may vary from atom to atom, simply because their routes to self consistency are different. Furthermore, it is not the strength of the perturbation alone that governs to which order the perturbation expansion should be taken. Clearly, for quantitatively acceptable results, the expansion must be taken to the second order, even for relatively weak perturbations, otherwise contributions from inter bond charge transfers will be excluded.

The total energy of propane calculated both directly and by the perturbation expansion is given in Table 2. The difference between the two calculations is 0.00453 a.u. As this is appreciably larger than the theoretical barrier to the rotation of a methyl group, it would appear that the perturbation method is unsuitable for theoretical studies on molecular conformation. However, more extensive calculations showed this difference remains constant, to within 0.00002 a.u., throughout the entire rotation of a methyl group in propane. Thus the potential curves, Fig. 2, for the rotation of a methyl group from the direct and perturbational calculations are almost indistinguishable, except that from the latter calculation is displaced by 0.00453 a.u. Similar results have been reported from SCF perturbation calculations on hydrogen bonded dimers.

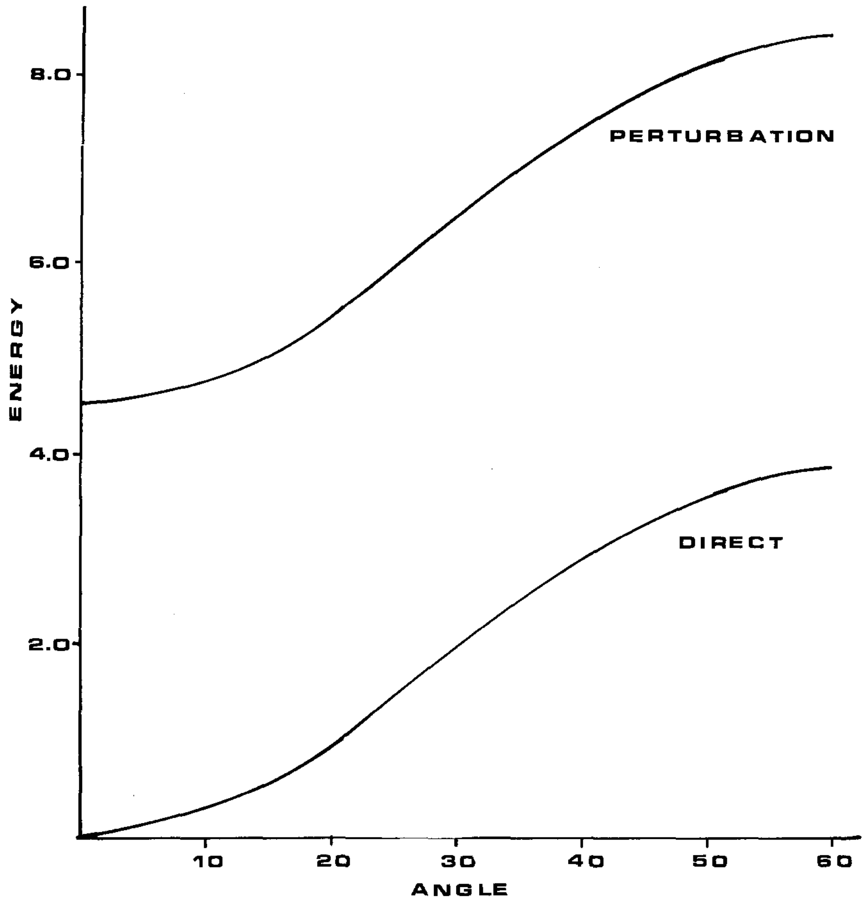

Fig. 2. Potential curves for the rotation of a methyl group in propane

5. Results for Ethynol, Ethylamine and Ethylfluoride

Certain of the inter bond interactions will get progressively stronger along the above series of molecules. The matrix element between two hybrid orbitals on the same atom, but different bonds, is directly proportional to the difference between the 2*s* and 2*p* orbital energies of the atom in question [5]. Since this difference increases along the series C, N, O and F, these will be corresponding increases in the strength of the perturbation for molecules containing these atoms. It is of interest to determine whether these increases are sufficient to appreciably affect the convergence of the perturbation expansion.

Table 2. Total molecular energies calculated directly and by means of SCF perturbation theory

Molecule	Et-CH_3	Et-NH_2	Et-OH	Et-F
Perturbation	−27.49125	−31.24343	−37.24451	−45.78253
Direct	−27.49577	−31.24814	−37.25048	−45.78821
Difference	0.00453	0.00471	0.00597	0.00568

The structural data necessary for calculations on $C_2H_5NH_2$, C_2H_5OH and C_2H_5F were generated from those of propane, Fig. 1, by the removal of the appropriate protons and the substitution of N, O and F, respectively, for C_9; the atomic numbering scheme of Fig. 1 was maintained. The sp^3 basis for the heavy atoms was retained. The lone pair orbitals of NH_2 and OH were directed along the tetrahedral directions of the absent hydrogens. In the case of F, the three lone pair orbitals were rotated out of these directions by 60°.

The molecular energies and total atomic populations, calculated both directly and by the perturbation method, are listed in Tables 2 to 5. The total energy calculated by the perturbation expansion shows only a slight deterioration in the level of convergence along the series. Surprisingly, the error reaches a maximum at Et-OH and a minimum at Et-NH_2. There is a significant increase in the perturbative error in the atomic populations along the series of molecules, but the perturbation expansion is acceptably convergent for all molecules. The atomic population of F is particularly interesting in this regard. Although the second and third order contributions, $Q_F^{(2)}$ and $Q_F^{(3)}$, are of the same sign and of comparable magnitude, the final error is an order of magnitude smaller than $Q_F^{(3)}$.

Table 3. Atomic populations calculated for ethylamine directly and by means of SCF perturbation theory (R_{NC} = 1.474 and R_{NH} = 1.0 Å)

Atom	1	2	6	7	9	10	11
$Q^{(0)}$	0.9076	0.98161	0.98252	0.98742	5.24100	3.96607	4.05824
$Q^{(2)}$	0.01389	0.01180	0.01368	0.01805	−0.02113	−0.05350	−0.02654
$Q^{(3)}$	−0.00169	0.00192	0.00219	0.00829	−0.00289	−0.01346	−0.00289
Total	0.91926	0.99532	0.99840	1.01377	5.21698	3.89912	4.02882
Direct	0.92059	0.99531	0.99982	1.01264	5.21630	3.89880	4.02802

Table 4. Atomic populations calculated for ethanol directly and by means of SCF perturbation theory (R_{CO} = 1.428 Å and R_{OH} = 0.97)

Atom	2	5	6	7	9	10	11
$Q^{(0)}$	0.97471	0.83241	0.98381	0.98932	6.28821	3.90924	4.05827
$Q^{(2)}$	0.00911	0.03913	0.00905	0.02296	−0.03683	−0.04938	−0.02612
$Q^{(3)}$	0.00223	−0.01592	−0.00344	0.01322	0.00772	−0.01956	0.00032
Total	0.98605	0.85562	0.98941	1.02549	6.25910	3.84030	4.03247
Direct	0.98530	0.86530	0.98987	1.02484	6.25451	3.83821	4.03184

Table 5. Atomic populations calculated for fluoroethane directly and by means of SCF perturbation theory (R_{CF} = 1.379 Å)

Atom	2	6	7	9	10	11
$Q^{(0)}$	0.96751	0.97756	0.97677	7.21756	3.84947	4.06685
$Q^{(2)}$	0.00929	0.00798	0.02087	−0.01211	−0.03211	−0.02409
$Q^{(3)}$	0.00178	−0.00645	0.01280	0.00535	−0.03111	0.00305
Total	0.97858	0.97909	1.01044	7.21081	3.78625	4.04582
Direct	0.97797	0.98126	1.01024	7.20862	3.78839	4.04534

6. Variable Hybridization Calculations for CH_3F

The above calculations are all in terms of a basis set which includes sp^3 hybrid orbitals. Direct CNDO calculations are independent of the choice of hybridization since this corresponds to an orthogonal transformation of the atomic orbital basis set. The same would be true of the perturbation calculations if they were taken to a sufficiently high order. However, as the perturbation expansion is truncated after the first few terms, some dependence on hybridization is to be expected. This could, perhaps, be turned to advantage by choosing the hybridization parameters so as to reduce the strength of the perturbation, thereby improving the convergence.

Preliminary calculations were made for CH_3F to investigate the effect of changes in hybridization on the convergence of the perturbation series. The sp^3 hybrids in the C–F bonds were replaced by hybrid orbitals of the type,

$$\text{Cos}\,\lambda X_{2s} + \text{Sin}\,\lambda X_{2p},$$

and the other hybrid orbitals changed to maintain orthogonality and to be consistent with the C_3 axis of symmetry. Calculations were repeated for a range of hybridization parameters, λ, for both C and F. Some of the results from these calculations, together with those obtained directly, are listed in Table 6. These data show that while some improvement can be effected in some matrix elements of P, it is largely at the expense of others; the overall improvement is small. This is largely because the limiting factor here is not the strength of the perturbation so much as the fact that certain routes to self consistency remain closed as long as the perturbation expansion is limited to the third order; the only way to improve the accuracy of the expansion is to take it to a higher order.

7. Summary and Conclusions

Molecular orbital calculations based on SCF perturbation theory have been reported for several molecules. The following conclusions have emerged from the results.

Table 6. Charge densities calculated for fluoromethane directly and by means of SCF perturbation theory with variable hybridization. P_{PC} and P_{PF} are, respectively, the total electron densities in the ϕ orbitals of carbon and fluorine

		Hybridization Parameters λ_C, λ_F		
	Direct	60,60	60,58	63,58
$P_{1S,1S}$	0.9969	0.9963	0.9963	0.9949
$P_{1S,1S'}$	−0.0347	−0.0295	−0.0297	−0.0308
$P_{1S,2Sc}$	0.5178	0.5189	0.5189	0.5175
$P_{1S,2SF}$	0.0022	0.0058	0.0048	0.0024
$P_{2Sc,2Sc}$	1.0402	1.0300	1.0289	1.0325
$P_{2Sc,2SF}$	0.1795	0.1679	0.1668	0.1732
$P_{2SF,2SF}$	1.8411	1.8519	1.8577	1.8573
P_{PC}	2.7771	2.7860	2.7811	2.7791
P_{PF}	5.3509	5.3432	5.3432	5.3464
W (a.u.)	−37.0935	−37.0905	−37.0898	−37.0904

1) The SCF perturbation method is sufficiently accurate when taken to the third order to make it a useful and interesting tool for the quantitative investigation of the electronic structure of molecules.
2) The truncation point in the perturbation series is not governed by the strength perturbation alone. For example, even the weakest perturbation must be taken at least to the second order to allow for inter-bond charge transfer.

References

1. Coulson, C. A., Redei, L. B., Stocker, D.: Proc. Roy. Soc. **A270**, 357 (1962)
2. Redei, L. B.: Proc. Roy. Soc. **A270**, 383 (1962)
3. Yoshizumi, Y.: Trans Faraday Soc. **53**, 125 (1957)
4. Dewar, M. J. S., Pettit, R.: J. Chem. Soc. 1625 (1954)
5. Pople, J. A., Santry, D. P.: Mol. Phys. **7**, 269 (1964)
6. McWeeny, R., Diercksen, G.: J. Chem. Phys. **44**, 3554 (1966)
7. Crowe, R. W., Santry, D. P.: Chem. Phys. **2**, 304 (1973)
8. O'Shea, S. F., Santry, D. P.: Theoret. Chim. Acta. (Berl.) **37**, 1 (1975)
9. Pople, J. A., Santry, D. P.: Mol. Phys. **9**, 311 (1965)
10. Roothaan, C. C. J.: Rev. Mod. Phys. **23**, 69 (1951)
11. Pople, J. A., Beveridge, D.: Approximate molecular orbital theory. New York: McGraw-Hill 1970
12. Middlemiss, K. M., Santry, D. P.: Chem. Phys. **1**, 128 (1973)
13. Middlemiss, K. M., Santry, D. P.: J. Chem. Phys. **61**, 5400 (1975)
14. Santry, D. P., in: Electronic structure of polymers and molecular crystals. J. Andre and J. Ladik, Eds. New York: Plenum Press 1975

Received March 17, 1975/February 23, 1976