Second-Order Perturbation Theory with a CASSCF Reference Function

Kerstin Andersson,* Per-Åke Malmqvist, Björn O. Roos, Andrzej J. Sadlej, and Krzysztof Wolinski[†]

Department of Theoretical Chemistry, Chemical Centre, P.O.B. 124, S-221 00 Lund, Sweden
(Received: November 1, 1989; In Final Form: March 29, 1990)

Second-order perturbation theory based on a CASSCF reference state is derived and implemented. The first-order wave function includes the full space of interacting states. Expressions for the contributions to the second-order energy are obtained in terms of up to four-particle density matrices for the CASSCF reference state. The zeroth-order Hamiltonian reduces to the Møller–Plesset Hamiltonian for a closed-shell reference state. The limit of the implementation is given by the number of active orbitals, which determines the size of the density matrices. It is presently around 13 orbitals. The method is illustrated in a series of calculations on H_2, H_2O, CH_2, and F^-, and the results are compared with corresponding full CI results.

1. Introduction

The simplest way to treat dynamical electron correlation effects in molecular systems is to use the second-order perturbation approach. Second and higher order perturbation theory for Hartree–Fock reference functions has gained today widespread use through program systems like GAUSSIAN.[1] The reliability of this approach for closed-shell systems is by now well established. In many cases quite useful results can be obtained already at the second-order level. However, the perturbation theory in this form is limited to cases where the Hartree–Fock wave function provides a valid starting point. Otherwise, the perturbation expansion converges only slowly or not at all and low-order approximations become rather unreliable.

Near-degeneracy correlation effects can be today effectively treated by using the multiconfigurational SCF method.[2] Normally, dynamical correlation is included through subsequent multireference (MR) CI calculations. Such an approach, however, has severe limitations as regards the size of the systems that can be treated, both with respect to the number of electrons and the basis set dimension. A general perturbation theory for the near-degenerate reference state was formulated a long time ago[3] but its applications in chemistry are developing only slowly.[4] The major problem is in devising a practical implementation of the theory. Such an implementation is most probably also going to suffer from similar bottlenecks with respect to the size of the system as the MR-CI and MR coupled cluster methods. It is therefore natural to investigate the possibility of extending low-order perturbation schemes to the case with a multiconfigurational reference function. In particular, as shown in the present paper, the second-order perturbation theory can without too much difficulty be implemented in an effective way, allowing applications to all systems that can be treated by the MCSCF method. An attempt was made some years ago to implement second-order perturbation theory in conjunction with CASSCF wave functions.[5] At the time it was not possible to include the full interacting space in the first-order wave function, because of the difficulty in computing the three- and four-particle matrices, which occur for matrix elements involving internal and semiinternal excitations. The results obtained in those earlier test applications were also rather disappointing, with only minor or no improvement in computed properties like dissociation energies, geometry parameters, and spectroscopic constants.

Later, Wolinski et al. derived a full second- and third-order treatment for open-shell SCF and small MCSCF reference functions and gave some illustrative results for several small molecules.[6] The results show, however, that their third-order contribution can be sizable, when the reference function is of the GVB type. The experience gained in the present second-order treatment partly supports this conclusion. A CASSCF second-order perturbation theory, based on the Møller–Plesset type partition, has also been presented by McDouall et al.[7] Their

approach is somewhat different, since in contrast to the present formulation it is based on an effective Hamiltonian formalism. The results presented by them, for the singlet–triplet splitting in CH_2 and the distance dependence of the correlation energy in H_2O, are not very encouraging. Not much improvement is obtained compared to the CASSCF results, except, of course, for the total energies.

The present work describes a second-order perturbation approach to dynamical correlation, with a single reference state given by a multiconfiguration CASSCF type wave function. Possible near-degeneracies are thus assumed to have been accounted for at the level of the CASSCF method. Recent developments in methods for computing N-particle density matrices[8] have made it possible to include the full first-order interacting space in such a treatment. The present approach is similar though not equivalent to the method suggested in the earlier work.[5] Thus, the zeroth-order Hamiltonian is constructed by using a Fock-type one-electron operator, which reduces to the Møller–Plesset HF operator for doubly occupied and unoccupied orbitals.[9] The first-order CI space is nonorthogonal and only block diagonal with respect to the zeroth-order Hamiltonian. This, however, does not constitute any serious problem since the separate blocks can be diagonalized without difficulty.

The method will be presented in detail in the next section. Several test calculations have been performed and compared with full CI results obtained with the same basis sets.

2. Theory

2.1. The First-Order Interacting Space. In the present study a multiconfigurational wave function $|0\rangle$, generated from a CASSCF calculation, is used as the zeroth-order wave function in a perturbation approach to the correlation problem. As the theory is formulated, it is valid for *any reference state* constructed as a full CI wave function in some orbital subspace. However,

(1) GAUSSIAN 86; Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Raghavachari, K.; Melius, C. F.; Martin, R. L.; Stewart, J. J. P.; Bobrowicz, F. W.; Rohlfing, C. M.; Kahn, R. L.; DeFrees, D. J.; Seeger, R.; Whiteside, R. A.; Fox, D. J.; Fleuder, E. M.; Pople, J. A. Carnegie-Mellon Quantum Chemistry Publishing Unit: Pittsburgh, PA, 1984.
(2) Three recent review articles by B. O. Roos, R. Shepard, and H.-J. Werner can be found in: *Ab Initio Methods in Quantum Chemistry, Part II; Adv. Chem. Phys.,* **69;** Lawley, K. P., Ed.; Wiley: Chichester, U.K., 1987.
(3) Brandow, B. H. *Rev. Mod. Phys.* **1967,** *39,* 771.
(4) See for example: Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981,** *32,* 359. Banerjee, A.; Simons, J. *J. Chem. Phys.* **1982,** *76,* 4548. Lindgren, I. *J. Phys. B* **1974,** *7,* 2441. Shavitt, I.; Redmon, L. T. *J. Chem. Phys.* **1980,** *73,* 5711. Mukherjee, D. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1986,** *20,* 409. Shepard, M. G.; Freed, K. F. *Int. J. Quantum Chem. Symp.* **1981,** *15,* 21. Kutzelnigg, W. *J. Chem. Phys.* **1984,** *80,* 822.
(5) Roos, B. O.; Linse, P.; Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Phys.* **1981,** *66,* 197.
(6) Wolinski, K.; Sellers, H. L.; Pulay, P. *Chem. Phys. Lett.* **1987,** *140,* 225. Wolinski, K.; Pulay, P. *J. Chem. Phys.* **1989,** *90,* 3647.
(7) McDouall, J. J. W.; Peasly, K.; Robb, M. A. *Chem. Phys. Lett.* **1988,** *148,* 183.
(8) Malmqvist, P.-Å.; Rendell, A.; Roos, B. O. *J. Phys. Chem.,* this issue.
(9) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934,** *46,* 618.

[†]Present address: Department of Chemistry, University of Arkansas, Fayetteville, AK 72701.

we shall use the fact that the orbitals are optimized to make some simplifications in the computational procedure.

The configuration space, in which the wave function is expanded, can be divided into four subspaces: V_0, V_K, V_{SD}, and $V_{TQ..}$, where (1) V_0 is the one-dimensional space spanned by the CAS reference function $|0\rangle$ for the state under consideration; (2) V_K is the space spanned by the orthogonal complement to $|0\rangle$ in the restricted full CI subspace used to generate the CAS wave function; (3) V_{SD} is the space spanned by all single and double replacement states generated from V_0, and (4) $V_{TQ..}$ is the space which contains all higher order excitations not included in 1–3.

It is only the functions in subspace 3 that interact with the reference function via the total Hamiltonian, and the zeroth-order Hamiltonian will be formulated in such a way that only V_{SD} contributes to the expansion of the first order wave function (vide infra).

All functions in V_{SD} can be generated from the functions, $\hat{E}_{pq}\hat{E}_{rs}|0\rangle$, where $\hat{E}_{pq}$ are the normal spin-averaged excitation operators.[5] Not all these functions are needed, however. Functions of V_K are, for example, generated when all four indexes p,q,r,s refer to active orbitals. The functions needed in the expansion of the first-order wave function can be divided into eight groups as follows

internal	$\hat{E}_{ti}\hat{E}_{uv}|0\rangle$	(1a)
	$\hat{E}_{ti}\hat{E}_{uj}|0\rangle$	(1b)
semiinternal	$\hat{E}_{at}\hat{E}_{uv}|0\rangle$	(1c)
	$\hat{E}_{ai}\hat{E}_{tu}|0\rangle, \quad \hat{E}_{ti}\hat{E}_{au}|0\rangle$	(1d)
	$\hat{E}_{ti}\hat{E}_{aj}|0\rangle$	(1e)
external	$\hat{E}_{at}\hat{E}_{bu}|0\rangle$	(1f)
	$\hat{E}_{ai}\hat{E}_{bt}|0\rangle$	(1g)
	$\hat{E}_{ai}\hat{E}_{bj}|0\rangle$	(1h)

where (i,j) are inactive, (t,u,v) are active, and (a,b) are secondary orbital indexes. The functions (1a–1h) are referred to as internal, semiinternal, and external when none, one, or two orbitals belong to the secondary subspace, respectively. It is easily realized that the single replacement states are linear combinations of the functions in (1a–1h), since $|0\rangle$ is an eigenfunction of the particle number operator

$$\sum_t \hat{E}_{tt} \tag{2}$$

with the eigenvalue equal to the number of active electrons and the index t runs over the active orbital space. The functions (1) have the same total spin and spin z component as the reference wave function and the space symmetry requirement is fulfilled by including only such excitations for which the product of the four symmetry labels contains the totally symmetric representation. We shall label the different subspaces in (1) as V_A, ..., V_H, respectively. The total dimension of this first-order space is dim V_{SD}.

In an earlier work[5] only the external subspaces V_F, V_G, and V_H of V_{SD} were used to construct the first-order wave function. This limitation was introduced for practical reasons, since the necessary matrix elements can then be expressed in terms of one- and two-particle density matrices only. It was felt at the time that the external subspace should be able to recover most of the correlation energy. However, the results obtained with this limitation clearly showed the importance of internal and semiinternal excitations.

A further partitioning of V_B, V_E, V_F, V_G, and V_H is possible. V_B is used here as an example, but the same procedure can be used for the other four subspaces.[5] For each quadruple of indexes in (1b) two states can be constructed, corresponding to different spin couplings of the new electron pair in the active space

$$(\hat{E}_{ti}\hat{E}_{uj} \pm \hat{E}_{tj}\hat{E}_{ui})|0\rangle \tag{3}$$

The plus sign leads to a singlet coupling of the inactive orbitals

i and j and the minus sign to a triplet coupling. The two sets of plus and minus combinations of states (1b) are thus functions in two mutually orthogonal subspaces V_{B^+} and V_{B^-} of V_B.

The first-order wave function is now expanded into a set of functions $|j\rangle$ from V_{SD}

$$|\Psi_1\rangle = \sum_{j=1}^{M} C_j|j\rangle \tag{4}$$

where $M \geq$ dim V_{SD} and $\{C_j, j = 1, ..., M\}$ is a solution of the system of linear equations

$$\sum_{j=1}^{M} C_j\langle i|\hat{H}_0 - E_0|j\rangle = -\langle i|\hat{H}|0\rangle, \quad i = 1, ..., M \tag{5}$$

where $E_0 = \langle 0|\hat{H}_0|0\rangle$ is the zeroth-order energy. The expansion functions $|j\rangle$ in (4) are not necessarily orthogonal and may also be linearly dependent.

2.2. The Zeroth-Order Hamiltonian. One of the major problems in formulating a low-order perturbation approach with a multiconfigurational reference function is the definition of the zeroth-order Hamiltonian. The most important condition is that it should lead to a perturbation expansion which converges rapidly. Of course, this feature can only be tested by performing actual calculations and comparing the results with corresponding full CI energies. The test calculations used in the present work are all of this kind. Secondly, the zeroth-order Hamiltonian should preferably be equivalent to the Møller–Plesset Hamiltonian in the limiting case of a closed shell reference function. One problem with the first-order equation (5) is the size of the matrix on the left-hand side of the equation. It is possible, as will be seen below, to construct the zeroth-order Hamiltonian in such a way that it leads to a further blocking of this matrix. This is of importance for the efficiency of the computer implementation of the method. The conditions given above are fulfilled for a zeroth-order Hamiltonian of the following form

$$\hat{H}_0 = \hat{P}_0\hat{F}\hat{P}_0 + \hat{P}_K\hat{F}\hat{P}_K + \hat{P}_{SD}\hat{F}\hat{P}_{SD} + \hat{P}_{TQ..}\hat{F}\hat{P}_{TQ..} \tag{6}$$

where $\hat{P}_0 = |0\rangle\langle 0|$ is the projector onto V_0, $\hat{P}_K$ is the projector onto V_K, $\hat{P}_{SD}$ is the projector onto V_{SD}, $\hat{P}_{TQ..}$ is the projector onto $V_{TQ..}$, and $\hat{F}$ is a sum of one-particle operators. With this choice of $\hat{H}_0$ only vectors belonging to V_{SD} will contribute to the first-order wave function and the second-order energy. Still the dimension of the matrix

$$(\langle i|\hat{H}_0 - E_0|j\rangle) \qquad |i\rangle, |j\rangle \in V_{SD} \tag{7}$$

is prohibitively big even for a small number of orbitals. However, if we define $\hat{F}$ to include only the diagonal terms of a one-particle operator, i.e.

$$\hat{F} = \sum_p \epsilon_p \hat{E}_{pp} \tag{8}$$

where p runs over the entire orbital space, and the off-diagonal terms are formally included in the perturbation term, we will get a blocking of the matrix (7) in 8 × 8 blocks where the off-diagonal blocks are zero, i.e.

$$\langle i|\hat{H}_0 - E_0|j\rangle = 0$$
$$\text{if } |i\rangle \in V_X, \ |j\rangle \in V_Y, \ X \neq Y \in \{A,...,H\} \tag{9}$$

If $\hat{F}$ also included the off-diagonal terms of the one-particle operator then (9) would not hold. Each of these eight diagonal blocks may now be further blocked by realizing that the matrix elements in (7) are zero if $|i\rangle \in V_{X^+}$, $|j\rangle \in V_{X^-}$, $X \in \{B,E,F,G,H\}$ and when the occupation numbers of the inactive and secondary orbitals in $|i\rangle$ and $|j\rangle$ are different.

With this choice of $\hat{F}$, the matrix 7 has a simple structure. The only nonzero elements are found in a rather large number of diagonal blocks whose dimensions are fairly small. The dimension of the largest block will be the cube of the number of active orbitals, n_a. When space symmetry is used, a further blocking of those smaller blocks is achieved. This is particularly advantageous for larger blocks involving the internal and semiinternal excitation spaces.

After the size of the matrices is reduced as far as it is possible within the present formalism, the remaining problem is then the construction of the orbital energy operator $\hat{F}$. The condition concerning the choice of this operator is that it should reproduce the results from the closed-shell second-order perturbation theory when the number of active orbitals is zero.

Let us introduce the operators

$$\hat{F}_{pq\sigma} = \hat{a}_{p\sigma}[\hat{H},\hat{a}_{q\sigma}^{\dagger}] - \hat{a}_{p\sigma}^{\dagger}[\hat{H},\hat{a}_{q\sigma}] \qquad (10)$$

whose spin-averaged expectation values we denote by

$$f_{pq} = \frac{1}{2}\sum_{\sigma}\langle 0|\hat{F}_{pq\sigma}|0\rangle \qquad (11)$$

The diagonal elements of (11) correspond to energy differences. If p represents an inactive orbital then f_{pp} is the energy difference between the system, represented by $|0\rangle$, and its positive ion. On the other hand, if p represents a secondary orbital then f_{pp} is the energy difference between the negative ion and the neutral system. For orbitals partly occupied in $|0\rangle$ f_{pp} corresponds to an interpolation between the two extremes. The evaluation of (11) yields the following explicit expression for the matrix elements f_{pq}

$$f_{pq} = h_{pq} + \sum_{rs} D_{rs}\left[(pq|rs) - \frac{1}{2}(pr|qs)\right] \qquad (12)$$

where **D** is the one-particle density matrix. The matrix 12 consists of three×three blocks corresponding to the three orbital subspaces. According to the generalized Brillouin theorem f_{pq} is zero for a converged CASSCF wave function, when one of the indexes represents an inactive orbital and the other a secondary orbital. The orbitals are determined by diagonalizing each of the diagonal blocks, one by one. This way of defining a unique set of orbitals is possible only for a CASSCF reference function, since the CAS CI space is invariant to such a transformation. It corresponds to using the canonical Hartree–Fock orbitals as the basis in the usual perturbation expansion for a single determinant closed-shell reference state. When the number of active orbitals is zero, the transformed matrix is diagonal and is identical with the canonical Fock operator in the Hartree–Fock theory.

The transformed matrix, **f'**, has the following property

$$f'_{pq} = \delta_{pq}\epsilon_p \qquad (13)$$

where the two indexes represent orbitals from the same orbital subspace. As was previously mentioned, we only include the diagonal terms (13) in the operator $\hat{F}$. Wolinski et al.[6] have argued that the full **f'** matrix should be used, since it leads to an orbital invariant formulation of the zeroth-order Hamiltonian. This would not seem to be necessary with the present procedure, which defines a unique set of orbitals. Notice also that the inactive secondary subblock of **f'** is zero for an orbital optimized CASSCF reference function, since these elements correspond exactly to the energy derivatives with respect to such orbital rotations.

2.3. The First-Order Wave Function and the Second-Order Energy.
For simplifying the notation we introduce the following matrices and vectors with elements

$$(H_0)_{ij} = \langle i|\hat{H}_0|j\rangle$$

$$S_{ij} = \langle i|j\rangle$$

C_i, the expansion coefficients for the first-order wave function; $i, j = 1, ..., M$ (14)

$$V_i = \langle i|\hat{H}|0\rangle$$

where M is the number of double replacement states (1). In determining the first-order wave function we have to solve the system of linear equations (5). Since in most cases $M > $ dim V_{SD} the double replacement states will be linearly dependent. This linear dependence (and near linear dependence) is removed by diagonalizing the overlap matrix **S** and deleting the eigenvectors which correspond to zero (or close to zero) eigenvalues. Equation 5 can now be written as

$$[(U\Lambda_s^{-1/2})^{\dagger}H_0(U\Lambda_s^{-1/2}) - E_0\mathbf{1}]\Lambda_s^{1/2}U^{\dagger}C = -\Lambda_s^{-1/2}U^{\dagger}V \qquad (15)$$

where

$$\Lambda_s = U^{\dagger}SU \qquad (16)$$

Λ_s is a diagonal $N \times N$ matrix and **U** is a $M \times N$ matrix where $N = $ dim V_{SD}. In cases with near-linear dependence $N < $ dim V_{SD}. The eq 15 corresponds to a transformation of the CI space to a orthonormalized form. The second-order energy is given as

$$E_2 = -(\Lambda_s^{-1/2}U^{\dagger}V)^{\dagger}[(U\Lambda_s^{-1/2})^{\dagger}H_0(U\Lambda_s^{-1/2}) - E_0\mathbf{1}]^{-1}(\Lambda_s^{-1/2}U^{\dagger}V) \qquad (17)$$

Because of the choice of $\hat{H}_0$, the contributions from each of the eight subspaces V_X of V_{SD} can be treated separately. As an illustrating example we choose the contribution from V_C.

$$E_2^C = -\sum_a ((U_C\Lambda_{s_c}^{-1/2})^{\dagger}V_C^a)^{\dagger}[\alpha^a\mathbf{1} + (U_C\Lambda_{s_c}^{-1/2})^{\dagger}B_C(U_C\Lambda_{s_c}^{-1/2})]^{-1}(U_C\Lambda_{s_c}^{-1/2})^{\dagger}V_C^a \qquad (18)$$

where

$$\Lambda_{s_c} = U_C^{\dagger}S_CU_C \qquad (19)$$

$$(S_C)_{tuv,xyz} = \langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{yz}|0\rangle \qquad (20)$$

$$(B_C)_{tuv,xyz} = -\epsilon_x\langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{yz}|0\rangle + \sum_w \epsilon_w\langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{ww}\hat{E}_{yz}|0\rangle \qquad (21)$$

$$\alpha^a = \epsilon_a - \sum_w \epsilon_w D_{ww} \qquad (22)$$

$$(V_C^a)_{tuv} = \sum_w [2P_{vutw} + \delta_{ut}D_{vw}][f_{aw} - \sum_y (ay|yw)] + \sum_{wxy}\langle 0|\hat{E}_{vu}\hat{E}_{tw}\hat{E}_{xy}|0\rangle(aw|xy) \qquad (23)$$

$$f_{aw} = h_{aw} + \sum_k [2(aw|kk) - (ak|kw)] \qquad (24)$$

and **D** and **P** are the one- and two-particle density matrices and k in (24) runs over the inactive and frozen orbital subspaces.

Since the summation in (18) is over the entire secondary orbital subspace it seems that a large number of matrix inversions would have to be performed. This can, however be avoided by a second unitary transformation **W** of the CI coefficients, which diagonalizes the transformed **B** matrix. Note that this transformation is also independent of the index a, and therefore has to be performed only once. Then, eq 18 simplifies to

$$E_2^C = -\sum_{a,p} |(\tilde{V}_C^a)_p|^2/(\alpha^a + (\Lambda_{\tilde{B}_C})_p) \qquad (25)$$

where

$$\Lambda_{\tilde{B}_C} = W_C^{\dagger}\tilde{B}_CW_C, \quad \tilde{B}_C = (U_C\Lambda_{s_c}^{-1/2})^{\dagger}B_C(U_C\Lambda_{s_c}^{-1/2}) \qquad (26)$$

and

$$\tilde{V}_C^a = W_C^{\dagger}\Lambda_{s_c}^{-1/2}U_C^{\dagger}V_C^a \qquad (27)$$

The same procedure can be used for all contributions. In the subspace spanned by vectors 1h the above transformation is given by a unit matrix, since the matrix representation of $\hat{H}_0 - E_0$ in V_H is automatically diagonal. The above procedure leads to a major reduction of the computational work needed to solve the first-order equations. A large number of matrix inversions is replaced by a small number of diagonalizations (two for each of the CI subspaces), with a simultaneous removal of all linear dependencies from the CI space. The expression 25 gives an illustrative formulation of the second-order energy. The denominator can be written as $\epsilon_a - \bar{\epsilon}_t$, where $\bar{\epsilon}_t$ represents an effective active orbital energy. The expression for the second-order energy with a multiconfigurational reference function thus has the same form as the usual Møller–Plesset expression in the closed-shell case, the difference being that the active orbital energies are replaced by a more complex expression involving a weighted sum over *all* active orbitals, where the weights are given by linear combinations of various (one through four particle) density matrix

TABLE I: Comparison of CASPT2 Results with Full CI (FCI) and Other MP2 Results for the H_2 Molecule

	SCF	CAS	CASPT2	FCI	MP2[b]	MP3[b]	exp[c]
R_e, Å	0.735	0.755	0.746	0.743	0.742	0.739	0.741
D_e, eV	3.63	4.14	4.57	4.67	4.56	4.62	4.75
ω_e, cm^{-1}	4616	4255	4407	4443			4401
$\omega_e x_e$, cm^{-1}	114.0	129.3	130.5	129.8			121.3
B_e, cm^{-1}	63.1	59.4	61.1	61.6			60.8
ΔE,[d] au	0.0382	0.0196	0.0037		0.0030	0.0007	

[a] Basis set: ANO (4s3p2d); active orbitals $1\sigma_g, 1\sigma_u$. [b] 6-31G** basis, GVB reference function; from ref 6. [c] Experimental results from ref 11. [d] Difference in total energy to FCI at $r(HH) = 1.4$ au.

elements. Note that such a formulation is possible for all types of excitations.

3. Some Test Applications

The test applications presented below have been performed with the purpose of comparing the present perturbation approach with full CI results obtained with the same basis set. The reference full CI results correspond to rather limited basis sets, which are incapable of yielding very accurate results for the properties studied. For this reason the performance of the CASSCF perturbation approach is studied merely from a formal point of view and the absolute accuracy of the present data is of lesser significance. The results obtained give strong evidence that the method used should be improved, before more extensive applications are made.

3.1. The Hydrogen Molecule. Calculations of the potential curve for H_2 have been performed using an ANO type basis set of the size 4s3p2d contracted from 8s4p3d.[10] The CASSCF calculations were made with the $1\sigma_g$ and $1\sigma_u$ orbitals active. They are virtually equivalent to the second-order calculations of Wolinski et al.[6] The results are presented in Table I, where a comparison is made with full CI (FCI) calculations with the same basis set, and also with the results of Wolinski et al.[6] The spectroscopic properties computed with second-order perturbation theory (CASPT2) show in general an improved agreement with the FCI results, as compared to the CASSCF data. The error in the bond distance is reduced from 0.012 to 0.003 Å, and the error in D_e from 0.53 to 0.10 eV. The present second-order results are also close to the MP2 results from ref 6. The deviation in total energy from FCI, at the HH distance 1.4 au, is only 0.0037 au, and corresponds to about 10% of the total correlation energy. Thus CASSCF plus CASPT2 is in this case capable of recovering 90% of the correlation energy. Clearly, the CASPT2 results must in this case lead to an increased value for the binding energy, since the CASPT2 energy at large distances is equal to the CASSCF energy. It is, however, not possible to conclude from these results that improved potential curves will always be obtained with this method. Actually, preliminary results for the dissociation of the nitrogen molecule indicate that in some cases the extra correlation energy obtained by CASPT2 is larger at large internuclear distances than it is arond equilibrium. In such cases the CASPT2 potential curve represents a deterioration with respect to the CASSCF results. The reason for this unbalance in the second-order energy at short and long internuclear separations is presently under investigation.

3.2. The Water Molecule in a DZ Basis. The second set of test calculations were performed for the water molecule in a DZ basis set. Here direct comparison can be made with the FCI results of Handy et al.[12] and also with the second-order results of McDouall et al.[7] (Table II). Three different active spaces were used in the CASSCF calculations, the first two being identical

(10) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. To be published.
(11) Huber, K. P.; Herzberg, G. *Constants of Diatomic Molecules*; van Nostrand Reinhold: New York, 1979.
(12) Handy, N. C. *Chem. Phys. Lett.* **1980**, *74*, 280. Harrison, R. J.; Handy, N. C. *Chem. Phys. Lett.* **1983**, *79*, 202.
(13) Bauschlicher, C. W., Jr.; Taylor, P. R. *Theor. Chim. Acta* **1987**, *71*, 273.

TABLE II: Comparison of CASPT2 Results with Full CI (FCI) and Other MP2 Results for H_2O in a DZ Basis (in au)

method[a]	r_e	$1.5r_e$	$2.0r_e$	100 Å
SCF	−76.0098	−75.8035	−75.5952	
MP2	−76.1493	−75.9946	−75.8525	
FCI[b]	−76.1579	−76.0145	−75.9052	−75.8648
MP2-FCI (2100/2020)	0.0086	0.0199	0.0527	
CASSCF	−76.0629	−75.9243	−75.8272	−75.7958
CASPT2	−76.1490	−76.0095	−75.9011	−75.8590
CASPT2-FCI	0.0089	0.0050	0.0041	0.0058
CASPT2-FCI (ref 7) (2000/2220)	0.0320	0.0302	0.0148	0.0113[c]
CASSCF	−76.0971	−75.9526	−75.8440	−75.8056
CASPT2	−76.1488	−76.0075	−75.9006	−75.8600
CASPT2-FCI	0.0091	0.0070	0.0046	0.0048
CASPT2-FCI (ref 7) (1000/4220)	0.0137	0.0117	0.0110	0.0075[c]
CASSCF	−76.1320	−75.9816	−75.8657	−75.8236
CASPT2	−76.1548	−76.0105	−75.9025	−75.8620
CASPT2-FCI	0.0031	0.0040	0.0027	0.0028

[a] Values within parentheses give the number of inactive and active orbitals used in the CASSCF and CASPT2 calculations, with inactive orbitals given first in symmetry order: a_1, b_1, b_2, and a_2. [b] Full CI results from ref 12. [c] This calculation was done at $r(OH) = 10$ Å, but the difference should be small.

TABLE III: Comparison of CASSCF, CASPT2, and FCI Binding Energies $(E(r) - E(100))$ for H_2O in a DZ Basis (kcal/mol)[a]

method[a]	r_e	$1.5r_e$	$2.0r_e$
FCI (2100/2020)	183.9	93.9	25.4
CASSCF	167.6	80.6	19.7
CASPT2	181.9	94.4	26.4
CASSCF-FCI	−16.3	−13.3	−5.7
CASPT2-FCI (2000/2220)	−2.0	0.5	1.0
CASSCF	182.9	92.2	24.1
CASPT2	181.2	92.6	25.5
CASSCF-FCI	−1.0	−1.7	−1.3
CASPT2-FCI (1000/4220)	−2.7	−1.3	0.1
CASSCF	193.5	99.1	26.4
CASPT2	183.7	93.2	25.4
CASSCF-FCI	9.6	5.2	1.0
CASPT2-FCI	−0.2	−0.7	0.0

[a] Active spaces as in Table II.

with those used in ref 7. The first calculation has four electrons in four active orbitals (2020, where the four numbers represent the number of active orbitals in the four different symmetries a_1, b_1, b_2, and a_2) and correlates only the electron pairs in the OH bonds. It is the smallest active space, which leads to correct atomic dissociation. The difference between the CASPT2 and FCI energies is with the present perturbation approach much smaller than with the effective Hamiltonian method used by McDouall et al.[7] The bond energy differences $(E(100) - E(r))$ are presented in Table III. The difference between the CASPT2 results and FCI ranges between −2.0 and +1.0 kcal/mol. The improvement compared to the CASSCF results is impressive.

The next larger active space, (2220), includes in addition two orbitals to describe the correlation of the oxygen π lone pair. However, this increase in the accuracy of the CASSCF wave function does not lead to a decrease in the difference between the CASPT2 and FCI energies and probably indicates that most of the error is located to the OH bonding region. The CASSCF potential curve obtained with this active space is fortuitously in very good agreement with the FCI results, and no further improvement is obtained at the CASPT2 level. The deviations, though somewhat smaller, are now closer to those obtained in ref 7.

The largest active space, (4220), has eight electrons in eight orbitals and adds to the preceding space two orbitals to correlate the oxygen σ lone pair. This CASSCF wave function is known

TABLE IV: The Polarizability of F⁻. A Comparison between CASPT2 and Full CI

method	polarizability, au
SCF[a]	9.89
CASSCF[b]	13.92
CASPT2	14.88
FCI[a]	16.30

[a] SCF and FCI results from ref 13. [b] 1s inactive; 2s,2p,3s,3p active.

from earlier work to give a very balanced description of the electronic structure of the water molecule.[14] The CASPT2 energies are now also much closer to the FCI values, the largest difference being 0.0040 au. The difference in binding energy is for all distances smaller than 1 kcal/mol (cf. Table III).

This set of calculations shows that in this particular case the CASPT2 energy behaves uniformly with respect to an increase of the active space. The difference in binding energy between CASSCF and FCI at r_e varies between −16.3 kcal/mol for the small active space and +9.6 kcal/mol for the largest space. The corresponding variation for CASPT2 is much smaller: from −2.7 kcal/mol (for the medium size active space) to −0.2 kcal/mol.

3.3. The Polarizability of F⁻. In the third test calculation an attempt is made to study the second-order perturbation correction to an electric property: the polarizability of F⁻. The correlation contribution to this property is substantial and it is interesting to see how much of this effect can be recovered in a second-order approach. Again we have the possibility to compare the CASPT2 data with full CI results,[13] obtained under the same conditions, that is, using the same basis set and the same electric field strength (0.005 au) in the finite field determination of the polarizability. The basis set employed by Bauschlicher and Taylor has segmented contraction and contains diffuse s, p, and d functions. The size is 5s3p2d and the details can be found in Table 1 of ref 13.

The CASSCF reference function has all eight valence electrons active and the active space consists of eight orbitals: 2s, 2s′, 2p, and 2p′ (4220 in C_{2v} symmetry). The polarizability was obtained as the second derivative of the energy with respect to the electric field strength.

The results of the calculation are presented in Table IV. The CASSCF polarizability already yields a substantial improvement compared to the SCF value. The error, as compared to the full CI result, has decreased from 6.41 au (SCF) to 2.38 au. The CASPT2 result adds only a modest improvement to this result by diminishing the difference to 1.42 au. Actually this seems to be a fortuitously good result. Preliminary results for the polarizability of other atoms and small molecules have shown that the CASPT2 perturbation correction is unstable. This is especially true in cases where an active orbital has an occupation number close to 2 and therefore easily rotates with inactive orbitals in the presence of the perturbing field. Such results strongly indicate that the reason for the instability is the neglect of the off-diagonal elements in the Fock type operator used to construct the zeroth-order Hamiltonian.

3.4. The Singlet–Triplet Splitting in CH₂. As a final test of the method we have studied the singlet–triplet splitting in the methylene radical. Again the results can be compared to the full CI calculations of Bauschlicher and Taylor.[13,15] The basis set used by them, and also here, has segmented contraction and the size (C/4s2p1d/H/2s1p). The active orbital space is constructed from the valence orbitals of carbon and hydrogen (6 electrons in 6 orbitals) and all valence electrons are correlated. The results, which are presented in Table V, are rather disappointing. The excitation energy computed with the CASSCF wave functions is already in very good agreement with the full CI results, and one would therefore hope that the perturbation correction should be small. This is not the case, however. The CASPT2 energy for the singlet state is above the FCI energy, but the situation is reversed for the triplet state. This unbalanced treatment of the

TABLE V: The Singlet–Triplet Separation in CH₂. A Comparison between CASPT2 and Full CI

	energy, au		exc energy, kcal/mol
	$^1A_1$	$^3B_1$	
CASSCF	−38.945 529	−38.965 954	12.82
PT2	−0.079 923	−0.085 211	
CASPT2	−39.025 452	−39.051 165	16.13
FCI[a]	−39.027 182	−39.046 259	11.97
FCI-CASPT2	−0.001 730	0.004 906	
OVB MP2[b]	−39.0154	−39.0366	13.42

[a] From ref 13. [b] Second-order result from ref 7.

dynamical correlation effect increases the excitation energy and destroys the good agreement with FCI obtained at the CASSCF level. Similar results have also been obtained by McDouall et al. in their second-order treatment,[7] which is based on an effective Hamiltonian formalism. The second-order correction to the excitation energy is in their treatment also positive instead of being negative (see the entry OVB MP2 in Table 5 of ref 7). The two formalisms are rather different and the similarity in the result may indicate that it is difficult in any second-order scheme to obtain a balance between corrections computed for states of different spin multiplicity.

4. Discussion

In this article we have presented a simple second-order perturbation method based on a CASSCF reference function (or more generally on a full CI function in a limited orbital subspace). Computationally the method is very efficient. The total time needed for a full calculation of the second-order energy is very small. The largest test we have performed to date was on the phenol molecule using a basis set consisting of 138 functions and coorelating all the 36 valence electrons. The CASSCF reference function was built by distributing eight π electrons in seven active orbitals. The total CASPT2 calculation, including the integral transformation, took 28 s of CPU time on an IBM 3090-17S computer. It is clear that a second-order calculation of the present type will add very little to the total time used in the preceding CASSCF calculation.

The results obtained in the test calculations presented above are, however, not entirely satisfactory. In all cases a substantial improvement in total energies has been obtained, with differences from the FCI results of the order of a few millihartrees. But even if this error is small on the absolute energy scale it is still large compared to the energy differences, which are of interest in most applications. Therefore in order to obtain an improvement in computed relative energies, the error in the second-order energy must behave in a balanced way. This has not always been the case for the selected examples.

One reason for the unbalance could be the neglect of off-diagonal contributions to the Fock operator (8) used to build the zeroth-order Hamiltonian. The number of off-diagonal elements has been reduced as much as possible by choosing the orbitals such that the diagonal blocks of f′ (inactive–inactive, active–active, and secondary–secondary) are diagonal. This is possible, since the CASSCF density matrices are invariant to the corresponding transformations. The inactive–secondary block of the f′ matrix is also automatically zero for an optimized CASSCF wave function, since it is identical with the energy derivatives with respect to the pertinent orbital rotations. However, the inactive–active and active–secondary blocks of the Fock matrix are not zero. Large rotations between orbitals in these subspaces may sometimes occur, due to some perturbation of the system (for example, an applied electric field, or a change in a bond distance). In those situations, contributions which are originally in the diagonal block of f′ may move to the nondiagonal blocks. The necessity to include the full f′ matrix in the zeroth-order Hamiltonian has been pointed out in the work of Wolinski et al.[6] with the argument that the theory is then invariant with respect to orbital rotations. Formally, this is not a problem by itself, since the orbitals are uniquely determined by the requirement that the

(14) Roos, B. *Int. J. Quantum Chem. Symp.* **1980**, *14*, 175.
(15) Bauschlicher, C. W., Jr.; Taylor, P. R. *J. Chem. Phys.* **1986**, *85*, 6510.

diagonal subblocks of the f' matrix are diagonal. However, it may be that this method does not represent a balanced choice of the orbitals.

The obvious remedy to the problem discussed above would be to include the off-diagonal blocks in the zeroth-order Hamiltonian. This can, and will, be done in the next version of the method, but it is not a trivial extension of the method, since the nice blocking of the zeroth-order matrix is then destroyed. The off-diagonal

elements will introduce couplings between the different types of excitations as given by (1). The possibility of treating this coupling in a computationally efficient way is presently being investigated.

Theoretical Studies of the Metallacyclopropenes c-[MX₂C₂H₂] (M = C, Si, Ge, Sn; X = H, F)

Jerry A. Boatz,[†] Mark S. Gordon,*

Department of Chemistry, North Dakota State University, Fargo, North Dakota 58105

and Lawrence R. Sita

Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213
(Received: November 1, 1989; In Final Form: March 15, 1990)

The geometries of the metallacyclopropenes c-[MX₂C₂H₂] (M = C, Si, Ge, Sn; X = H, F) are predicted by using the 3-21G(d) basis set and SCF wave functions. The nature of the ring bonding is investigated via analysis of the total electron density and is found to have little or no π-complex character. As a further probe of the electronic structure, bent bond lengths and intrinsic vibrational frequencies are computed. The classical barrier heights and the thermodynamics of the reaction MX₂ + HC≡CH → c-[MX₂C₂H₂] are predicted by using MP2/3-21G(d) energies.

I. Introduction

The nature of the bonding in small, highly strained ring compounds has been the topic of several recent theoretical[1-5] and experimental[6-9] papers. These studies have focused on establishing the degree of π-complex versus "classical" σ-bonding character in saturated carbon-based heterocycles[1,5] and their silicon analogues.[3-5] In particular, the degree of π-complex character as a function of heteroatom X in the cyclopropanes c-[XHₙC₂H₄] and the silicon analogues c-[XHₙSi₂H₄] has been investigated. Consistent with the Dewar[10] and Chatt–Duncanson[11] models of metal–olefin bonding described within the framework of acceptor/donor interactions, these works have predicted increasing π-complex character in these rings as the electronegativity of X increases, especially for the silicon-based rings. However, a recent paper from this laboratory[5] suggests that the apparent π-complex character is actually due to the presence of highly bent bonds.

Very recently, the first stannacyclopropene derivative, Δ¹,⁷-2,2,6,6-tetramethyl-4-thia-8,8-bis(bis(trimethylsilyl)methyl)-8-stannabicyclo[5.1.0]octene,[9] was synthesized. This derivative is observed to be in thermal equilibrium in solution with its parent stannylene and acetylene fragments, with the latter being favored at temperatures above −16 °C.[9] This observation has been explained by proposing that the bonding in the three-membered ring portion of the stannacyclopropene molecule is largely π-complex in nature. According to this picture, the facile dissociation of the ring is a result of the relatively weak π-complex binding (compared to the more usual σ-ring bonding, see structures **1** and **2**. Furthermore, replacement of the bis(trimethylsilyl) groups at the tin

center with chlorines prohibits formation of a stannacyclopropene product. The rationale for this observation is that chlorine substitution reduces the stability of an already marginally stable π-complex by decreasing the electron-donating capability of the stannylene fragment. The present work examines the bonding in the metallacyclopropenes c-[MX₂C₂H₂] (M = C, Si, Ge, Sn; X = H, F). In particular, bent bond lengths (obtained from total electron densities[12]), intrinsic frequencies[13] (obtained from computed Hessian matrices), and the energetics of the reactions MX₂ + HC≡CH → c-[MX₂C₂H₂] are calculated.

Of the ring compounds considered in this work, cyclopropene is the most studied, both experimentally and theoretically. The geometry of cyclopropene has been determined from its microwave spectrum.[14] The harmonic frequencies, normal modes of vibration,

π-complex　　　　　　　　　σ-ring
1　　　　　　　　　　　**2**

†Present address: Department of Chemistry, University of Utah, Salt Lake City, UT 84112.

(1) Cremer, D.; Kraka, E. *J. Am. Chem. Soc.* **1985,** *107,* 3800.
(2) Cremer, D.; Kraka, E. *J. Am. Chem. Soc.* **1985,** *107,* 3811.
(3) Cremer, D.; Gauss, J.; Cremer, E. *THEOCHEM* **1988,** *169,* 531.
(4) Grev, R. S.; Schaefer, H. F. III *J. Am. Chem. Soc.* **1989,** *109,* 6577.
(5) Boatz, J. A.; Gordon, M. S. *J. Phys. Chem.* **1989,** *93,* 3025.
(6) (a) Yokelson, H. B.; Millevolte, A. J.; Gillette, G. R.; West, R. *J. Am. Chem. Soc.* **1987,** *109,* 6865. (b) West, R.; De Young, D. J.; Haller, K. J. *J. Am. Chem. Soc.* **1985,** *107,* 4942.
(7) Aue, D. H.; Webb, H. M.; Davidson, W. R.; Vidal, M.; Bowers, M. T.; Goldwhite, H.; Vertal, L. E.; Douglas, J. E.; Kollman, P. A.; Kenyon, G. L. *J. Am. Chem. Soc.* **1980,** *102,* 5151.
(8) Gleiter, R.; Schäfer, W.; Baudler, M. *J. Am. Chem. Soc.* **1985,** *107,* 8043.
(9) Sita, L. R.; Bickerstaff, R. D. *J. Am. Chem. Soc.* **1988,** *110,* 5208.
(10) Dewar, M. S. *J. Bull. Soc. Chim.* **1951,** *18,* C71.
(11) Chatt, J.; Duncanson, L. A. *J. Chem. Soc.* **1953,** 2939.
(12) See, for example; (a) Bader, R. F. W.; Slee, T. S.; Cremer, D.; Kraka, E. *J. Am. Chem. Soc.* **1983,** *105,* 5061. (b) Cremer, D.; Kraka, E.; Slee, T. S.; Bader, R. F. W.; Lau, C. D. H.; Nguyen-Dang, T. T.; McDougall, P. J. *J. Am. Chem. Soc.* **1983,** *105,* 5069. (c) Bader, R. F. W.; Tang, T.-H.; Tal, Y.; Biegler-König, F. W. *J. Am. Chem. Soc.* **1982,** *104,* 940. (d) Bader, R. F. W.; Tang, T.-H.; Tal, Y.; Biegler-König, F. W. *J. Am. Chem. Soc.* **1982,** *104,* 946. (e) Bader, R. F. W.; Anderson, L.; Sjoberg, P.; Laurence, P. R. *Chem. Phys. Lett.* **1979,** *101,* 1389. (f) Runtz, G. R.; Bader, R. F. W.; Messer, R. R. *Can. J. Chem.* **1977,** 3040.
(13) (a) Boatz, J. A.; Gordon, M. S. *J. Phys. Chem.* **1989,** *93,* 1819. (b) Boatz, J. A.; Gordon, M. S. *J. Phys. Chem.* **1989,** *93,* 5774.