Theoret. Chim. Acta (Berl.) 44, 97–107 (1977)

THEORETICA CHIMICA ACTA

The Optimization of MCSCF Functions by Application of the Generalized Brillouin Theorem: The LiH_2 Potential Energy Surface

Paul J. A. Ruttink and Joop H. van Lenthe

Theoretical Chemistry Group, State University, Padualaan 8, Utrecht, The Netherlands

The generalized Brillouin theorem is used to construct an optimization procedure for MCSCF functions by iterative contracted CI calculations. Special attention is paid to the MO transformation step in each iteration. In this method the MCSCF calculation may easily be augmented by a restricted CI calculation involving a configuration set which is uniquely determined by the trial function. An application to the calculation of the potential energy surface for linear LiH_2 in the reaction $LiH + H \leftrightarrows Li + H_2$ leads to the conclusion that this restricted CI is necessary in order to obtain satisfactory results for the potential energy barrier in this reaction.

Key words: MCSCF method, optimization by Brillouin's theorem – LiH_2, potential energy surface for

1. Introduction

Since it is difficult to obtain the necessary improvements over HF results by the straightforward CI method, there is a growing interest in the MCSCF method which optimizes the MO's in a multi-configuration wave function. We here present an MCSCF optimization method based on the generalized Brillouin theorem [1]. Two advantages of this method over effective one-electron operator methods [2] are that it is conceptually simple and that it is easy to augment such a calculation by an effective restricted CI calculation. In Sect. 4 this method is applied to the calculation of the potential energy surface for linear LiH_2 in the region of interest for the exchange reaction

$$LiH + H \leftrightarrows Li + H_2.$$

2. Optimization Procedure

Our method is a combination of a suggestion made by Hinze [3] and the method used by Grein [4].

The optimization consists of the following steps:

1) An AO basis $\{\chi\}$ of dimension n is chosen and an orthogonal set of starting MO's

$\{\psi\}$ is generated. This set is divided into subsets of doubly occupied, variably occupied and empty orbitals by defining

$$\{\psi_d | n_d^0 = 2\}, \qquad \{\psi_v | 0 < n_v^0 < 2\}, \qquad \{\psi_e | n_e^0 = 0\} \tag{1}$$

where n^0 is the occupation number of an MO in the MCSCF trial function Ψ_0. The sets of occupation numbers and spin coupling schemes for the configurations Φ_k (i.e. symmetry adapted linear combinations of Slater determinants) are selected and Ψ_0 is constructed according to

$$\Psi_0 = \sum_k a_k \Phi_k \tag{2}$$

2) The coefficients a_k in (2) are determined by a (small) CI calculation.

3) All possible singly excited states Ψ_{ij}

$$\Psi_{ij} = (C_{i \to j} - C_{j \to i}) \Psi_0 \tag{3}$$

are constructed. Here $C_{i \to j}$ is a transition operator in second quantization representation [5]. This definition implies that we are working in a spin-restricted formalism. The "Brillouin states" Ψ_{ij} are unnormalized linear combinations of a (not necessarily orthogonal) set of configurations, which is uniquely determined by the set $\{\Phi_k\}$ in Eq. (2) and by the dimension of the MO basis $m (m \leqslant n)$. This set, augmented by the set $\{\Phi_k\}$, will be denoted by $\{\Phi_l\}$.

4) The H-matrix for the set $\{\Phi_l\}$ is constructed and transformed to the basis $\{\Psi_0, \Psi_{ij}\}$.

5) The expansion coefficients b_{ij} in

$$\Psi_B = b_0 \Psi_0 + \sum_{i<j} b_{ij} \Psi_{ij} \tag{4}$$

are determined by solving the corresponding secular problem.

Since we are working in a spin-restricted formalism, the Brillouin states are generally not mutually orthogonal (contrary to [3]). For the order d_B of the matrix to be diagonalized we have

$$d_B \leqslant \tfrac{1}{2} m(m-1), \qquad (m \leqslant n).$$

For the true MCSCF-MO's all coefficients b_{ij} in the eigenvector corresponding to the lowest eigenvalue vanish, because these MO's satisfy the generalized Brillouin theorem [1]

$$\langle \Phi_0 | H | \Psi_{ij} \rangle = 0 \qquad (1 \leqslant i, j \leqslant m). \tag{5}$$

6) From the information contained in the CI coefficients b_0 and b_{ij} an orthogonal MO transformation matrix T is constructed in order to obtain a better approximation for the MCSCF-MO's in the next iteration by

$$\psi' = \psi T. \tag{6}$$

Two methods to obtain T are discussed in the next section.

7) The set $\{\psi\}$ is replaced by $\{\psi'\}$ and the procedure is repeated from 3) until convergence is obtained, i.e. $|b_{ij}| < t$, $1 \leqslant i, j \leqslant m$, where t is a certain threshold.

8) The calculation may be completed by a CI calculation using the set $\{\Phi_l\}$:

$$\Psi_{\mathrm{CI}} = \sum_l a_l \Phi_l. \tag{7}$$

This set constitutes a useful subset of the full configuration set because of the following reasons:

a) It may be expected to contain the most important singly, doubly and triply excited configurations with respect to the HF determinant, if Ψ_0 contains the most important doubly excited configurations.

b) The number of independent parameters a_l in Eq. (7) is generally much larger than the number of parameters b_{ij} in Eq. (4), since roughly

$$d_l \approx d_k d_B$$

where d_l and d_k are the dimensions of the sets $\{\Phi_l\}$ and $\{\Phi_k\}$ respectively. Therefore Ψ_{CI} may be expected to lead to an appreciable energy lowering with respect to the MCSCF function Ψ_0 (which is identical to Ψ_B for the converged MCSCF-MO's, cf. Eq. (4)).

c) The calculation of Ψ_{CI} involves only an additional diagonalization, since the corresponding H-matrix has already been calculated in step 4) of the optimization procedure.

d) Ψ_{CI} has the same invariance properties under MO transformations as Ψ_0. This may be seen as follows. The MO's belonging to the subset $\{\psi_v\}$ defined in Eq. (1) are uniquely determined by the MCSCF optimization, but the MO's in $\{\psi_d\}$ and $\{\psi_e\}$ are not. A transformation within the set $\{\psi_e\}$ corresponds to a transformation of those configurations in $\{\Phi_l\}$, which have just one ψ_e singly occupied. Since all single excitations with respect to the set $\{\Phi_k\}$ are present in $\{\Phi_l\}$, this set is closed under these transformations. An analogous argument holds for $\{\psi_d\}$. Thus Ψ_{CI} is invariant under transformations which leave Ψ_0 invariant. Since the individual MO's in the subset $\{\psi_d\}$ and $\{\psi_e\}$ are not uniquely determined by our optimization procedure, this is a necessary condition to obtain results for Ψ_{CI} which are independent of the choice of starting MO's for the MCSCF optimization.

3. The MO Transformation Step

In step 5) of the optimization procedure the coefficients b_{ij} in Eq. (4) are used to generate the MO transformation matrix T in Eq. (6). The coefficients a_k in Eq. (2) are kept fixed in this step. In principle we would like to choose T such that the following criterium is satisfied

$$\langle \Psi_B - \Psi_0' | \Psi_B - \Psi_0' \rangle \quad \text{minimal} \tag{8}$$

where Ψ_0' is defined by

$$\Psi_0' = \sum a_k \Phi_k' \tag{9}$$

Φ_k' contains the improved MO's ψ' and Ψ_0' may be varied by varying T. Since T is orthogonal, Ψ_0' is normalized, and Eq. (8) is equivalent to

$$\langle \Psi_B | \Psi_0' \rangle \quad \text{maximal.} \tag{10}$$

The orthogonal transformation T in one-electron space corresponds to an orthogonal transformation T_N in N-electron space, i.e. the space spanned by the full configuration set. Therefore we have

$$\delta \langle \Psi_B | \Psi_0' \rangle = \delta \langle \Psi_B | T_N | \Psi_0 \rangle = \langle \Psi_B | \delta T_N | \Psi_0 \rangle = 0 \tag{11}$$

for variations T_N such that $T_N + \delta T_N$ still corresponds to an orthogonal transformation $T + \delta T$ in one-electron space.

Generally the relations between the matrix elements of T and T_N are rather complicated, leading to a set of coupled Nth order equations. However, since the coupling between the MO transformation and the coefficients a_k is neglected anyway in this step, it will not be efficient to solve these equations exactly. In the following, two approximate methods are discussed, which have proved to be useful in practical applications.

3.1. First-Order Method (A)

Suppose we take

$$T = a_0 I + A \tag{12}$$

where A is skew symmetrical, $|a_{ij}| \ll 1$. Then if we apply T to $\{\psi\}$ and expand Ψ_0' to terms of first order in a_{ij} we obtain

$$\Psi_0' = a_0^{N-1} (a_0 \Psi_0 + \sum_{i<j} a_{ji} \Psi_{ij}). \tag{13}$$

From Eq. (6) it is clear that Ψ_0' (apart from the normalization) may be identified with Ψ_B. Therefore, in order to make the transition from Ψ_B of Eq. (4) to Ψ_0', we take

$$T_{ij} = a_{ij} = -T_{ji} = -b_{ij} \qquad (i<j) \tag{14}$$

i.e. the non-diagonal elements of T are identified with the coefficients b_{ij} of the unnormalized Brillouin states, as defined by Eq. (3). This result has also been found by Grein [4] from the convergence properties of several alternative procedures. Eq. (12) gives slightly better results than the method used by Grein, because his formula implies $a_0 = 1$ in Eqs. (12) and (13). The matrix T defined by Eq. (12) is generally not strictly orthogonal and this has to be corrected before we apply T to $\{\psi\}$. For the orthogonalization the Löwdin method seems to be particularly suited, because it maximisses the overlap between non-orthogonal and the orthogonalized MO's. In practice, a somewhat better rate of convergence is obtained with successive Löwdin and Schmidt orthogonalizations for the various subsets defined in Eq. (1). Method A is very simple to apply, but it has the disadvantage that it is difficult to see how the orthogonalization affects the energy of the trial function, especially in cases where it is difficult to choose good starting MO's. Moreover, it may be shown that if there exists a T, such that

$$\Psi_0' = T_N \Psi_0 = \Psi_B \tag{15a}$$

i.e.

$$\langle \Psi_B - \Psi_0' | \Psi_B - \Psi_0' \rangle = 0 \tag{15b}$$

then method A will not always lead to this transformation (see appendix for details).

3.2. Density Matrix Method (B)

The norm in N-electron space of Eq. (8) may be replaced by a norm in one-electron space

$$\langle \rho(\Psi_B) - \rho(\Psi_0') | \rho(\Psi_B) - \rho(\Psi_0') \rangle \quad \text{minimal} \tag{16}$$

where $\rho(\Psi_B)$ and $\rho(\Psi_0')$ are the one-electron spinless density functions of Ψ_B and Ψ_0' respectively. Two reasons for choosing Eq. (16) are:

a) Eq. (16) leads to linear conditions on the elements of T, as will be shown below.
b) If there exists a transformation T such that Eq. (15) holds, Eqs. (8) and (16) are equivalent, since in that case

$$\langle \Psi_B - \Psi_0' | \Psi_B - \Psi_0' \rangle = \langle \rho(\Psi_B) - \rho(\Psi_0') | \rho(\Psi_B) - \rho(\Psi_0') \rangle = 0.$$

It will now be shown how Eq. (16) may be satisfied. By expressing the density functions in Eq. (16) in some arbitrary orthogonal one-electron basis $\{\phi\}$

$$\rho = \sum_{ij} P_{\phi,ij} \phi_i \phi_j$$

we obtain in terms of the density matrices $\boldsymbol{P}$

$$\mathrm{tr}\,[\boldsymbol{P}_\phi(\Psi_B) - \boldsymbol{P}_\phi(\Psi_0')]^2 \quad \text{minimal.} \tag{17}$$

If the configurations Φ_k in Ψ_0 are built from the set $\{\phi\}$ the elements of $\boldsymbol{P}_\phi(\Psi_0)$ may be expressed in the form

$$P_{\phi,ij}(\Psi_0) = \langle \Psi_0 | C_{i\to j} | \Psi_0 \rangle = \sum_k \sum_{k'} a_k a_{k'} \langle \Phi_k | C_{i\to j} | \Phi_{k'} \rangle.$$

These elements depend only on the coefficients a_k and on the occupation numbers and spin coupling schemes of the set $\{\Phi_k\}$. Since these are kept fixed in the MO transformation step (cf. Eq. (9)) we have

$$\boldsymbol{P}_{\psi'}(\Psi_0') = \boldsymbol{P}_\psi(\Psi_0). \tag{18}$$

With

$$\boldsymbol{P}_{\psi'}(\Psi_B) = \boldsymbol{T}^\dagger \boldsymbol{P}_\psi(\Psi_B)\, \boldsymbol{T},$$

Eq. (18) and using the orthogonality of T, it is easily verified that Eq. (16) is equivalent to

$$\mathrm{tr}\,[\boldsymbol{T}^\dagger \boldsymbol{P}_B \boldsymbol{T} \boldsymbol{P}_0] \quad \text{maximal} \tag{19}$$

where $\boldsymbol{P}_B \equiv \boldsymbol{P}_\psi(\Psi_B)$ and $\boldsymbol{P}_0 \equiv \boldsymbol{P}_\psi(\Psi_0)$.

Varying T in Eq. (19) yields

$$\mathrm{tr}\,[\delta \boldsymbol{T}^\dagger \boldsymbol{P}_B \boldsymbol{T} \boldsymbol{P}_0 + \boldsymbol{T}^\dagger \boldsymbol{P}_B \delta \boldsymbol{T} \boldsymbol{P}_0] = 0. \tag{20}$$

The orthogonality of T may be conserved by restricting δT to

$$\delta T = TA \tag{21}$$

where A is an arbitrary skew symmetrical matrix. Substitution of Eq. (21) into Eq. (20) yields

$$\mathrm{tr}[A\,[P_0, T^\dagger P_B T]\,] = 0. \tag{22}$$

Since P_0 and P_B are symmetrical, the commutator in Eq. (22) is skew symmetrical. Since A is an arbitrary skew symmetrical matrix, Eq. (22) can only be satisfied if

$$[P_0, T^\dagger P_B T] = 0. \tag{23}$$

Eq. (23) implies that P_0 and $T^\dagger P_B T$ have a common set of eigenvectors. If the eigenvectors and eigenvalues of P_0 and P_B are denoted by T_0, T_B, N_0 and N_B respectively, it is easily verified that Eq. (23) is satisfied by

$$T = T_B T_0^\dagger \tag{24}$$

and that the maximal value for Eq. (19) is obtained as

$$\mathrm{tr}\,[T_0 T_B^\dagger P_B T_B T_0^\dagger P_0] = \mathrm{tr}(N_B N_0). \tag{25}$$

In order to find the absolute maximum of Eq. (19), in constructing T the eigenvectors in T_0 and T_B must be ordered such that the corresponding eigenvalues in N_0 and N_B are ordered in the same way, i.e.

$$N_{0i} \leqslant N_{0j} \qquad \text{if} \qquad N_{Bi} \leqslant N_{Bj}. \tag{26}$$

Eq. (24) is consistent with the convergence criterium $|b_{ij}| < t$ $(1 \leqslant i, j \leqslant m)$, for in the limit $b_{ij} = 0$ we have

$$\Psi_B = \Psi_0, \qquad T_B = T_0 \tag{27a}$$

$$T = I, \qquad \psi' = \psi \tag{27b}$$

i.e. the MO's have converged if the Brillouin theorem Eq. (5) is satisfied. Generally Eq. (24) involves two diagonalizations in order to find the required MO transformation. The diagonalization of P_0 involves only the set $\{\psi_v\}$. Since the solution is invariant under transformations of the sets $\{\psi_d\}$ and $\{\psi_e\}$, P_B need not be completely diagonalized. In practice, however, a full diagonalization of P_B is required because T is not uniquely determined unless Eq. (26) is used.

The density matrix method has the disadvantage that the density function of Ψ_B is approximated in the mean by using the quadratic norm (16). This norm is not sensitive to variations in MO's with small occupation numbers in Ψ_0. This means that MO's occurring only in configurations with small coefficients in Ψ_0 (i.e. the correlation orbitals) are given less weight than they should have, according to the energy criterium which has been replaced by Eq. (16) for the determination of Ψ_0'. Consequently this method gives non-optimal convergence for those orbitals which (especially in the first few iterations) change most.

Another disadvantage is the fact that there is not always a one-to-one correspondence between Eqs. (4) and (27b), i.e. there are cases for which Eq. (27b) holds identically (see appendix for details). This corresponds to the fact that Eq. (23) may be satisfied identically, i.e. if P_0 and P_B are unit matrices, T is not uniquely determined by Eq. (23) and diagonalization of P_0 and P_B leads to Eq. (27b).

Obviously the density matrix method does not converge in these cases. This applies especially to MCSCF functions consisting mainly of configurations with more than one singly occupied orbital, *e.g.* VB type calculations. By comparing the two methods for determining T described above for a number of test cases we come to the following conclusions. Although with the same starting vectors rather different paths in the orbital parameter space are followed both methods generally need about the same (small) number of iterations to reach convergence. Method B appears to be somewhat safer in the case of inaccurate starting vectors, but it also seems to be slightly slower.

4. Results for Linear LiH_2

Some calculations have been carried out for linear LiH_2 with the method described above. This system, to our knowledge, has not been treated before at the MCSCF or CI level of approximation. For one geometry (R_{LiH} = 3.1 a.u., R_{HH} = 2.7 a.u.) in the saddle point region for the reaction

$$LiH + H \rightleftharpoons Li + H_2 \tag{28}$$

the results of a number of MCSCF and CI calculations were compared. The MCSCF results for the energy barrier $E(LiH_2) - E(LiH + H)$, where $E(LiH_2)$ is the energy of the saddle point geometry, were also compared with SCF and CI results.

As AO basis a split valence GTO basis was used, viz. Li(6, 2) → [3,2], H(3) → [2]. The Li 1*s* and H 1*s* exponents and contraction coefficients were taken from [6], whereas the Li 2*s* and 2*p* exponents were partially optimized in LiH at R_{LiH} = 4 a.u. with $\alpha_{2s} = \alpha_{2p}$ for each GTO. This was done because $E(Li + H_2) \ll E(LiH + H)$ and consequently the barrier height is rather sensitive to exponent variations in LiH. The Li inner shell MO was kept fixed to the SCF results, since this MO is not expected to affect the form of the potential energy surface appreciably.

The full configuration set (with the Li inner shell MO doubly occupied) contains 168 configurations. The results for various MCSCF and CI calculations with various MO basis sets for the fixed geometery are given in Table 1. The correlation energies are given in terms of the total correlation energy

$$E^t_{corr} = 0.03016 \text{ a.u.} = 0.821 \text{ eV} = 18.9 \text{ kcal/mol}$$

defined as the energy difference between the full CI result and the HF energy in this AO basis. The HF, MCSCF (set A, to be defined below), restricted CI (set $\{\Phi_l\}$) and full CI calculations were carried out for a large number of points in the region of interest for reaction (28). The results for the barrier heights and saddle point geometries are given in Table 2.

Table 1. Correlation energies for LiH_2 (R_{LiH} = 3.1 a.u., R_{HH} = 2.7 a.u.)

Method	Conf. Set	n (Conf.)	MO-Set	ΔE (% of E^t_{corr})
HF	HF	1	HF	0
CI	A	3	HF	5
CI	B	3	HF	8
CI	C	4	HF	9
CI	HF + s^a	20	HF	29
CI	$A + s^a$	48	HF	57
CI	HF + d^a	79	HF	61
MCSCF	A	3	MCSCF	62
MCSCF	C	4	MCSCF	74
INO	$A + s^a$	48	INO (HF)[c]	86
CI	$\{\Phi_l, A\}^b$	43	MCSCF	88
INO	$A + s^a$	48	INO (MCSCF)[c]	90
CI	HF + s + d^a	98	HF, INO, MCSCF	98
CI	full	168	HF, INO, MCSCF	100

[a] s and d stand for all singly and doubly excited configurations respectively.

[b] $\{\Phi_l, A\}$ is the set $\{\Phi_l\}$ generated from set A (see text).

[c] the MO-sets in parentheses are the starting MO's.

Table 2. Barrier heights (ΔE) and saddle point geometries for LiH + H $\rightleftarrows$ Li + H_2

Method	ΔE(a.u.)	R_{LiH}(a.u.)	R_{HH}(a.u.)
HF(Φ_1)	0.0101	3.30	2 60
MCSCF (set A)	0.0094	3.38	3.09
CI $\{\Phi_l, A\}$	0.0032	3.30	3.27
CI (full)	0.0029	3.30	3.25

For the MCSCF calculations several selections from the following configuration set have been used

$$\begin{aligned} \Phi_1 &= |1\,\bar{1}\,2\,\bar{2}\,3| \\ \Phi_2 &= |1\,\bar{1}\,4\,\bar{4}\,3| \\ \Phi_3 &= 6^{-1/2}\{2|1\,\bar{1}\,2\,4\,\bar{3}| - |1\,\bar{1}\,\bar{2}\,4\,3| - |1\,\bar{1}\,2\,\bar{4}\,3|\} \\ \Phi_4 &= 6^{-1/2}\{2|1\,\bar{1}\,2\,5\,\bar{3}| - |1\,\bar{1}\,\bar{2}\,5\,3| - |1\,\bar{1}\,2\,\bar{5}\,3|\}. \end{aligned} \tag{29}$$

The HF determinant Φ_1 is dominant at all geometries considered. For the reactant LiH or product H_2 the MO's ψ_2 and ψ_4 are the bonding and antibonding orbitals respectively. With this choice Φ_2 ensures a qualitatively correct dissociation behaviour into individual atoms. Φ_3 and Φ_4 are only important in the saddle point region and may therefore be expected to affect the barrier height. The spin coupling schemes of Φ_3 and Φ_4 are chosen such that they are orthogonal to the Brillouin states Ψ_{24} and Ψ_{25}, defined with respect to the full set $\{\Phi_k\}$ of Eqs. (29), which involve the same occupation numbers as Φ_3 and Φ_4.

In the set $A = \{\Phi_k | k = 1, 2, 3\}$ only one correlation orbital is involved, *viz.* ψ_4. This is a restriction, because the mixing of Φ_1 with Φ_2 or Φ_3 separately leads to different forms for this MO. Within the set $B = \{\Phi_k | k = 1, 2, 4\}$ there is also a restriction, i.e. the correlation orbitals ψ_4 and ψ_5 are assumed to be orthogonal. A function without these restrictions, i.e. a function containing the occupation numbers of set B with non-orthogonal correlation orbitals may be represented in an orthogonal MO basis by set $C = \{\Phi_k | k = 1, 2, 3, 4\}$ The results for these sets are as follows (ΔE = percentage of the correlation energy E^t_{corr}). For the fixed geometry (Table 1) all sets give poor results with HF – MO's (ΔE = 5-9%). After the MCSCF optimization set A gives ΔE = 62%, set B leads to a highly divergent iterative process and set C gives ΔE = 74%. The divergence in the calculation with set B is caused by the fact that Φ_2 and Φ_3 compete with each other in lowering the energy. Apparently the restriction in set A is less severe in this case than the restriction in set B, even though set B contains a larger number of variational parameters than set A. Although the performance of set A seems to be quite impressive, the result for the barrier height (Table 2) is only slightly better than the HF result, indicating that this set contains primarily intramolecular correlation. Since it is very unlikely that set C gives significantly better results for the barrier height, the MCSCF method with the set $\{\Phi_k\}$ of Eqs. (29) is unsatisfactory in this respect.

The MCSCF-MO's were also used for restricted CI calculations with the set $\{\Phi_l\}$ corresponding to MCSCF function A. This set $\{\Phi_l, A\}$ contains 43 configurations and yields 88% of the correlation energy at the fixed geometry. With (Schmidt orthogonalized) HF-MO's this configuration set gives only 57% of the correlation energy. With MCSCF-MO's the barrier height is found to be only 0.0003 a.u. = 0.2 kcal/mol too high (cf. Table 2). This corresponds to a barrier height of 31% of the HF result and to a barrier height lowering with respect to the HF result of 96% of the full CI result. The saddle point geometry is also well predicted by this restricted CI calculation. Since these results are much better than the MCSCF results, we see that our method for restricted CI calculations with the set $\{\Phi_l\}$ as defined in Sect. 2 leads to much better agreement with full CI results than a simple MCSCF calculation. Also, by comparing the correlation energy contributions and the lowering of the barrier heights for the MCSCF and restricted CI calculation respectively, it may be seen that the intermolecular correlation contributions are relatively better taken care of in the restricted CI calculation than in the MCSCF calculation. It is, however, clearly difficult to separate inter- and intramolecular effects in this case.

The calculations for the fixed geometry were also compared with INO calculations [7] with a fixed configuration set, similar to the set $\{\Phi_l, A\}$. Owing to program restrictions we had to include some extra configurations into this set, leading to a set of 48 configurations. These calculations appear not to lead to unique results, as might have been expected from the fact that the INO method does not use T as defined in Eq. (24) for the MO transformation step. After a number of iterations the energy reaches a minimum and after that the process diverges. If the energy minimum is taken as the final energy, the result appears to depend on the choice of the starting MO's. Thus we find ΔE = 86% and ΔE = 90% for HF and MCSCF-MO's respectively. This problem is less serious in a large CI calculation e.g. including Φ_1 and all single and double excitations

(98 configurations). However, in this case the result turns out to be insensitive to the choice of MO's, since HF, MCSCF and INO-MO's all lead to the same result ($\Delta E = 98\%$).

5. Conclusions

We have shown that the spin-restricted MCSCF optimization method based on the generalized Brillouin theorem for an orthogonal MO-basis leads to an iterative process with satisfactory convergence properties. The rate of convergence depends in most cases only slightly on the method used for the MO transformation step whereas in extreme cases the two methods considered behave quite differently. From the configuration set chosen for the trial function a useful configuration set for restricted CI calculations may be generated, which has been proved to give much better results than a simple MCSCF calculation for the characteristics of the potential energy surface for the exchange reaction

$$\mathrm{LiH} + \mathrm{H} \rightleftarrows \mathrm{Li} + \mathrm{H}_2.$$

Acknowledgements. The authors wish to thank Dr. J. A. Vliegenthart and Dr. F. B. van Duijneveldt for many stimulating discussions and helpful comments. The investigations were supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

Appendix

The performance of the two methods for choosing T in Eq. (6) considered above if Eqs. (15) are satisfied may conveniently be compared by considering a simple (and therefore rather extreme) example.

We choose

$$\Psi_0 = 2^{-1/2}[|a\bar{b}| - |\bar{a}b|] \tag{A1}$$

and from Eqs. (3), (4) and (9) we find, confining our attention for simplicity to the mixing of the orbitals a and b,

$$\begin{aligned} \Psi_{ab} &= 2^{1/2}[|b\bar{b}| - |a\bar{a}|] \\ \Psi_B &= b_0\Psi_0 + b_{ab}\Psi_{ab} \qquad \text{(A2)} \\ \Psi_0' &= \cos 2\alpha\Psi_0 + \tfrac{1}{2}\sin 2\alpha\Psi_{ab}. \end{aligned}$$

Eqs. (15) may now be satisfied by choosing α such that

$$\mathrm{tg}2\alpha = 2b_{ab}/b_0.$$

The first-order substitution method leads to

$$\mathrm{tg}\,\alpha = b_{ab}/b_0.$$

This is only a good approximation to the exact result for small values of α.

The density matrix method is formally correct in this case. However, for the above example it may easily be shown that it does not work in practice, since P_0 and P_B

are both unit matrices in this case. Therefore Eq. (23) is identically satisfied and consequently α cannot be determined from Eq. (24).

This result for the density matrix method may be generalized as follows.

For any function

$$\Psi_0 = \left| \prod_{i=1}^{n} \psi_i \Theta \right| \tag{A3}$$

where Θ is a linear combination of spin products and the set $\{\psi\}$ is an orthogonal set of space functions of dimension n, we have

$$P_0 = P_B = I. \tag{A4}$$

This result is independent of the coefficients b_{ij} in Eq. (4). For the proof of Eq. (A4) the commutation rules for the excitation operators $C_{i \to j}$ and the anti-Hermitian property of the operator $C_{i \to j} - C_{j \to i}$ in Eq. (3) may be used. Since it is straightforward but somewhat lengthy it will not be reproduced here.

References

1. Levy, B., Berthier, G.: Intern. J. Quantum Chem. **2**, 307 (1968)
2. Das, G., Wahl, A. C.: J. Chem. Phys. **56**, 1769 (1972)
3. Hinze, J.: J. Chem. Phys. **59**, 6424 (1973)
4. Banerjee, A., Grein, F.: Int. J. Quantum Chem. **10**, 123 (1976); Grein, F., Chang, T. C.: Chem Phys. L. **12**, 44 (1971)
5. Ruttink, P. J. A.: Theoret. Chim. Acta (Berl.) **36**, 289 (1975)
6. van Duijneveldt, F. B.: IBM Technical Report RJ 945 (1971)
7. Bender, C. F., Davidson, E. R.: J. Phys. Chem. **70**, 2675 (1966)

Received June 9, 1976/September 14, 1976