Theor Chim Acta (1989) 76: 95–111

Theoretica
Chimica Acta

An efficient first-order CASSCF method based on the renormalized Fock-operator technique

U. Meier and V. Staemmler
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-4630 Bochum,
Federal Republic of Germany

(Received August 1, 1988; revised March 8/Accepted April 13, 1989)

Summary. A new efficient first-order CASSCF method (multiconfiguration SCF (self consistent field) in a complete active space) is described. Its main characteristics are (i) use of the generalized Brillouin theorem (Fock-operator method), (ii) renormalization of single excitations, (iii) fast microiterations containing only two-index transformations, i.e. M^3N^2 steps. Convergence is generally reached in eight to twelve macroiterations. The method is applied to several examples (LiH, N_2, AlO) and compared to other MCSCF (multiconfiguration SCF) methods.

Key words: First-order CASSCF — Brillouin conditions — Renormalization of single excitations

1. Introduction

Both first-order (linearly convergent) and second-order (quadratically convergent) MCSCF methods are widely used in modern electronic structure calculations. For very detailed recent reviews we refer to the papers of Olsen et al. [1], Werner [2] and Shepard [3] concerning second-order methods and Roos [4] concerning first-order methods. First-order methods are probably more frequently used in numerical applications, since they are conceptually simpler and easier to implement. However, general first-order MCSCF methods converge very slowly, if they converge at all; reasonable convergence behaviour could only be achieved for the special class of CASSCF (complete active space SCF) wavefunctions [4]. Therefore much effort has been concentrated in the past few years on improving second-order MCSCF methods. New developments of first-order methods were limited to two-configuration SCF wavefunctions [5–8].

In this paper we present a simple first-order CASSCF method based on the Fock-operator technique. The method is closely related both to the formalism proposed by Hinze [9, 10] (but the single excitations are renormalized) and to the super-CI (configuration interaction) CASSCF method of Roos et al. [11–13]. The main idea is that the time consuming partial integral transformation in the microiterations can be completely avoided with the consequence that the microiterations become very fast. This idea has not yet been used in first-order methods, but is closely related to techniques applied in second-order MCSCF programs. Convergence is generally reached in eight to twelve (macro) iterations which corresponds to the convergence rate of standard closed shell SCF calculations. This also holds in many cases for excited states which are not the lowest of their symmetry as well as for "average of configuration" states.

We will show for some small and medium size examples that our first-order method is competitive with second-order methods as far as convergence rate and computational effort are concerned. In addition it is much simpler from the point of view of programming since no second derivatives are needed and the Fock equations are solved iteratively as in conventional closed-shell SCF methods. The simplicity of the algorithm also allows highly vectorized computer code.

2. Fock-operator method

For the definition of the Fock matrix elements we adopt the notation of Roos et al. [12, 13]. The total orbital space is decomposed into three subspaces containing inactive (core), active, and virtual orbitals. We will use p, q, r, s as general MO (molecular orbital) indices, i, j, k, l as core MO indices, t, u, v, w as active MO indices, and a, b, c, d as virtual MO indices.

The multiconfiguration wavefunction ψ_0 is a sum of Slater determinants (SD) or configuration state functions (CSF).

$$\psi_0 = \sum_I C_I \phi_I. \tag{1}$$

The wave function ψ_0 that minimizes the energy has to satisfy the orbital restricted Brillouin condition [14]

$$\langle \psi_0 | \hat{H} | (\hat{E}_{pq} - \hat{E}_{qp}) \psi_0 \rangle = 0, \tag{2}$$

where $\hat{E}_{pq}$ is a spin averaged excitation operator. The Brillouin condition is modified in the way suggested by Roos [12]

$$\psi_{pq} = \frac{1}{\sqrt{|\gamma_{pp} - \gamma_{qq}|}} \hat{E}_{pq} \psi_0, \tag{3}$$

where γ_{pp} is the occupation number of the orbital p, such that the singly excited functions are normalized

$$\langle \psi_{pq} | \psi_{pq} \rangle = 1. \tag{4}$$

The modified Brillouin condition can be written as

$$B_{pq} = \langle \psi_0 | \hat{H} | (\psi_{pq} - \psi_{qp}) \rangle = 0. \tag{5}$$

For later use we define the first- and second-order reduced density matrices γ_{pq} and Γ_{pqrs} by

$$\gamma_{pq} = \sum_{I,J} C_I \langle \phi_I | \hat{E}_{pq} | \phi_J \rangle C_J, \tag{6}$$

$$\Gamma_{pqrs} = \frac{1}{2} \sum_{I,J} C_I \langle \phi_I | \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} | \phi_J \rangle C_J. \tag{7}$$

For the evaluation of the matrix elements B_{pq} one makes use of special properties of the density matrices

$$\gamma_{ii} = 2, \tag{8}$$

$$\Gamma_{iijj} = 2, \qquad \Gamma_{ijji} = -1, \qquad \Gamma_{iiii} = 1, \qquad \Gamma_{iitu} = \gamma_{tu}, \qquad \Gamma_{itui} = -\tfrac{1}{2}\gamma_{tu}. \tag{9}$$

All other density matrix elements with combinations of inactive indices different from those in Eqs. (8) and (9) vanish as well as all density matrix elements having at least one virtual index. This holds whenever the core orbitals are completely occupied and the virtual orbitals completely empty, i.e. $\gamma_{aa} = n_a = 0$ for the virtual orbitals and $\gamma_{ii} = n_i = 2$ for the core orbitals in each SD (or CSF) entering in (1).

In the case of a CASSCF (complete active space SCF) wavefunction [12] where the CI-expansion within the active orbital subspace is complete only three types of singly excited functions ψ_{pq} interact with ψ_0, namely core-virtual, active-virtual and core-active excitations. The modified Brillouin condition (5) for these excitations can be expressed with the help of two operators

$$F_{pq}^{\text{core}} = h_{pq} + \sum_i \{2(pq|ii) - (pi|qi)\} \tag{10}$$

$$G_{pq}^{\text{act}} = \sum_{tu} \gamma_{tu} \{(pq|tu) - \tfrac{1}{2}(pt|qu)\} \tag{11}$$

as

core → virtual $(i \to a)$

$$B_{ai} = \frac{1}{\sqrt{2}} \{2F_{ai}^{\text{core}} + 2G_{ai}^{\text{act}}\} \tag{12}$$

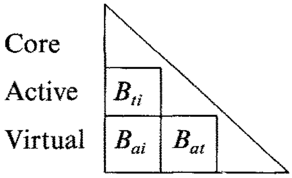
active → virtual $(t \to a)$

$$B_{at} = \frac{1}{\sqrt{\gamma_{tt}}} \left\{ \sum_u \gamma_{tu} F_{au}^{\text{core}} + 2 \sum_{uvw} \Gamma_{tuvw} (au|vw) \right\} \tag{13}$$

core → active $(i \to t)$

$$B_{ti} = \frac{1}{\sqrt{2 - \gamma_{tt}}} \left\{ 2F_{ti}^{\text{core}} + 2G_{ti}^{\text{act}} - \sum_u \gamma_{tu} F_{ui}^{\text{core}} - 2 \sum_{uvw} \Gamma_{tuvw} (iu|vw) \right\}. \tag{14}$$

We now define a hermitian matrix $\boldsymbol{B}$ containing the matrix elements (12)–(14) in the off-diagonal blocks [15, 16]



If this matrix is diagonal, obviously, the Brillouin conditions (12)–(14) are satisfied. One can even add arbitrary diagonal blocks (B_{ij}, B_{tu}, B_{ab}) without changing the Brillouin conditions (12)–(14). However, a reasonable choice of the diagonal blocks might accelerate the convergence rate. In our treatment they are set to zero except for the diagonal elements which are chosen as

$$B_{pp} = \frac{1}{\sqrt{2}} \{2F^{\text{core}}_{pp} + 2G^{\text{act}}_{pp}\}. \tag{15}$$

It should be emphasized that in general MCSCF wavefunctions the block B_{tu} contains Brillouin conditions for rotations between partly occupied orbitals and cannot be chosen arbitrarily.

The definition of the off-diagonal blocks of the operator $\boldsymbol{B}$ given so far is equivalent to the one used in other MCSCF methods [11–13, 17–23]. We have only repeated those formulas which are necessary to understand the simplification of the microiterative equations which we shall derive in the next section; for more details we refer the reader to [12, 13].

3. Microiterative equations

The diagonalization of the $\boldsymbol{B}$-matrix

$$\boldsymbol{B}\boldsymbol{C}' = \varepsilon \boldsymbol{C}' \tag{16}$$

defines a unitary transformation from the old orbitals $\boldsymbol{C}$ to a new set of orbitals $\boldsymbol{C}'$

$$\boldsymbol{C}' = \boldsymbol{U}\boldsymbol{C} = (\boldsymbol{1} + \boldsymbol{D})\boldsymbol{C}, \tag{17}$$

$$|dp\rangle = |p'\rangle - |p\rangle = \sum_q |q\rangle D_{pq}. \tag{18}$$

Note, that according to the definition of $\boldsymbol{B}$ elements D_{pq} connecting the same orbital subspace are zero to first order. In conventional SCF or MCSCF procedures a microiteration consists of a diagonalization of the $\boldsymbol{B}$-matrix and a recalculation of $\boldsymbol{B}'$ in the new vectors $|p'\rangle$ without changing the density matrix elements γ_{tu} and Γ_{tuvw} that enter into Eqs. (13) and (14).

If we apply this procedure for B_{at} (Eq. (13)) and take only linear terms in the D_{pq} explicitly (assuming for a moment that F^{core} is a constant one-electron operator), we get

$$B'_{a't'} = B_{at} + \frac{1}{\sqrt{\gamma_{tt}}} \left\{ \sum_{pu} \gamma_{tu} [F^{\text{core}}_{up} D_{pa} + F^{\text{core}}_{ap} D_{pu}] \right.$$
$$+ 2 \sum_{puvw} \Gamma_{tuvw} [(pu|vw) D_{pa} + (ap|vw) D_{pu}$$
$$\left. + (au|pw) D_{pv} + (au|vp) D_{pw}] \right\} + O(D^2), \tag{19}$$

or if we express the two-electron integrals by J and K operators

$$B'_{a't'} = B_{at} + \frac{1}{\sqrt{\gamma_{tt}}} \left\{ \sum_{pu} \gamma_{tu} [F^{\text{core}}_{up} D_{pa} + F^{\text{core}}_{ap} D_{pu}] \right.$$
$$+ 2 \sum_{puvw} \Gamma_{tuvw} [\langle p|J_{vw}|u\rangle D_{pa} + \langle a|J_{vw}|p\rangle D_{pu}$$
$$\left. + \langle a|K_{uw}|p\rangle D_{pv} + \langle a|K_{uv}|p\rangle D_{pw}] \right\} + O(D^2), \tag{20}$$

where p runs over all orbital subspaces. If we neglect the D_{pq} elements which connect orbitals of the same subspace, the only exchange integrals that enter into Eq. (20) are of the following type

$$\langle a|K_{uv}|b\rangle D_{bw}, \qquad \langle a|K_{uv}|i\rangle D_{iw}. \tag{21}$$

Apparently, they have two indices in the active space and two indices outside the active space. These integrals can be assumed to be small, especially when the orbital subspaces are well separated and a certain shell structure is reproduced by the choice of the orbital spaces. $B'_{a't'}$ can thus be approximated by

$$B'_{a't'} \simeq \tilde{B}_{a't'} = B_{at} + \frac{1}{\sqrt{\gamma_{tt}}} \left\{ \sum_{pu} \gamma_{tu} [F^{\text{core}}_{up} D_{pa} + F^{\text{core}}_{ap} D_{pu}] \right.$$
$$\left. + 2 \sum_{puvw} \Gamma_{tuvw} [\langle p|J_{vw}|u\rangle D_{pa} + \langle a|J_{vw}|p\rangle D_{pu}] \right\}. \tag{22}$$

According to the definitions of $\boldsymbol{B}$ and $\boldsymbol{C}'$ this can be rewritten as

$$\tilde{B}_{a't'} = \frac{1}{\sqrt{\gamma_{tt}}} \left\{ \sum_{u} \gamma_{tu} \langle a'|F^{\text{core}}|u'\rangle + 2 \sum_{uvw} \Gamma_{tuvw} \langle a'|J_{vw}|u'\rangle \right\}. \tag{23}$$

The changes in F^{core} and G^{act} can be neglected for the same reasons.

A similar consideration leads to

$$\tilde{B}_{t'i'} = \frac{1}{\sqrt{2 - \gamma_{tt}}} \left\{ 2\langle t'|F^{\text{core}}|i'\rangle + 2\langle t'|G^{\text{act}}|i'\rangle - \sum_{u} \gamma_{tu} \langle u'|F^{\text{core}}|i'\rangle \right.$$
$$\left. - 2 \sum_{uvw} \Gamma_{tuvw} \langle u'|J_{vw}|i'\rangle \right\}, \tag{24}$$

$$\tilde{B}_{a'i'} = \frac{1}{\sqrt{2}} \{ 2\langle a'|F^{\text{core}}|i'\rangle + 2\langle a'|G^{\text{act}}|i'\rangle \}. \tag{25}$$

Thus the matrix $\boldsymbol{B}'$ can be approximated by the matrix $\tilde{\boldsymbol{B}}$ which is simply evaluated by transforming the J_{vw}-operators as well as F^{core} and G^{act} into the new basis $\boldsymbol{C}'$. In contrast to conventional, e.g., closed-shell, SCF methods this transformation does not mean just a change of the basis for a fixed operator, but because of the u-summation in Eqs. (23) and (24) a construction of new operators (except for Eq. (25)).

The matrix $\tilde{\boldsymbol{B}}$ defines a new unitary transformation and therefore a new microiterative cycle. This is repeated until $\tilde{\boldsymbol{B}}$ becomes diagonal. This does not imply that $\boldsymbol{B}'$ itself is diagonal at the end of a microiterative cycle. However, since our approximations affect only the second derivatives of the energy with respect to the orbitals and since the first derivatives, i.e. B_{at} etc., are calculated exactly once in each macroiteration, the whole process converges to the correct solution. It should be noted that it is necessary to choose the "diagonal" blocks of $\boldsymbol{B}$ as zero (except for the diagonal elements); otherwise the D_{pq} connecting orbitals of the same subspace are not zero to first order and the approximations leading to (23)–(25) are not justified. This implies that our procedure is only reasonable for CASSCF wavefunctions. This point has been mentioned before by Werner and Meyer [17].

Thus, our first-order CASSCF method consists of the conventional interplay between macroiterations and microiterations: each macroiteration contains a complete CI calculation in the active space, the calculation of the density matrices γ and Γ and the construction of the operators F^{core}, G^{act} and all J_{vw} with v and w running over the active orbitals. In the microiterations only transformations of the operators F^{core}, G^{act} and J_{vw} are necessary, this needs only M^2N^3 floating point operations, where N is the number of basis functions and M the number of active orbitals. Equations (23)–(25) show that only J-operator matrix elements containing three active indices are needed, which cuts the number of floating point operations further to $\frac{1}{2}M^3N^2$ if the matrix elements of the second-order density matrix, Eq. (7), are symmetrized in the first two and in the last two indices. Since the microiterations need no N^4 steps, i.e. no sweep through the AO (atomic orbital) integral list, they are very fast and it not so very crucial to keep their number as small as possible. If one is not too far from the final solution, 10–15 microiterations are sufficient, however the approximations in Eqs. (19)–(22) result in much worse convergence if one is far from the desired solution.

It should be noted that Werner and Knowles [19, 20] use similar "one-index transformations" in their orbital optimization. But since they calculate second derivatives they also have to calculate and transform the K operators; furthermore, their microiterations contain CI-coefficient updates which are completely absent in our scheme.

The renormalization (3) of the single excitations resulting in the factor $(|\gamma_{pp} - \gamma_{qq}|)^{-1/2}$ in the B_{pq} elements is crucial for the convergence behaviour of our method. This factor gives a large weight to those rotations that do not affect the energy very much, i.e. those between core orbitals and almost doubly occupied active orbitals as well as between virtual orbitals and nearly empty active orbitals. Without introducing these factors, in many cases with weakly occupied active orbitals convergence could not be achieved, but this renormalization almost guarantees convergence.

Finally, we applied the usual methods (damping, energy denominator shifts, DIIS [24, 25] etc.) to improve convergence both of the microiterations and of the macroiterations. For the microiterations Pulay's DIIS method [24] is switched on automatically as soon as the squared norm of the residual vector (the parameter λ in [24]) is smaller than $\sim 10^{-5}$ (in most cases after 5–6 microiterations). The same threshold is used for the macroiterations, but is generally reached only after 7–10 iterations.

4. Test calculations

First tests of our method have been performed for the two lowest $^1\Sigma^+$ states of LiH at an internuclear distance $R = 3.0a_0$ which is close to the equilibrium distance of the $X^1\Sigma^+$ ground state. At this distance the ground state wavefunction is dominated by the configuration $1\sigma^2 2\sigma^2$, all other configurations – even $1\sigma^2 3\sigma^2$ which is necessary for the proper dissociation into Li and H – are rather weakly occupied. The inclusion of the very weakly occupied 4σ and 5σ orbitals in the active space causes severe convergence problems in conventional first- and second-order MCSCF procedures.

For comparison we have also used the second-order MCSCF program of Shepard et al. [21–23] in a version adjusted by Schmitz [26] to the Cyber 205. These calculations are denoted by UEXP2 in the tables. The calculations with our new first-order method are denoted by RFO1 ("renormalized Fock operator, first-order").

The Gaussian basis set for LiH has been taken form [18] and consists of

Li: $11s5p$ in the contraction 5, $6 \times 1/2$, 3×1,

H: $7s2p$ in the contraction 3, $4 \times 1/2 \times 1$.

Slight differences between our final energies and those reported by other authors or calculated by Shepard's program are due to the fact that we are using Gaussian lobe functions [27, 28] instead of Cartesian Gaussians [29, 30].

For all RFO1 and UEXP2 calculations on LiH as presented in the Tables 1a–d we always started from the (occupied and virtual) SCF orbitals of LiH^+. Table 1a shows that our convergence is similar to that obtained with Shepard's program for the calculation with three active orbitals (CAS3) for the $X^1\Sigma^+$ ground state.

We have performed the same CAS3 calculation under four different conditions:

i) With renormalization (Eq. (3)) and with microiterations as described in the previous section: 9 macroiterations were sufficient, as documented in Table 1a.

ii) Without renormalization, but with microiterations: 16 macroiterations ($\Delta E = -3.7E - 10$).

Table 1. CASSCF calculations for LiH[a]

	a. $X^1\Sigma^+$ active space: $(2\sigma, 3\sigma, 4\sigma)$		b. $X^1\Sigma^+$ active space: $(2\sigma, 3\sigma, 4\sigma, 5\sigma)$				c. $A^1\Sigma^+$ active space: $(2\sigma, 3\sigma, 4\sigma, 5\sigma)$	d. $X^1\Sigma^+$ and $A^1\Sigma^+$ active space: $(2\sigma, 3\sigma, 4\sigma, 5\sigma)$	
Macroiteration	RFO1	UEXP2	RFO1		UEXP2		RFO1	RFO1	
	ΔE/a.u.	ΔE/a.u.	E/a.u.	ΔE/a.u.	E/a.u.	ΔE/a.u.	ΔE/a.u.	$X^1\Sigma^+$ ΔE/a.u.	$A^1\Sigma^+$ ΔE/a.u.
1	$-3.9E-2$	$-3.8E-2$	-7.998928	$-4.0E-2$	-7.993830	$-3.8E-2$	$-1.5E-2$	$-4.3E-2$	$-1.5E-2$
2	$-6.2E-3$	$-8.4E-3$	-8.007319	$-8.3E-3$	-8.001785	$-7.9E-3$	$-3.6E-4$	$-5.2E-3$	$+3.3E-4$
3	$-4.5E-3$	$-1.1E-3$	-8.008198	$-8.8E-4$	-8.003338	$-1.5E-3$	$-5.3E-4$	$-6.7E-4$	$-1.0E-3$
4	$-3.5E-4$	$-3.7E-3$	-8.008379	$-1.8E-4$	-8.007536	$-4.2E-3$	$-5.3E-4$	$-3.9E-5$	$-2.5E-4$
5	$-7.2E-5$	$-1.5E-4$	-8.008503	$-1.2E-4$	-8.008362	$-8.5E-4$	$-4.1E-4$	$-1.1E-5$	$-4.6E-5$
6	$-1.6E-5$	$-5.6E-6$	-8.008842	$-3.4E-4$	-8.008420	$-5.8E-5$	$-1.3E-5$	$-5.1E-6$	$-5.3E-6$
7	$-4.9E-5$[b]	$-3.5E-7$	-8.008914	$-7.2E-5$	-8.008424	$-4.1E-6$	$-1.5E-6$	$-4.1E-6$[b]	$+1.8E-6$[b]
8	$-1.6E-8$	$-2.7E-8$	-8.008933	$-2.0E-5$	-8.008760	$-3.4E-4$	$-2.2E-7$[b]	$-9.8E-8$	$+8.9E-8$
9	$-1.4E-10$	$-2.2E-9$	-8.008941	$-7.6E-6$	-8.008943	$-1.8E-4$	$-3.7E-8$	$+1.7E-8$	$-1.8E-8$
10			-8.008945	$-3.6E-6$	-8.008953	$-1.0E-5$	$-5.8E-9$	$+4.0E-9$	$-4.2E-9$
11			-8.008946	$-1.8E-6$	-8.008954	$-9.8E-7$	$-3.4E-10$	$-6.0E-10$	$+5.4E-10$
12			-8.008949	$-2.2E-6$[b]	-8.008954	$-7.5E-8$			
13			-8.008949	$-1.1E-9$	-8.008954	$-5.0E-9$			
Final energy/a.u.	-8.008417	-8.008423					-7.887889	-8.007776	-7.887040

[a] Basis set see text; $R = 3.0a$

[b] A DIIS extrapolation step was performed in this macroiteration

iii) With renormalization, but only one Fock-operator diagonalization per macroiteration: after 100 macroiterations ΔE was still as large as $-2.3E-8$.

iv) Without renormalization and with only one Fock-operator diagonalization per macroiteration: $\Delta E = -9.2E-6$ after 100 macroiterations.

Pulay's convergence acceleration [24] could not be applied to the two calculations without microiterations because of strong linear dependence of the consecutive Fock operators.

In the calculation with four active orbitals (CAS4), Table 1b, the 5σ orbital is very weakly occupied ($n_{5\sigma} = 0.000391$). Shepard's method seems to converge first to the CAS3 (2σ, 3σ, 4σ) solution and to start optimizing the 5σ-orbital only after seven macroiterations. Due to the renormalization, Eq. (3), our method converges much more smoothly to the CAS4 result. Table 1b shows that our first-order methods need no more macroiterations than Shepard's second-order treatment. The final convergence is obtained in the RFO1 calculations by application of Pulay's DIIS [24]. Convergence of the first excited ${}^1\Sigma^+$ state is reached in 11 iterations as shown in Table 1c. With our version of Shepard's program this state did not converge.

In a density matrix formulated MCSCF algorithm a simultaneous optimization of several states can easily be performed by averaging the first- and second-order reduced density matrices [18]

$$\gamma_{pq} = \sum_i w^i \gamma^i_{pq}, \tag{26}$$

$$\Gamma_{pqrs} = \sum_i w^i \Gamma^i_{pqrs},$$

$$\sum_i w^i = 1. \tag{28}$$

The w^i are arbitrary weighting factors for the states i. The energies obtained for each state i are upper bounds for the true MCSCF energies. Table 1d shows a calculation for $X\,{}^1\Sigma^+$ and $A\,{}^1\Sigma^+$ with equal weights for both states. A similar calculation with a smaller basis set has been performed by Werner and Knowles [20] who could obtain convergence after three of their macroiterations. We repeated the calculation with the small basis given in [20]; the convergence pattern is the same as for the larger basis in Table 1d.

Our second example is the ground state of N_2. Tables 2–4 contain our results for different basis sets and different active spaces together with results reported in the literature. All calculations performed by us (RFO1, UEXP2) have been started with SCF orbitals from N_2^+, ${}^2\Sigma_g^+$ throughout.

The comparison of our first-order CASSCF with second-order methods for the CAS6 (six active orbitals) calculation in Table 2 shows that we need about twice as many macroiterations (including one Pulay extrapolation in iteration 8) as the second-order programs. The new technique of Werner and Knowles [19, 20] converges even faster. By including CI updates during the microiterations they are generally able to achieve convergence after three to four

Table 2. CASSCF calculations for N_2, $X^1\Sigma_g^+$, $R = R_e = 2.094a_0$; active space $(3\sigma_g, 1\pi_u, 1\pi_g, 3\sigma_u)$[a]

Macroiteration	RFO1[a] ΔE/a.u.	UNEXP2[a] ΔE/a.u.	Jørgensen et al.[b] ΔE/a.u.	Werner/Knowles[c] ΔE/a.u.
1	$-5.8E-2$	$-5.2E-2$	$-1.9E-2$	$-5.4E-2$
2	$-2.9E-2$	$-2.8E-2$	$-1.7E-2$	$-2.0E-2$
3	$-1.5E-2$	$-2.1E-2$	$-1.7E-2$	$-1.9E-6$
4	$-5.3E-3$	$-1.0E-2$	$-2.7E-3$	$<-1.0E-12$
5	$-1.5E-3$	$-4.3E-4$	$-3.8E-5$	
6	$-3.6E-4$	$-6.9E-7$	$-1.3E-8$	
7	$-8.5E-5$			
8	$-2.3E-5$[d]			
9	$-7.3E-8$			
10	$-3.2E-9$			
Final energy/a.u.	-109.103792	-109.103994	-109.095718	-109.095718

[a] Basis set I: Huzinaga (51111/311) + 1d (exponent 0.95)

[b] Fletcher optimization [31], basis set see [31]

[c] [20], basis set see [31]

[d] A DIIS extrapolation step was performed in this macroiteration

Table 3. CASSCF calculations for N_2, $X^1\Sigma_g^+$, $R = R_e = 2.094a_0$; active space $(2\sigma_u, 3\sigma_g, 1\pi_u, 1\pi_g, 3\sigma_u)$

Macroiteration	ROF1[a] ΔE/a.u.	Roos et al.[b] ΔE/a.u.
1	$-4.1E-1$	$-1.8E+0$
2	$+1.5E-4$	$-1.5E-1$
3	$-2.3E-3$	$-2.3E-2$
4	$-3.2E-2$	$-1.0E-2$
5	$-4.2E-3$	$-2.3E-2$
6	$-3.4E-3$	$-1.5E-2$
7	$-3.8E-5$	$-5.3E-3$
8	$-5.5E-6$	$-1.4E-3$
9	$-1.4E-6$[c]	$-3.0E-4$
10	$-1.5E-8$	$-7.0E-5$
11	$-3.5E-10$	$-3.0E-5$
12		$-2.0E-5$
13		$-1.0E-5$
14		$<-1.0E-5$
Final energy/a.u.	-109.110819	-109.11271

[a] Basis set I, see Table 2

[b] Super-CI method [12], basis set [12]

[c] A DIIS extrapolation step was performed in this macroiteration

Table 4. CASSCF calculations for N_2, $X^1\Sigma_g^+$, $R = R_e = 2.094a_0$; active space ($2\sigma_g, 2\sigma_u, 3\sigma_g, 1\pi_u, 1\pi_g, 3\sigma_u$)

Macroiteration	ROF1[a] ΔE/a.u.	UEXP2[a] ΔE/a.u.
1	$-6.1E-2$	$-6.6E-2$
2	$-1.5E-2$	$-3.2E-2$
3	$-1.7E-2$	$-1.8E-2$
4	$-1.9E-2$	$-2.6E-3$
5	$-4.7E-3$	$-9.4E-5$
6	$-6.5E-4$	$-1.0E-6$
7	$-1.4E-4$	
8	$-6.6E-5$	
9	$-7.6E-5$[b]	
10	$-2.4E-7$	
11	$-1.6E-8$	
Final energy/a.u.	-109.125968	-109.126147

[a] Basis set II: Huzinaga (411111/311) + 2*s* (exponents 0.07, 0.025) + 1*p* (exponent 0.05) + 2*d* (exponents 1.2, 0.3)

[b] A DIIS extrapolation step was performed in this macroiteration

macroiterations. As discussed in the previous section, one macroiteration in our first-order scheme is faster than in all second-order methods since the evaluation of second derivatives is completely avoided.

The CAS7 (seven active orbitals) calculation (Table 3) can be compared with a similar first-order calculation by Roos et al. [12] using the super-CI technique. In both schemes the convergence pattern is slightly irregular due to the normalization condition (4). However, a reliable and reasonably fast convergence is achieved in 10 macroiterations.

Enlargement of the active space does not change the convergence properties of our method as shown for the CAS8 calculation of N_2 $X^1\Sigma_g^+$ in Table 4. A CAS10 calculation for the ground state of N_2 (active space: $2\sigma_g, 2\sigma_u, 3\sigma_g, 1\pi_u, 1\pi_g, 3\sigma_u, 4\sigma_g, 4\sigma_u$) also converged to the same level of accuracy within 10 macroiterations.

In Table 5 we present CASSCF results for various electronic states of $N_2(X^1\Sigma_g^+, A^3\Sigma_u^+, B^3\Pi_g, C^3\Pi_u)$ at their respective equilibrium distances in order to show how dissociation energies D_e and excitation energies T_e improve when the active space is augmented. For the calculation of the dissociation energies, the ^4S ground state and the first excited ^2D state of the N atom are needed. The active spaces for the CAS6 and CAS8 calculations contain only the 2*s* and 2*p* valence orbitals on N, so they dissociate into "SCF atoms". The SCF energies are for basis set I (see Table 2): $E_{\mathrm{SCF}}(^4S) = -54.394540$ a.u., $E_{\mathrm{SCF}}(^2D) = -54.290042$ a.u., ($\Delta E = 2.84$ eV); and for basis set II (see Table 4): $E_{\mathrm{SCF}}(^4S) = -54.395644$ a.u., $E_{\mathrm{SCF}}(^2D) = -54.291485$ a.u., ($\Delta E = 2.83$ eV). The CAS10 calculation needs a CAS5 calculation (2*s*, 2*p*, 3*s*) for the N atom; the energies for basis set II are: $E_{\mathrm{CAS5}}(^4S) = -54.403323$ a.u., $E_{\mathrm{CAS5}}(^2D) = -54.295335$ a.u., ($\Delta E = 2.94$ eV, experimental $\Delta E = 2.38$ eV [32]).

Table 5. Comparison of CASSCF results with SCF, PNO-CI, CEPA and experimental results for various electronic states of N_2

State	$X^1\Sigma_g^+$	$A^3\Sigma_u^+$	$B^3\Pi_g$	$C^3\Pi_u$
R_e/a_0	2.094	2.470	2.324	2.171
E(SCF)/a.u.[a]	−108.964168		−108.681451	−108.511616
T_e/eV			7.69	12.31
D_e/eV	4.76		−0.09	−4.71
E(SCF)/a.u.[b]	−108.974579		−108.686038	−108.521180
T_e/eV			7.85	12.34
D_e/eV	4.99		−0.03	−4.52
E(CAS6)/a.u.[a]	−109.103792	−108.871532	−108.762377	
T_e/eV		6.32	9.29	
D_e/eV	8.56	2.24	2.12	
E(CAS6)/a.u.[b]	−109.113508	−108.880123	−108.766961	
T_e/eV		6.35	9.43	
D_e/eV	8.77	2.42	2.17	
E(CAS8)/a.u.[a]	−109.116397	−108.882557	−108.824522	−108.683709
T_e/eV		6.36	7.94	11.77
D_e/eV	8.91	2.54	3.81	−0.02
E(CAS8)/a.u.[b]	−109.125968	−108.891123	−108.828690	−108.690510
T_e/eV		6.39	8.09	11.85
D_e/eV	9.11	2.72	3.85	0.09
E(CAS10)/a.u.[b]	−109.162109	−108.921932	−108.851202	−108.715201
T_e/eV		6.54	8.46	12.61
D_e/eV	9.67	3.14	4.15	0.45
E(PNOCI)/a.u.[a]	−109.267410		−108.992183	−108.839411
T_e/eV			7.49	11.65
E(PNOCI)/a.u.[b]	−109.291185		−109.011847	−108.863125
T_e/eV			7.60	11.65
E(CEPA)/a.u.[a]	−109.298878		−109.030665	−108.892240
T_e/eV			7.30	11.07
E(CEPA)/a.u.[b]	−109.326138		−109.054731	−108.921119
T_e/eV			7.38	11.02
T_e/eV exp[c]		6.22	7.39	11.05
D_e/eV exp	9.91	3.68	4.90	1.24

[a] Basis set I, see Table 2
[b] Basis set II, see Table 4
[c] [34]

As it has also been shown by other authors [19] the convergence of the CASSCF results when the active space is enlarged to a "full correlation limit" is very poor. Although the dissociation energies of the CASSCF calculations improve with larger active spaces in the N_2 example, the excitation energies show an irregular and almost unpredictable behaviour; for CAS10 they are worse than for CAS8 for all states and even worse than at SCF level for $B^3\Pi_g$ and $C^3\Pi_u$. The reason is that the fraction of the correlation energy contained in the same active space may be quite different for different electronic states. Note that this phenomenon also occurs for the calculation of the $^4S \rightarrow {}^2D$ excitation energy of the N atom.

For the states $X^1\Sigma_g^+$, $B^3\Pi_g$, and $C^3\Pi_u$ which can be described reasonably well by an SCF reference wavefunction containing only one Slater determinant we have also performed PNO-CI and CEPA calculations [33]. Table 5 shows that the CEPA excitation energies are much superior to the CASSCF results though the computational effort is smaller than for the CAS calculations. (The almost perfect agreement between the CEPA and experimental T_e values for $B^3\Pi_g$ and $C^3\Pi_u$ might be partly a lucky coincidence.) $A^3\Sigma_u^+$ cannot be treated with our CEPA program since the $\pi^3\pi^1\,{}^3\Sigma_u^+$ state is a one-configuration two-determinant reference which is not included in our CEPA program [33].

It seems unfair to compare the quality of CASSCF and CEPA calculations since the two methods are designed for two different purposes: CEPA calculates dynamical correlation energies while CASSCF is applied to describe quasi-degeneracy, i.e. non-dynamical effects. However, in practice it is very hard to distinguish between these two sources of electron correlation. Of course, proper dissociation can only be treated with an MCSCF method. But it seems impossible to understand—prior to extensive calculations and an analysis of the wavefunction—why CASSCF works so well for the excitation energy for the A state of N_2, but fails for the B and C states, even though all four states are one-configuration reference states. Why is the "dynamical" correlation similar in the X and A states but different from that in B and C? We have therefore taken the more pragmatical point of view of applying different methods and looking at their results, and to abandon the badly defined distinction between dynamical and structural correlation effects.

As a third less trivial benchmark test we have treated the two lowest Σ states of the AlO molecule for which calculations with several other MCSCF programs have been published [35, 36]. At its equilibrium distance ($R_e = 3.06a_0$) AlO has a $X^2\Sigma^+$ ground state, with the leading configuration $1\sigma^2 2\sigma^2 3\sigma^2 1\pi^4 4\sigma^2 5\sigma^2 2\pi^4 6\sigma^2 7\sigma^1$ and a low lying $B^2\Sigma^+$ state, with the leading configuration $1\sigma^2 2\sigma^2 3\sigma^2 1\pi^4 4\sigma^2 5\sigma^2 2\pi^4 6\sigma^1 7\sigma^2$, which differs from the ground state by a single excitation. Table 6a gives the results of a CAS7 calculation (active space $2\pi 6\sigma 7\sigma 3\pi 8\sigma$) for both states at an internuclear distance of $R = 3.00a_0$, Table 6b is for the corresponding CAS8 calculation (active space $5\sigma 2\pi 6\sigma 7\sigma 3\pi 8\sigma$). The basis set was taken from [35] with a contraction scheme for Al of 6, 3, 4 × 1/4, 2, 3 × 1/2 × 1 and for O of 6, 2, 3 × 1/4, 3 × 1/2 × 1. As in [36] the ground state calculations have been started with an SCF wavefunction of the $AlO^- X^1\Sigma^+$ ion. The excited state calculation starts with the ground state calculation. A maximum number of 20 microiterations per macroiteration was sufficient to achieve the convergence of Tables 6a, b.

The CAS7 ground state calculation can be compared with second-order calculations of Lengsfield and Liu [35] as well as with recent first-order calculations reported by Yamamoto et al. [36] performed with their super-CI-MCSCF program named JASON2 [36, 37]. JASON2 converges ($|\Delta E| < 10^{-5}$ a.u.) after 27 macroiterations, while RFO1 needs only 11 macroiterations for convergence ($|\Delta E| < 10^{-8}$ a.u.). The second-order Newton–Raphson type MCSCF program of GAMESS which is used for comparison by Yamamoto [36, 38] needs 9

Table 6. CASSCF calculations for AlO, $X^2\Sigma^+$, $B^2\Sigma^+$, $R = 3.00a_0$

Macroiteration	a. Active space ($2\pi, 6\sigma, 7\sigma, 3\pi, 8\sigma$)[a]		b. Active space ($5\sigma, 2\pi, 6\sigma, 7\sigma, 3\pi, 8\sigma$)[a]	
	$X^2\Sigma^+$ ΔE/a.u.	$B^2\Sigma^+$ ΔE/a.u.	$X^2\Sigma^+$ ΔE/a.u.	$B^2\Sigma^+$ ΔE/a.u.
1	$-1.2E-1$	$-5.9E-2$	$-1.4E-1$	$-2.5E-2$
2	$-1.6E-2$	$-4.8E-3$	$-6.6E-3$	$-2.1E-3$
3	$-2.8E-3$	$-1.6E-3$	$-4.9E-4$	$-4.4E-4$
4	$-6.6E-4$	$-5.7E-4$	$-9.4E-5$	$-1.2E-4$
5	$-2.2E-4$	$-3.2E-4$	$-2.5E-5$	$-3.8E-5$
6	$-8.2E-5$	$-1.6E-4$	$-7.6E-6$	$-1.3E-5$
7	$-3.4E-5$	$-8.5E-5$	$-2.6E-6$	$-7.4E-6$[b]
8	$-1.6E-5$	$-4.4E-5$	$-1.4E-6$[b]	$-1.1E-7$
9	$-1.7E-5$[b]	$-2.2E-5$	$-1.4E-8$	$-2.2E-8$
10	$-4.9E-8$	$-2.2E-5$[b]	$-2.9E-9$	$-7.5E-9$
11	$-4.4E-9$	$-2.2E-7$		
12		$-3.0E-8$		
13		$-1.2E-8$		
14		$-6.2E-9$		
Maximum off-diagonal B-matrix element/a.u.	$3.8E-5$	$4.0E-5$	$2.5E-5$	$5.0E-5$
Final energy/a.u.	-316.865086	-316.738570	-316.868893	-316.749979

[a] Basis set see text

[b] A DIIS extrapolation step was performed in this macroiteration

macroiterations ($|\Delta E| < 10^{-5}$ a.u.). Lengsfield and Liu [35] do not report the convergence behaviour of their calculations.

The CPU times for our CAS7 $X^2\Sigma^+$ calculation are listed in Table 7 and compared with the CPU times which are reported in [36] for the JASON2 calculations. Note that:

i) JASON2 uses a direct CI algorithm with CSF's, RFO1 a conventional CI with Slater determinants.

ii) The 4-index transformation is programmed proportional to MN^4 in the JASON2 program, in the RFO1 program it is a fast M^2N^4 (M = number of active orbitals, N = number of basis functions) step avoiding I/O.

iii) The expenditure for the microiterations in RFO1 is compared with building up and diagonalizing the SX-matrix of the super-CI formalism [12, 36].

iv) The total time for RFO1 contains 11 orbital optimizations (11 macroiterations), 12 integral transformations and 12 CI diagonalizations. During the

Table 7. Comparison of CPU times (s) for the programs RFO1 and JASON2 for a CAS7 calculation of AlO $X^2\Sigma^+$

	CPU times in s	
	ROF1	JASON2
Integral transformation	3.8	4.3
CI diagonalization	3.7[a]	11.4[b]
Microiterations/SX-Matrix[c]	7.5	4.3
Number of macroiterations	11	27
Total time	172.7	540.0
	Performance in MFLOPS[d]	
	CYBER 205	HITAC S810/10
Peak	200[e]	315[f]
Average	25[g]	50[g]

[a] CI dimension 1225 SDs, symmetry selection (C_{2v}) reduces the dimension to 321 SDs and the diagonalization time to 0.63s.

[b] CI dimension 784 CSFs

[c] Explanation see text and [12, 36]

[d] MFLOPS = 10^6 floating point operations per s

[e] [39]

[f] [36]

[g] Estimated according to Ohno [40]

calculation the number of microiterations per macroiteration decreases as well as the CI diagonalization time. Average values for one macroiteration are quoted in Table 7.

Finally the program package RFO1 is much more compact than JASON2 or UEXP2 (RFO1: 9500 FORTRAN statements, JASON2: 44000 statements [36], UEXP2: 32000 statements [26]).

5. Conclusions

In this paper we have presented a linearly convergent (first-order) CASSCF program based on the Fock-operator method. Its main characteristics are the following:

1. Convergence rate comparable to single configuration SCF algorithms in 8–12 macroiterations.

2. Simple formalism for the evaluation of the Fock matrix and extremely simple microiterative equations.

3. Renormalization of the Fock matrix elements with the occupation numbers of the active orbitals.

4. Efficient application of the approximate second-order convergence acceleration of Pulay.

Feature 2–4 make our formalism competitive to second-order methods with respect to overall computational effort.

Acknowledgements. The authors thank F. Schmitz, Bochum, who carried out the calculations with R. Shepard's program, and Prof. R. Ahlrichs, Karlsruhe, who provided us with a copy of this program. All calculations were performed on the CYBER 205 vectorcomputer of the 'Rechenzentrum der Ruhr-Universität Bochum'. This work was financially supported by the Sonderforschungsbereich 42 at Wuppertal.

References

1. Olsen J, Yeager DL, Jørgensen P (1983) Optimization and characterization of a multiconfigurational self-consistent field (MCSCF) state. Adv Chem Phys 54:1
2. Werner HJ (1987) Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration-interaction methods. Adv Chem Phys 69:1
3. Shepard R (1987) The multiconfiguration self-consistent field method. Adv Chem Phys 69:63
4. Roos BO (1987) The complete active space self-consistent field method and its applications in electronic structure calculations. Adv Chem Phys 69:399
5. Almlöf J, Roos B, Siegbahn P (1978) Comput Chem 2:89
6. Hamilton TP, Pulay P (1986) J Chem Phys 84:5728
7. Fernandez Rico J, Paniagua M, Garcia De La Vega JM, Fernandez-Alonso JI, Fantucci P (1986) J Comp Chem 7:201
8. Wasilewski J, Staemmler V (1988) Acta Phys Pol A74:355
9. Hinze J (1973) J Chem Phys 59:6424
10. Hinze J, Yurtsever E (1979) J Chem Phys 70:3188
11. Roos BO (1980) Int J Quantum Chem Symp 4:175
12. Roos BO, Taylor PR, Siegbahn PEM (1980) Chem Phys 48:157
13. Siegbahn PEM, Almlöf J, Heiberg A, Roos BO (1981) J Chem Phys 74:2384
14. Lévy B, Berthier G (1968) Int J Quantum Chem 2:307
15. Staemmler V (1977) Theor Chim Acta 45:89
16. McWeeny R (1964) In: Löwdin PO, Pullman A (eds) Molecular orbitals in chemistry, physics and biology. Academic Press, New York, p 305
17. Werner HJ, Meyer W (1980) J Chem Phys 73:2342
18. Werner HJ, Meyer W (1981) J Chem Phys 74:5794
19. Knowles P, Werner HJ (1985) Chem Phys Lett 115:259
20. Werner HJ, Knowles P (1985) J Chem Phys 82:5053
21. Shepard R, Simons J (1980) Int J Quantum Chem Symp 14:211
22. Shepard R, Shavitt I, Simons J (1982) J Chem Phys 76:543
23. Brown FB, Shavitt I, Shepard R (1984) Chem Phys Lett 105:363
24. Pulay P (1980) Chem Phys Lett 73:393
25. Pulay P (1982) J Comp Chem 3:556
26. Schmitz F (1988) Diplom thesis, Bochum
27. Ahlrichs R (1974) Theor Chim Acta 33:157
28. Kutzelnigg W, Schindler M, Klopper W, Meier U, Wallmeier H (1986) Experiences with the CYBER 205 for quantum chemical calculations. Lect Notes Chem, vol 44. Springer, Berlin Heidelberg New York, p 55

29. Dupuis M, Rys J, King HF (1976) J Chem Phys 65:111
30. Ahlrichs R, Böhm HJ, Ehrhardt C, Scharf P, Schiffer H, Lischka H, Schindler M (1985) J Comp Chem 6:200
31. Jørgensen P, Swanstrøm P, Yeager DL (1983) J Chem Phys 78:347
32. Moore CE (1971) Atomic energy levels, Nat Bur Stand (US) Circular 467, vol 1. Washington
33. Staemmler V, Jaquet R (1981) Theor Chim Acta 59:487
34. Huber KP, Herzberg G (1979) Constants of diatomic molecules. Van Nostrand Reinhold, New York
35. Lengsfield III BH, Liu B (1982) J Chem Phys 77:6083
36. Yamamoto S, Nagashima U, Aoyama T, Kashiwagi H (1988) J Comp Chem 9:627
37. Yamamoto S, Teraoka J, Kashiwagi H (1988) J Chem Phys 88:303
38. Dupuis M, Spangler D, Wendoloski JJ (1980) NRCC Software Catalog, vol 1, Program RGO1 (GAMESS)
39. Hockney RW, Jesshope CR (1981) Parallel computers. Hilger, Bristol, p 124
40. Ohno K (1986) Japanese supercomputers and molecular orbital calculations. Lect Notes Chem, vol 44. Springer, Berlin Heidelberg New York, p 50