Physica Scripta

A Comparison of the Super-CI and the Newton-Raphson Scheme in the Complete Active Space SCF Method

To cite this article: Per Siegbahn *et al* 1980 *Phys. Scr.* **21** 323

View the article online for updates and enhancements.

Related content

- The relativistic many body problem in molecular theory
  Werner Kutzelnigg
- Quantum theory of molecular electronic structure
  R McWeeny and B T Pickup
- Ab-initio Calculated Infrared Transition Probabilities, Dipole Moments and Spectroscopic Constants for the $X^2_^+$ State of BeH
  Mats Larsson

Recent citations

- Pseudotetrahedral cobalt(ii) complexes with PNP-ligands showing uniaxial magnetic anisotropy
  Yuan-Qi Zhai *et al*
- Large Scale Electron Correlation Calculations: Rank-Reduced Full Configuration Interaction
  B. Scott Fales *et al*
- Origin of the green persistent luminescence of Eu-doped SrAl2O4 from a multiconfigurational ab initio study of 4f7 4f65d1 transitions
  Lixin Ning *et al*

This content was downloaded from IP address 130.64.11.153 on 01/07/2018 at 20:02

Physica Scripta. Vol. 21, 323–327, 1980

A Comparison of the Super-CI and the Newton-Raphson Scheme in the Complete Active Space SCF Method

Per Siegbahn and Anders Heiberg*

Institute of Theoretical Physics, Vanadisvägen 9, S-113 46 Stockholm, Sweden

Björn Roos

Department of Physical Chemistry 2, Chemical Centre, S-220 07 Lund, Sweden

and

Bernard Levy

Laboratoire de Chimie de l'ENSJF, 1 rue Maurice Arnoux, 92 120 Montrouge, France

Received June 11, 1979

Abstract

A comparison of the super-C I and the Newton–Raphson scheme in the complete active space SCF method. Per Siegbahn and Anders Heiberg (Institute of Theoretical Physics, Vanadisvägen 9, S-113 46 Stockholm, Sweden); Björn Roos (Department of Physical Chemistry 2, Chemical Centre, S-220 07 Lund, Sweden); and Bernard Levy (Laboratoire de Chimie de l'ENSJF, 1 rue Maurice Arnoux, 92 120 Montrouge, France). *Physica Scripta (Sweden) 21*, 323–237, 1980.

A density matrix formulation is presented of the super-C I and Newton–Raphson methods in complete active space SCF (CASSCF) calculations. The CASSCF method is a special form of the MC–SCF method, where the C I wave function is assumed to be complete in a subset of the orbital space (the active space), leaving the remaining orbitals doubly occupied in all configurations.

Explicit formulas are given for all matrix elements in the super-C I method and the first and second derivatives in the Newton–Raphson formulation. The similarities between the two methods are pointed out and the differences in the detailed formulations are discussed. Especially interesting is the fact, that while the second derivatives can be expressed in terms of first and second order density matrices, the matrix elements between the super-C I states involve also the third order density matrix in some cases.

1. Introduction

Recent progress in computational quantum chemistry, both concerning models for treating chemical problems and methods for solving these problems, has led to a renewed interest in the MC–SCF method. Whereas it has earlier always been emphasized that the power of the MC–SCF method is its ability to describe the electronic structure of an atomic or molecular system with only a small number of configurations, we will here rather stress the use of only a few orbitals with variable occupation numbers. Obviously the distribution of a small number of electrons among few orbitals does not necessarily lead to a short C I expansion. Eight electrons in eight orbitals, as an example, gives rise to 1764 singlet coupled configurations if there is no spatial symmetry reduction. In conventional MC–SCF procedures one selects from this long list the few which are assumed to be the most important. This selection is however often difficult to make, e.g., when studying a more complicated chemical reaction. In the present paper a method will be presented where this selection is not necessary; all configurations will be included in the C I expansion, the only remaining problem being the selection of the orbitals. As a consequence the method must be capable of treating long expansions in an efficient way.

A convenient way of handling long expansions in the MC–SCF method is in terms of reduced density matrices [1–3] and this formulation will be used here. An iteration in an MC–SCF calculation can be divided into two different parts. In the first part the C I-coefficients are obtained and in the second part the orbital optimization is performed. In most MC–SCF calculations performed to this date the first part is done in only a small fraction of the time spent in the second part. The advantage of writing the information from the C I-calculation in terms of reduced density matrices is that the reference to the individual C I-coefficients is avoided in the orbital optimization section. This means that a much longer C I expansion can be used, within a fixed orbital space, without increasing the time spent in the dominating orbital part. The use of a C I-expansion, which is complete in a subspace of the orbital space, may actually speed up the orbital section, since orbital rotations which leave the energy invariant can be removed. This probably also leads to a more rapid convergence of the orbitals.

In a series of papers different MC–SCF methods using the density matrix formalism will be discussed, for the case of a C I-expansion which is complete in part of the orbital space. The term active will be used for these orbitals. The remainder of the occupied orbital space will be referred to as inactive. These orbitals are doubly occupied in all configurations. The active and inactive orbitals constitute the primary orbital space. The rest of the orbital space, which is not occupied in the C I wave function is termed secondary.

In the present paper we will be mainly concerned with the formal aspects of the MC–SCF method. In particular two seemingly quite different approaches to the orbital optimization part will be discussed, the Newton–Raphson approach [3–5] and the super-C I method [6]. Explicit formulas for the relevant derivatives and matrix elements will be derived. It will be shown that the two approaches are indeed very closely related, but that they differ in significant details. The most important difference is the appearance of third order density

* Permanent address: Institute of Chemistry, University of Oslo, Blindern, Oslo 3, Norway.

matrix elements in the super-CI method, which do not appear in the Newton–Raphson approach. For a recent review and discussion of the MC–SCF method see [19].

2. The super-CI approach

In order to derive the matrix elements appearing in the super-CI approach it is convenient to adopt the unitary group notation [7, 8]. A generator of the unitary group for the spatial orbitals φ_p and φ_q is defined as a spin product of creation and annihilation operators,

$$\hat{E}_{pq} = \sum_{\sigma} a^{+}_{p\sigma} a_{q\sigma} \tag{1}$$

The spin-independent Hamiltonian can be expressed in terms of these generators as

$$\hat{H} = \sum_{p,q} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{p,q,r,s} (pq|rs)[\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}] \tag{2}$$

where the one-and two-electron integrals have their usual definitions,

$$h_{pq} = \langle \varphi_p|\hat{h}|\varphi_q\rangle \tag{3a}$$

$$(pq|rs) = \langle \varphi_p(1)\varphi_r(2)|1/r_{12}|\varphi_q(1)\varphi_s(2)\rangle \tag{3b}$$

The first and second-order reduced density matrices for a CI wave function

$$|0\rangle = \sum_K A_K|K\rangle$$

can consequently be written in the form

$$D_{pq} = \langle 0|\hat{E}_{pq}|0\rangle \tag{4a}$$

$$P_{pqrs} = \tfrac{1}{2}\langle 0|(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps})|0\rangle \tag{4b}$$

The generators of the unitary group further fulfil a commutation relation, which will be extensively used in the derivations of the super-C I matrix elements:

$$\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps} = \hat{E}_{rs}\hat{E}_{pq} - \delta_{ps}\hat{E}_{rq} \tag{5}$$

As discussed in the introduction the C I expansion $|0\rangle$ treated in this paper is assumed to be complete in the small active orbital space. the letters $t, u, v, \ldots$ will be used to denote active orbitals, whereas the inactive orbitals will be labelled $i, j, k, \ldots$, and orbitals in the secondary space $a, b, c, \ldots$. Indices $p, q, r, \ldots$ will be used when the subspace is not specified.

Clearly density matrix elements with any of its indices in the secondary subspace are zero. The following relations also hold:

$$D_{ip} = 2\delta_{ip} \tag{6a}$$

$$P_{ijpq} = \delta_{ij}D_{pq} - \delta_{iq}\delta_{jp} \tag{6b}$$

$$P_{pijq} = 2\delta_{ip}\delta_{jq} - \tfrac{1}{2}\delta_{ij}D_{pq} \tag{6c}$$

If these relationships are used there is no reason to explicitly keep and store any density matrix elements with inactive indices. This is an important result with the consequence that, if the number of active orbitals is small which is usually the case, the list of necessary first and second order density matrix elements can easily be kept in main storage during the calculation. This was otherwise considered to be one of the major problems in connection with the density matrix formulated MC–SCF method suggested by Kendrick and Hillier [9], which will be further discussed later.

With these preliminary definitions made the super-C I approach to MC–SCF can now be described. The starting point for this method is the generalized Brillouin theorem [10]:

$$\langle 0|(\hat{E}_{qp} - \hat{E}_{pq})\hat{H}|0\rangle = 0 \tag{7}$$

which is the necessary and sufficient condition for an orbital optimized wave function. In the super-C I method we define (unnormalized) Brillouin states as

$$|pq\rangle = (\hat{E}_{pq} - \hat{E}_{qp})|0\rangle \tag{8}$$

and solve a secular problem with these states as the basis set, to obtain a wave function

$$\Psi_{\mathrm{SCI}} = |0\rangle + \sum_{p,q} X_{pq}|pq\rangle \tag{9}$$

The orbitals are then improved with the contraction formula [6]

$$\varphi_p^{(k)} = \varphi_p^{(k-1)} + \sum_{q\neq p} X_{pq}\varphi_q^{(k-1)} \tag{10}$$

It is clear from the secular equation system that the generalized Brillouin theorem is fulfilled at convergence, that is, when all coefficients X_{pq} are zero. The contraction formula (10) for the optimization of the orbitals is a natural way to pack the information from the total wave function (9) into the starting function $|0\rangle$. This was the rather intuitive way the method was first suggested and only more recently has the close connection to the Newton–Raphson procedure been emphasized [6]. An alternative approach to the orbital improvement numbers as the new orbitals [11]. In this way the largest overlap between the super C I wave function (9) and the improved $|0\rangle$ is ensured.

Since the inactive orbitals are doubly occupied in all configurations comprising $|0\rangle$ and further the expansion is complete in the active subspace, the energy will be invariant towards rotations within these two subspaces. Thus only three types of Brillouin states $|pq\rangle$ have to be considered namely, $|ai\rangle$, $|at\rangle$ and $|ti\rangle$. Consequently only nine types of matrix, element can be written as

$$\langle pq|\hat{H}|rs\rangle = \langle 0|(\hat{E}_{qp} - \hat{E}_{pq})\hat{H}(\hat{E}_{rs} - \hat{E}_{sr}|0\rangle \tag{11}$$

where actually only one of the generators in each of the Brillouin states gives a non-zero contribution, due to the limitations of the pq values.

Writing the Hamiltonian operator in the form (2) and using the commutation relation (5) and the definitions (4) and (6) of the density matrix elements it is straightforward, though rather tedious, to derive expressions for these matrix elements. Introducing also a third order density matrix with elements

$$Q_{vxyztu} = \tfrac{1}{2}\langle 0|(\hat{E}_{vx}\hat{E}_{yz} - \delta_{xy}\hat{E}_{vz})\hat{E}_{tu}|0\rangle \tag{12}$$

and two Fock type operators as

$$F^{\mathrm{I}}_{pq} = h_{pq} + \sum_k [2[(pq|kk) - \tfrac{1}{2}(pk|qk)] \tag{13a}$$

$$F^{\mathrm{A}}_{pq} = \sum_{v,x} D_{vx}[(pq|vx) - \tfrac{1}{2}(pv|qx)] \tag{13b}$$

where I stands for inactive and A for active, the final results for the matrix elements can be written as

$$\langle 0|\hat{H}\hat{E}_{ti}|0\rangle = 2(F^{\mathrm{I}}_{ti} + F^{\mathrm{A}}_{ti}) - \sum_{v} D_{tv}F^{\mathrm{I}}_{iv} - 2\sum_{v,x,y} P_{tvxy}(iv|xy) \tag{14a}$$

$$\langle 0|\hat{H}\hat{E}_{ai}|0\rangle = 2(F^{\mathrm{I}}_{ai} + F^{\mathrm{A}}_{ai}) \tag{14b}$$

$$\langle 0|\hat{H}\hat{E}_{at}|0\rangle = \sum_{v} D_{tv}F^{\mathrm{I}}_{av} + 2\sum_{v,x,y} P_{tvxy}(av|xy) \tag{14c}$$

$$\begin{aligned}\langle 0|\hat{E}_{it}\hat{H}'\hat{E}_{uj}|0\rangle = {} & 2\delta_{ij}F^{\mathrm{A}}_{tu} - 2\delta_{ij}\sum_{v,x,y} P_{vxty}(vx|uy) \\ & - \sum_{v}(\delta_{uv} - D_{uv})[(vt|ij) - 2(vj|ti)] \\ & - \sum_{v}(\delta_{tv} - D_{tv})[(vu|ij) - 2(vi|uj)] \\ & - \sum_{v,x}\{2[\delta_{tu}D_{vx} - P_{tuvx}](vx|ij) \\ & - [\delta_{tu}D_{vx} - 2P_{vutx}](vj|xi)\} \\ & - (2\delta_{tu} - D_{tu})F^{\mathrm{I}}_{ij} + \delta_{ij}\sum_{v}[2\delta_{tv}F^{\mathrm{I}}_{uv} \\ & - D_{tv}F^{\mathrm{I}}_{uv} - D_{uv}F^{\mathrm{I}}_{tv}] + \delta_{ij}\sum_{v,x}[D_{tu}D_{vx} \\ & - 2P_{tuvx}]F^{\mathrm{I}}_{vx} + \delta_{ij} \\ & \sum_{v,x,y,z}(P_{vxyz}D_{tu} - Q_{vxyztu})(vx|yz)\end{aligned} \tag{14d}$$

$$\begin{aligned}\langle 0|\hat{E}_{it}\hat{H}\hat{E}_{aj}|0\rangle = {} & \sum_{v} D_{tv}[(av|ij) - 2(aj|vi)] \\ & - 2\delta_{ij}\sum_{v,x,y} P_{tvxy}(av|xy) + 2\delta_{ij}(F^{\mathrm{I}}_{at} \\ & + F^{\mathrm{A}}_{at}) - \delta_{ij}\sum_{v} D_{tv}F^{\mathrm{I}}_{av} - 2[(at|ij) \\ & - 2(aj|ti)]\end{aligned} \tag{14e}$$

$$\begin{aligned}\langle 0|\hat{E}_{it}\hat{H}\hat{E}_{au}|0\rangle = {} & -2\sum_{v,x} P_{tvux}(ax|vi) - \sum_{v} D_{uv}[(at|vi) \\ & - 2(av|ti)]\end{aligned} \tag{14f}$$

$$\begin{aligned}\langle 0|\hat{E}_{ia}\hat{H}'\hat{E}_{bj}|0\rangle = {} & 4(ai|bj) - 2(ab|ij) + 2\delta_{ij}(F^{\mathrm{I}}_{ab} + F^{\mathrm{A}}_{ab}) \\ & - 2\delta_{ab}(F^{\mathrm{I}}_{ij} + F^{\mathrm{A}}_{ij})\end{aligned} \tag{14g}$$

$$\begin{aligned}\langle 0|\hat{E}_{ia}\hat{H}\hat{E}_{bt}|0\rangle = {} & \sum_{v} D_{tv}[2(ai|bv) - (ab|iv)] \\ & - 2\delta_{ab}\sum_{v,x,y} P_{tvxy}(vi|xy) - \delta_{ab}\sum_{v} D_{tv}F^{\mathrm{I}}_{vi}\end{aligned} \tag{14h}$$

$$\begin{aligned}\langle 0|\hat{E}_{ta}\hat{H}'\hat{E}_{bu}|0\rangle = {} & \sum_{v,x}[2P_{tuvx}(ab|vx) + 2P_{txvu}(ax|bv)] \\ & + D_{tu}F^{\mathrm{I}}_{ab} + \delta_{ab}\sum_{v,x}[2P_{tuvx} - D_{tu}D_{vx}]F^{\mathrm{I}}_{vx} \\ & + \delta_{ab}\sum_{v,x,y,z}[Q_{vxyztu} - D_{tu}P_{vxyz}](vx|yz) \\ & - 2\delta_{ab}\sum_{v,x,y} P_{uvxy}(tv|xy)\end{aligned} \tag{14i}$$

where

$$\hat{H}' = \hat{H} - \langle 0|\hat{H}|0\rangle$$

The calculation of these matrix elements is straightforward, provided that the density matrices can be kept in the main storage. This is normally not possible for the third order density matrices, however. A possibility, which will be further discussed in a forthcoming paper [12], is to approximate the expressions for the matrix elements between the Brillouin states. Apart from the problem with the third order density matrices, the matrix elements (14) are easily evaluated. The two-electron integrals are read sequentially and the contributions added to a core load of matrix elements. The number of necessary reads of the integral list is equal to the number of core loads needed to create the complete matrix. An alternative approach is to solve the secular problem using the direct method, that is, evaluating the vector **HX** directly from the integrals [13]. In any case the calculation is easy, since the size of the super-C I problem is very limited, in most cases only a few hundred terms.

It should be emphasized again that the calculation is independent of the size of the CAS wave function $|0\rangle$. In the conventional use of the super-C I method this is not the case. Here one starts by generating all configurations singly excited with respect to all configurations comprising $|0\rangle$. There can be a large number of such singly excited states. The matrix elements between these states are then generated and in the last step a contraction is made to yield matrix elements between the Brillouin states. Obviously this method is more time consuming than the density matrix formulation, and also severely limits the possible size of the CAS wave function. A more detailed account of the density matrix formulated super-C I method with applications to energy surfaces for chemical reactions will be presented elsewhere [12].

3. The Newton–Raphson approach

In the orbital optimization part of the Newton–Raphson approach we look for the unitary transformation **U** of the initial orbital set which minimizes the energy. This unitary transformation can be parameterized as [3],

$$\mathbf{U} = \exp(\mathbf{X}) = \mathbf{1} + \mathbf{X} + \tfrac{1}{2}\mathbf{X}^2 + \ldots \tag{15}$$

where **X** is a skew-symmetric matrix ($\mathbf{X}^{\dagger} = -\mathbf{X}$). Taking the derivative of the Taylor expansion of the energy up to second order in the matrix elements X_{pq} leads to the set of linear equations,

$$\left(\frac{\partial E}{\partial X_{pq}}\right)_0 + \sum_{r<s}\left(\frac{\partial^2 E}{\partial X_{pq}\partial X_{rs}}\right)_0 X_{rs} = 0 \tag{16}$$

If the equation system (16) is solved exactly we have by definition a quadratically convergent method.

The simplest way of deriving the expressions for the first and second derivatives appearing in (16) is by first writing the energy in terms of density matrices as

$$E = \sum_{p,q} D_{pq}h_{pq} + \sum_{p,q,r,s} P_{pqrs}(pq|rs) \tag{17}$$

If the derivatives of the density matrices are neglected, as is usually done and has been shown to be a good approximation [14], the derivation is straightforward. Kendrick and Hillier [9] used this approach with the additional assumption that in the expressions for the second derivatives the first derivatives are zero. This, however, seems to be an unnecessary assumption and

gives no obvious simplification of the final formulas. Based on Fock-like orbital equations Hinze [4] derived similar expressions, where the expression (15) is truncated at first order and the change of the Fock operators is neglected. These approximations also seem unnecessary and were in fact found to affect the rate of convergence in some cases [9]. The method of Das and Wahl [5] finally is closely related to the Hinze approach.

In order to illustrate the difference between the super-C I method and the Newton–Raphson approach it is useful to explicitly write down the expressions for the derivatives in terms of the generators of the unitary group. The derivation of these expressions follows the same lines as used for the first derivatives in [10]. The details for both first and second derivatives can be found in [15], and the results are

$$\left(\frac{\partial E}{\partial X_{pq}}\right)_0 = 2\langle 0|\hat{H}(\hat{E}_{pq} - \hat{E}_{qp})|0\rangle \tag{18}$$

$$\begin{aligned}\left(\frac{\partial^2 E}{\partial X_{pq}\partial X_{rs}}\right)_0 &= 2\langle 0|(\hat{E}_{qp} - \hat{E}_{pq})\hat{H}(\hat{E}_{rs} - \hat{E}_{sr})|0\rangle \\ &+ \langle 0|(\hat{E}_{qp} - \hat{E}_{pq})(\hat{E}_{sr} - \hat{E}_{rs})\hat{H}|0\rangle \\ &+ \langle 0|(\hat{E}_{sr} - \hat{E}_{rs})(\hat{E}_{qp} - \hat{E}_{pq})\hat{H}|0\rangle\end{aligned} \tag{19}$$

We notice that, apart from an unimportant factor of 2, the expression for the first derivative and the first term in the second derivative are exactly equal to the matrix-elements in the super-C I method. The two remaining terms in the second derivative corresponds partly to interactions between the MC state $|0\rangle$ and doubly excited states not included in the super-C I approach [18], but also include gradient type terms which may be equally important.

The explicit expressions in terms of integrals and density matrices for the six different second derivatives can be derived as in Section 2, using the commutation relation (5) between the generators of the unitary group. The same expressions can of course be obtained by taking the derivatives directly of the energy expression (17) and this is probably a simpler method in practice. The results are listed below.

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{it}\partial X_{ju}} &= 2\sum_{v,x}[P_{utvx}(vx|ij) + (P_{uxvt} + P_{uxtv})(vi|xj)] \\ &+ \sum_v \{(\delta_{tv} - D_{tv})[4(vi|uj) - (ui|vj) - (uv|ij)] \\ &+ (\delta_{uv} - D_{uv})[4(vj|ti) - (tj|vi) - (tv|ij)] \\ &+ D_{tu}F^{\mathrm{I}}_{ij} - 2\delta_{ij}\sum_{v,x,y} P_{tvxy}(uv|xy) \\ &+ 2\delta_{ij}(F^{\mathrm{I}}_{tu} + F^{\mathrm{A}}_{tu}) - \delta_{ij}\sum_v D_{uv}F^{\mathrm{I}}_{tv} - 2\delta_{tu}(F^{\mathrm{I}}_{ij} \\ &+ F^{\mathrm{A}}_{ij})\end{aligned} \tag{20a}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{it}\partial X_{ja}} &= \sum_v (2\delta_{tv} - D_{tv})[4(aj|vi) - (av|ij) - (ai|vj)] \\ &- \delta_{ij}\sum_{v,x,y} P_{tvxy}(av|xy) + 2\delta_{ij}(F^{\mathrm{I}}_{at} \\ &+ F^{\mathrm{A}}_{at}) - \tfrac{1}{2}\delta_{ij}\sum_v D_{tv}F^{\mathrm{I}}_{av}\end{aligned} \tag{20b}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{it}\partial X_{ua}} &= -2\sum_{v,x}\{P_{tuvx}(ai|vx) + (P_{tvux} + P_{tvxu})(ax|vi)\} \\ &- \sum_v D_{uv}[(ai|tv) - 4(av|ti) + (at|vi)] - D_{tu}F^{\mathrm{I}}_{ai} \\ &+ \delta_{tu}(F^{\mathrm{I}}_{ai} + F^{\mathrm{A}}_{ai})\end{aligned} \tag{20c}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{ia}\partial X_{jb}} &= 2[4(ai|bj) - (ab|ij) - (aj|bi)] + 2\delta_{ij}(F^{\mathrm{I}}_{ab} + F^{\mathrm{A}}_{ab}) \\ &- 2\delta_{ab}(F^{\mathrm{I}}_{ij} + F^{\mathrm{A}}_{ij})\end{aligned} \tag{20d}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{ia}\partial X_{tb}} &= \sum_v D_{tv}[4(ai|bv) - (av|bi) - (ab|vi)] \\ &- \delta_{ab}\sum_{v,x,y} P_{tvxy}(vi|xy) - \delta_{ab}(F^{\mathrm{I}}_{ti} + F^{\mathrm{A}}_{ti}) \\ &- \tfrac{1}{2}\delta_{ab}\sum_v D_{tv}F^{\mathrm{I}}_{vi}\end{aligned} \tag{20e}$$

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 E}{\partial X_{ta}\partial X_{ub}} &= 2\sum_{v,x}\{P_{tuvx}(ab|vx) + (P_{txvu} + P_{txuv})(ax|bv)\} \\ &+ D_{tu}F^{\mathrm{I}}_{ab} - \delta_{ab}\sum_{v,x,y}\{P_{tvxy}(uv|xy) \\ &+ P_{uvxy}(tv|xy)\} - \tfrac{1}{2}\delta_{ab}\sum_v [D_{tv}F^{\mathrm{I}}_{uv} + D_{uv}F^{\mathrm{I}}_{tv}]\end{aligned} \tag{20f}$$

The above direct expressions for the derivatives in terms of density matrices and integrals have been found to be more useful in practical applications than the expressions in terms of two indexed Fock operators [5, 9] since the storage of these operators is not trivial.

When the first and second derivatives have been constructed according to the formulas (20) the Newton–Raphson equation system is solved for **X** by an iterative Gauss–Seidel method, which has exactly the same convergence criterion as the Newton–Raphson method, namely that the Hessian matrix is positive definite. Orthonormality of the orbitals is finally preserved by using the iterative procedure suggested in [3]. Details of the Newton–Raphson method including examples from applications will be given in a subsequent paper [16].

4. Comparison of the methods

With the formulas given in Sections 2 and 3 we are now in a position to make a detailed comparison between the super-C I and the Newton–Raphson methods. Two aspects will be emphasized, the formal differences in the expressions (14) and (20), and the amount of computational work involved in each iteration of the orbital optimization procedure.

The by far most time-consuming part of an MC–SCF iteration is to our experience the four-index transformation of two-electron integrals. This is at least true if the number of configurations is not more than a few hundred. An important aspect of an MC–SCF method is therefore which two-electron integrals are needed, or more specifically how many of the indices in these integrals belong to the largest orbital subspace, the secondary. Comparing expressions (14) and (20) shows that the two methods are the same in this respect, both requiring integrals with at most two secondary orbitals. It should be noticed, however, that the most important terms, the Brillouin matrix elements, contain integrals with only one index outside the active orbital subspace. Experience shows that the transformation of these integrals can be performed in a small fraction of the time needed for integrals appearing in the other matrix-

elements (or second derivatives). It is therefore interesting to look for approximations of these matrix-elements (second-derivatives) that do not involve other types of integrals. The judgement of the goodness of such approximate methods will be the rate of convergence of the iterative orbital optimization, compared to the calculation with all terms included. Further details on these aspects of the methods will be given in a forthcoming paper [12].

We now turn our attention to the actual orbital optimization step, which can be divided into two parts. In the first part the matrix elements or derivatives are constructed and in the second part the super-C I matrix is diagonalized or alternatively the Newton–Raphson equation system solved. The last part is clearly very similar in the two methods. The problems have the same dimensionality, and a common way to solve an eigenvalue problem iteratively is to treat it as a set of linear equations, as is done for example in the Nesbet method [17]. The only difference is that the diagonalization needs a simultaneous updating of the energy, which is however no problem. For some reason the difficulty to solve the linear equation system in the Newton–Raphson procedure seems to have been overestimated in the literature. Solving for the lowest eigenvalue of a secular problem with a few hundred rows has for a long time been a minor problem requiring only a few seconds on modern computers. Clearly this part will not be the bottleneck for large scale MC–SCF methods.

Looking now at the details of the expressions (14) and (20) significant differences are found. Some of the super-C I matrix element expressions contain third order density matrix elements. It is not realistic to compute and store this very large matrix, and the super-C I method in the density matrix formulation can therefore not be handled without approximations. For the derivatives in the Newton–Raphson scheme there are no such problems and an exact treatment is therefore in principle possible. The Newton–Raphson method is further quadratically convergent (or almost so, since the derivatives of the density matrix elements have been neglected) which does not hold for the super-C I method. The super-C I method can at best be said to be approximately quadratically convergent since some of the matrix elements are rather closely related to second derivatives. The conclusion must therefore be that if it turns out to be feasible to handle the Newton–Raphson procedure exactly it must be the recommended method, at least based on purely theoretical arguments. In practice, of course, the situation can be different. Also, if it turns out that the calculation of the second derivatives according to (18), including the four-index integral transformation, is too time-consuming to tbe practically useful, and approximations have to be introduced, the super-C I method may be an equally efficient choice.

Acknowledgement

One of us (P.S.) would like to express his gratitude to l'ENSJF for financing a visit to Paris, where part of this work was performed. This work has been supported by a grant from the Swedish Natural Science Research Council.

References

1. Löwdin, P. O., Phys. Rev. **97**, 1474 (1955).
2. McWeeny, R., Rev. Mod. Phys. **32**, 335 (1960).
3. Levy, B., Intern. J. Quantum Chem. **4**, 297 (1970).
4. Hinze, J., J. Chem. Phys. **59**, 6424 (1973).
5. Wahl, A. C. and Das, G., in Modern Theoretical Chemistry, Vol. 3, H. F. Schaefer (ed.), Plenum Press, New York, N.Y., (1977); J. Chem. Phys. **56**, 1769 (1972).
6. Grein, F. and Chang, T. C., Chem. Phys. Lett. **12**, 44 (1971); Bannerjee, A. and Grein, F., Intern. J. Quantum Chem. **10**, 123 (1976).
7. Moshinsky, M., Group Theory and the Many-Body Problem, Gordon and Breach, New York, 1968.
8. Paldus, J., in Theoretical Chemistry: Advances and Perspectives, H. Eyring and D. J. Henderson (eds.), Academic Press, New York, 1976, Vol. 2, p. 131.
9. Kendrick, J. and Hillier, I., Chem. Phys. Lett. **41**, 283 (1976).
10. Levy, B. and Berthier, G., Intern. J. Quantum Chem. **2**, 307 (1968).
11. Ruedenberg, K. and Sundberg, K. R., in Quantum Science, Methods and Structure, (eds. J.-L. Calais, O. Goscinski, J. Linderberg and Y. Öhrn), Plenum Press, New York, 1976.
12. Roos, B. O., Siegbahn, P. E. M. and Taylor, P. R. (to be published).
13. Roos, B. O. and Siegbahn, P. E. M., in Modern Theoretical Chemistry 3: Methods of Electronic Structure Theory (ed. Henry F. Schaefer III), Plenum Press, New York, 1977.
14. Levy, B., Chem. Phys. Lett. **18**, 59 (1973); see however also Yeager, D. L. and Yörgensen, P., to be published.
15. Levy, B., thesis, CNRS No. AO 5271, Paris (1971).
16. Heiberg, A., Roos, B. O. and Siegbahn, P. E. M. (to be published).
17. Nesbet, R. K., J. Chem. Phys. **61**, 5231 (1974).
18. Chang, T. C. and Schwarz, W. H. E., Theoret. Chim. Acta **44**, 45 (1977).
19. Dalgaard, E. and Jörgensen, P., J. Chem. Phys. **69**, 3833 (1978).