

## A COMPLETE ACTIVE SPACE SCF METHOD (CASSCF) USING A DENSITY MATRIX FORMULATED SUPER-CI APPROACH

Björn O. ROOS, Peter R. TAYLOR\*

*Department of Physical Chemistry 2, Chemical Centre, S-220 07 Lund 7, Sweden*

and

Per E.M. SIEGBAHN

*Institute of Theoretical Physics, S-113 46 Stockholm, Sweden*

Received 19 November 1979

A density matrix formulation of the super-CI MCSCF method is presented. The MC expansion is assumed to be complete in an *active* subset of the orbital space, and the corresponding CI secular problem is solved by a direct scheme using the unitary group approach. With a density matrix formulation the orbital optimization step becomes independent of the size of the CI expansion. It is possible to formulate the super-CI in terms of density matrices defined only in the small active subspace; the doubly occupied orbitals (the inactive subspace) do not enter. Further, in the unitary group formalism it is straightforward and simple to obtain the necessary density matrices from the symbolic formula list. It then becomes possible to treat very long MC expansions, the largest so far comprising 726 configurations. The method is demonstrated in a calculation of the potential curves for the three lowest states ( ${}^1\Sigma_g^+$ ,  ${}^3\Sigma_g^+$  and  ${}^3\Pi_g^-$ ) of the  $N_2$  molecule, using a medium-sized gaussian basis set. Seven active orbitals were used yielding the following results:  $D_e$ : 8.76 (9.90), 2.43 (3.68) and 3.39 (4.90) eV;  $r_e$ : 1.108 (1.098), 1.309 (1.287) and 1.230 (1.213) Å;  $\omega_e$ : 2333 (2359), 1385 (1461) and 1680 (1733)  $\text{cm}^{-1}$ , for the three states (experimental values within parentheses). The results of these calculations indicate that it is important to consider not only the dissociation limit but also the united atom limit in partitioning the occupied orbital space into an active and an inactive part.

### 1. Introduction

One of the basic concepts in modern theory of electronic structure in molecular systems is the molecular orbital. Since its introduction by Hund and Mulliken [1, 2] it has played a dominant role in qualitative as well as quantitative descriptions of "what electrons are doing in molecules". The primary tool for the calculation of molecular orbitals, in ordinary closed-shell molecules, is Hartree-Fock theory [3, 4], especially in the form developed for molecules by Roothaan [5]. Such calculations are today routinely performed in many

laboratories with existing ab initio self-consistent field programs. Further, certain classes of open-shell systems can be treated with a similar technique [5].

Hartree-Fock theory neglects the instantaneous interaction between the electrons, the so-called electron correlation. However, even in a more exact theory, where electron correlation is included, the molecular orbital concept survives - in terms of the natural orbitals [6]. The natural orbital concept has become a powerful tool for analysing complex wavefunctions, and has also been used in developing methods for determining accurate wavefunctions. Examples are the iterative natural orbital method [7] and the PNO-CI method [8, 9].

The Hartree-Fock model gives a surprisingly accurate zeroth-order wavefunction for stable

\* Present address: Lehrstuhl für Theoretische Chemie der Universität Karlsruhe, 7500 Karlsruhe, Postfach 6380, Germany.

closed-shell molecules and also in many cases for open-shell systems. On the other hand, it breaks down completely in situations where the electrons undergo substantial rearrangements, as in the formation or breaking of chemical bonds. A qualitatively correct description of these situations requires a more complicated wavefunction, even at the zeroth-order level of approximation. A single configuration description of the electronic structure is not sufficient here; the wavefunction must comprise several electronic configurations in order to be even qualitatively correct. This leads immediately to a multi-configurational self-consistent field (MCSCF) theory as the natural extension of the Hartree-Fock method [10]. The molecular orbital concept is preserved in such an extension, but the orbitals now have non-integral occupation numbers. In addition, the concept of an orbital energy can be maintained, through the use of the extended Koopmans' theorem [11]. MCSCF theory has been reviewed in a number of recent papers in which further details may be found [12–14].

Many different methods have been utilized for optimizing the orbitals in an MCSCF wavefunction. In this paper we shall deal exclusively with the so-called “super-CI” method developed by Grein et al. [15, 16]. The basis of this method is the Brillouin theorem [17], in the form generalized to multi-configurational wavefunctions by Levy and Berthier [18]. This generalized Brillouin theorem will be referred to here as the Brillouin–Levy–Berthier (BLB) theorem, a name which has recently been suggested [19]. The super-CI method of Grein et al. uses the coefficients from a CI calculation which comprises all single excitations out of the reference MC function to construct a unitary rotation of the orbitals in each iteration. Cheung et al. [19] have recently suggested an alternative approach, which has proved to be rapidly convergent in a number of applications [20]. In their approach, the natural orbitals (with highest occupation numbers) of the singly-excited CI (super-CI) wavefunction are chosen as the improved orbitals. This ensures maximum overlap between the super-CI wavefunction and the improved MC function.

• In the present work a detailed procedure will be discussed for MCSCF calculations employing a

special type of MC wavefunction for which the natural orbital super-CI approach is a convenient choice. The MC wavefunction to be considered is constructed from two sets of occupied orbitals: the *inactive* and *active* orbitals. The inactive orbitals are doubly occupied in all configurations in the MC wavefunction. Conversely, the active orbitals define a subset of the total orbital space, in which the configuration expansion is chosen to be complete. The wavefunction constructed in this way will be called the complete active space (CAS) wavefunction. This type of MC expansion, which has also been discussed by Ruedenberg and Sundberg [21], has several conceptual as well as computational advantages. First, orbital rotations within the inactive and active subspaces leave the energy unchanged, and therefore do not need to be considered. This leads to a simplification of the orbital optimization step, and may also, in fact, give faster convergence. By choosing the orbitals to be natural orbitals these remain well-defined during the iterations, and converge to the natural orbitals of the MCSCF wavefunction. Secondly, with the use of a complete CI wavefunction in the active subspace, the sometimes difficult choice of the “dominant” configurations is completely avoided and the only remaining “chemical” problem is the choice of active orbitals.

One difficulty may be the length of a complete CI expansion since even with a rather small number of active orbitals, the wavefunction may comprise several thousand terms: thus distributing eight electrons among eight orbitals in all possible ways yields 1764 singlet configurations if no reduction by symmetry is possible. This in turn gives rise to two computational problems. An efficient method must be found to solve the MC secular problem. The graphical unitary group approach (GUGA) developed by Shavitt [22], following the theoretical development of Paldus [23], is extremely well suited for complete CI calculations, and as it is a direct method rather long expansions can be handled in an efficient way. A GUGA program has therefore been written to solve the MC secular problem in the present work. For the more recent development of the unitary group approach, see for example, refs. [48–50].

The second, more serious problem concerns the solution of the super-CI secular problem. In the conventional super-CI approach [15, 16, 19], the super-CI wavefunction is constructed by considering first all possible single excitations out of each configuration in the MC wavefunction: obviously the number of such excitations easily becomes exceedingly large as the number of terms in the MC function increases. Another approach is adopted here: the matrix elements between the super-CI states are expressed directly in terms of reduced density matrices for the MC state, and the calculation thereby becomes independent of the length of the MC expansion. In practice, this cannot be achieved without approximating some of the matrix elements, as the exact expressions involve third-order density matrices, which would present considerable computational problems.

Different levels of approximation of the super-CI matrix elements are possible. A practical choice is governed by the rate of convergence and the ease of computation. One possible choice to be discussed has the advantage of reducing the two-electron transformation problem substantially: as the transformation problem is one of the major bottlenecks in applications of the MCSCF method, this approximation is of special interest.

In the following sections the density matrix formulated super-CI approach will be presented and discussed in detail. The different levels of approximation will be studied and some calculations on  $N_2$ ,  $F_2$  and  $CuF_2$  will be presented as examples.

## 2. The one-electron basis and the hamiltonian

The one-electron orbital space used to build the configurations in the CAS wavefunction will be termed the “primary space” in the following discussion. The remainder of the total orbital space, obtained from the atomic basis by the LCAO procedure, is termed *secondary*. This latter space plays no role in the final (orbital-optimized) wavefunction but is used, during the optimization procedure, to improve the primary space. The primary space is subdivided into *inactive* and *active* subspaces. The inactive orbitals are defined as those orbitals which are doubly occupied in all configura-

tions which comprise the CAS wavefunction. In actual applications of the method these would be orbitals which have occupation numbers close to two and which are not expected to change during the chemical process studied. The active orbitals constitute the space in which the CI wavefunction is complete, that is, all possible orbital occupations and spin-couplings consistent with the desired state are included. These orbitals have been termed “reaction orbitals” in the work of Ruedenberg and Sundberg [21]: the term active is preferred here to denote a more general orbital space, which would commonly include those orbitals undergoing a radical change during the course of a chemical reaction, but could also comprise *all* valence orbitals (a full valence CI study) or all orbitals defining a number of excited states, etc.

Labels  $i, j, k$  and  $l$  will be used for inactive orbitals,  $t, u, v$  and  $x$  for active orbitals and  $a, b, c$  and  $d$  for secondary orbitals. The inactive, active and secondary subspaces will be designated as  $\{II\}$ ,  $\{Ia\}$  and  $\{II\}$ , respectively. Labels are also needed for orbitals belonging to the complete space  $\{Ii\} + \{Ia\} + \{II\} = \{I\} + \{II\}$ . Letters  $p, q, r$  and  $s$  will be used in this case. The number of orbitals in each subspace is given as  $n_{ii}$ ,  $n_{ia}$  and  $n_{ii}$ , respectively, with  $n = n_{ii} + n_{ia} + n_{ii}$ .

The non-relativistic spin-independent hamiltonian for the molecular system,

$$\hat{H} = \sum_{\alpha} \hat{h}_{\alpha} + \frac{1}{2} \sum_{\alpha, \beta} \hat{g}_{\alpha\beta}, \quad (1)$$

will be defined in terms of the generators of the unitary group<sup>†</sup>.

$$\hat{E}_{pq} = \sum_{\sigma} \hat{a}_{p\sigma}^{\dagger} \hat{a}_{q\sigma}, \quad (2)$$

where  $\hat{a}_{p\sigma}^{\dagger}$  ( $\hat{a}_{p\sigma}$ ) is a creation (annihilation) operator for an electron in orbital  $\varphi_p$  with spin quantum number  $\sigma$ . The operator  $\hat{E}_{pq}$  can also be considered as an excitation operator replacing one orbital  $\varphi_q$  with the orbital  $\varphi_p$  and leaving the spin unchanged. This property of  $\hat{E}_{pq}$  will be used later in the discussion. The hamiltonian (1) in second quantized form is given in terms of the operators (2) as:

$$H = \sum_{p,q} h_{pq} \hat{E}_{pq} + \frac{1}{2} \sum_{p,q,r,s} (pq|rs) [\hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps}], \quad (3)$$

<sup>†</sup> For a thorough discussion of the unitary group and the generators of its Lie algebra see, for example, ref. [47].

where the one- and two-electron integrals are defined as

$$h_{pq} = \langle p | \hat{h} | q \rangle, \quad (4a)$$

$$(pq | rs) = \langle \varphi_p(1) \varphi_r(2) | 1/r_{12} | \varphi_q(1) \varphi_s(2) \rangle. \quad (4b)$$

The generators  $\hat{E}_{pq}$  obey the commutation relations

$$[\hat{E}_{pq}, \hat{E}_{rs}] = \delta_{qr} \hat{E}_{ps} - \delta_{ps} \hat{E}_{rq}. \quad (5)$$

The configuration state functions (CSF's) included in the CAS wavefunction contain no orbitals from the secondary space, and orbitals from the inactive subspace will always be doubly occupied. Denoting such a CSF by  $|\mu\rangle$  the following relations are immediately obtained:

$$\hat{E}_{pa} |\mu\rangle \equiv 0 \quad (\varphi_a \text{ in } \{\text{II}\}), \quad (6a)$$

$$E_{ip} |\mu\rangle = 2\delta_{ip} |\mu\rangle \quad (\varphi_i \text{ in } \{\text{Ii}\}). \quad (6b)$$

From these equations and the commutation relation (5) the following scheme is obtained for "contracting" products of generators which contain secondary or inactive orbitals:

$$\hat{E}_{pb} \hat{E}_{aq} |\mu\rangle = \delta_{ab} \hat{E}_{pq} |\mu\rangle \quad (\varphi_a \text{ and } \varphi_b \text{ in } \{\text{II}\}), \quad (7a)$$

$$\hat{E}_{ip} \hat{E}_{qj} |\mu\rangle = \{\delta_{ij}(2\delta_{pq} - \hat{E}_{qp}) + 2\delta_{ip} \hat{E}_{qj}\} |\mu\rangle \quad (\varphi_i \text{ and } \varphi_j \text{ in } \{\text{Ii}\}). \quad (7b)$$

The relations (5)–(7) given above will be used later to derive explicit expressions for Brillouin matrix elements and for matrix elements between states which are singly-excited with respect to the CAS wavefunction.

### 3. The complete active space (CAS) CI wavefunction

The wavefunction on which the present method is based is constructed as a complete CI wavefunction in the active subspace. Thus it contains all possible distributions of  $N_a$  electrons among the  $n_{ia}$  active orbitals, satisfying all possible spin-couplings which correspond to a total spin quantum number  $S$ . The inactive orbitals are always doubly occupied; the number of electrons in such orbitals being  $N_i$ . The total number of electrons is  $N = N_i + N_a$ . The CAS wavefunction is written as

$$|0\rangle = \sum_{\mu} C_{\mu} |\mu\rangle, \quad (8)$$

where the sum extends over the complete configuration space  $|\mu\rangle$ .

Shavitt [22] has presented a detailed method for the solution of the secular problem which determines the expansion coefficients  $C_{\mu}$ , using a direct configuration interaction method based on the unitary group approach of Paldus [23]. In this method the matrix elements are given as

$$H_{\mu\nu} = \sum_{p,q}^{(\text{I})} h_{pq} \langle \mu | \hat{E}_{pq} | \nu \rangle + \frac{1}{2} \sum_{p,q,r,s}^{(\text{I})} (pq | rs) \langle \mu | (\hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps}) | \nu \rangle. \quad (9)$$

In the direct CI approach [24] contributions to the vector  $\sigma = \mathbf{H}\mathbf{C}$  are obtained directly from the list of one- and two-electron integrals, according to

$$\sigma_{\mu} = \sum_{p,q}^{(\text{I})} h_{pq} \sum_{\nu} C_{\nu} A_{pq}^{\mu\nu} + \sum_{p,q,r,s}^{(\text{I})} (pq | rs) \sum_{\nu} C_{\nu} B_{pqrs}^{\mu\nu}. \quad (10)$$

Comparing (9) and (10) one notes that the coupling coefficients  $A_{pq}^{\mu\nu}$  and  $B_{pqrs}^{\mu\nu}$  are given as matrix elements of single generators and the generator products by

$$A_{pq}^{\mu\nu} = \langle \mu | \hat{E}_{pq} | \nu \rangle \quad (11a)$$

and

$$B_{pqrs}^{\mu\nu} = \langle \mu | \frac{1}{2}(E_{pq}E_{rs} - \delta_{qr}E_{ps}) | \nu \rangle. \quad (11b)$$

In Shavitt's approach a graphical representation of the Gelfand basis (the CSF's) is employed. This representation facilitates the generation of the coupling coefficients for each of the integrals in (10), making a direct calculation of (10) possible. The information is stored as a list of formulae, containing the coupling constants corresponding to a given integral, followed by data necessary for a straightforward evaluation of the interacting CSF's  $|\mu\rangle$  and  $|\nu\rangle$ .

The only information required in performing the singly-excited state CI is the first- and second-order density matrices for the active subspace. These matrices can be generated directly from the formulae created in the CI step. The extra work involved therefore corresponds to one extra iteration in the iterative solution of the secular equation. This can be demonstrated explicitly by starting from the energy expression

$$E_0 = \sum_{\mu,\nu} C_\mu H_{\mu\nu} C_\nu. \quad (12)$$

Inserting the expression (9) for the hamiltonian matrix elements we obtain directly the first and second-order density matrices as

$$D_{tu}^{(0)} = \sum_{\mu,\nu} C_\mu \langle \mu | \hat{E}_{tu} | \nu \rangle C_\nu, \quad (13a)$$

$$P_{tuvx}^{(0)} = \sum_{\mu,\nu} C_\mu \langle \mu | \frac{1}{2} [\hat{E}_{tu} \hat{E}_{vx} - \delta_{uv} \hat{E}_{tx}] | \nu \rangle C_\nu. \quad (13b)$$

Thus the list of coupling coefficients enables a direct calculation of these matrices; all information needed for a given matrix element is stored sequentially on the “formula tape”.

As the CI wavefunction is complete in the active subspace, it is invariant to unitary transformations among the active orbitals. These orbitals are thus only defined to within such a unitary transformation. As is evident in later discussion, a useful choice is those orbitals which diagonalize the matrix  $\mathbf{D}^{(0)}$ , that is, the natural orbitals. The iterative method to be discussed converges to such orbitals, which have several advantages, as follows. The calculations can be simplified by assuming this condition to be fulfilled also *before* convergence. Normally the density matrix (or, in cases of degenerate states, a symmetry-averaged density matrix) transforms as the totally symmetric irreducible representation of the point group under consideration. If it does not have this property the result can be said to be symmetry unstable, a situation similar to that sometimes encountered in Hartree–Fock calculations. Such instabilities are generally due to low-lying excited states; if they occur in a CASSCF calculation, this indicates that important orbitals are missing from the active subspace. However, in the case of a totally symmetric density matrix, the natural active orbitals are also symmetry orbitals.

On the other hand, in a limited MCSCF calculation the orbitals often tend to localize, and thus do not transform as irreducible representations of the point group. This reduces the extent to which symmetry can be used in the calculation. Such a problem does not arise here – the full symmetry can always be used. This is especially important for the efficient calculation of AO integrals and the subsequent transformation of these integrals to MO basis.

Finally, the natural orbitals are particularly suited to obtaining a “chemical insight” into the results of the calculation. Ruedenberg has discussed this advantage of the natural orbitals of a CAS type wavefunction, and has introduced the term “natural reaction orbitals” [21]. To conform with the nomenclature introduced above they are here termed *natural active orbitals* (NAO’s).

#### 4. The Brillouin–Levy–Berthier theorem and the super-CI method

One of the more effective methods used for the solution of the multi-configurational SCF problem is the so-called “super-CI” method, as formulated by Grein and co-workers [15, 16]. The method is based on the Brillouin–Levy–Berthier (BLB) or “generalized Brillouin” theorem [18], which can be written in second quantized form as

$$\langle 0 | \hat{H}(\hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} - \hat{a}_{q\sigma}^\dagger \hat{a}_{p\sigma}) | 0 \rangle = 0. \quad (14)$$

The super-CI method begins with a calculation of the expansion coefficients in the MC wavefunction (8). A new secular problem is then constructed using  $|0\rangle$  and the singly-excited states (SX-states),

$$(\hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} - \hat{a}_{q\sigma}^\dagger \hat{a}_{p\sigma}) | 0 \rangle, \quad (15)$$

as a configurational basis. The expansion coefficients are determined by the usual variational method. Improved orbitals can then be obtained either by performing a unitary transformation of the original MO’s, based on the expansion coefficients for the SX states [15], or by a natural orbital analysis of the resulting super-CI wavefunction [19]. The latter method will be used in the present work. Detailed formulae for the super-CI density matrix elements are presented in appendix B, where they are expressed in terms of the super-CI expansion coefficients and the first- and second-order density matrices (13) for the CAS wavefunction.

The selection of orbitals for the next iteration from the super-CI NO’s is straightforward. The  $n_{ii}$  orbitals with the highest occupation numbers are chosen as the inactive orbitals and the next  $n_{ia}$  as the active orbitals. The order within each group is immaterial. Again this is due to the fact that the

CAS wavefunction is complete. None of the selection problems discussed by Cheung et al. [19] are therefore encountered in the present scheme.

In practical applications, however, the super-CI method has some serious disadvantages. It is necessary to construct first the matrix elements between CSF's singly-excited with respect to all CSF's occurring in (8). The number of such singly-excited configurations can be exceedingly large if many terms are used in the multi-configurational wavefunction. The method is therefore limited to short expansions of the form (8), which in turn means that for a wavefunction of the CAS type only a few active orbitals can be used, severely limiting the applicability of the method.

Further, in order to calculate these matrix elements between SX-states, a list of two-electron integrals ( $pq|rs$ ) is required in which two of the indices correspond to secondary orbitals. A partial transformation of the AO basis to such MO integrals has to be performed in each iteration; as this transformation is very time-consuming when the basis set is large it may actually become the bottleneck of such a calculation.

It is, however, possible to remedy these limitations of the method to a certain extent, when the multi-configurational wavefunction is of CAS type. As will be shown, for this case it is possible to express the BLB matrix elements directly in terms of the first- and second-order density matrices for the CAS functions. Approximate expressions for the interaction between the SX-states can also be obtained in this form. In this way, the need to construct explicit matrix elements between all singly-excited configurations is avoided. Further, by introducing effective one-electron operators it is also possible to reduce the number of MO two-electron integrals to those which have at most *one* index in the secondary space.

The BLB theorem as stated in eq. (14) corresponds to an unrestricted variation of all spin-orbitals occupied in the MCSCF wavefunction  $|0\rangle$ . However, in most applications the variations are restricted to the orbitals defining the space part of the wavefunction, and the same space orbitals are used independent of the spin labels. The first order change in energy due to unitary transformations among these orbitals is then given by (14), summed over spin. It follows that an *orbital restricted* BLB

theorem can be expressed in terms of the generators  $\hat{E}_{pq}$  as defined in (2),

$$\langle 0 | \hat{H}(\hat{E}_{pq} - \hat{E}_{qp}) | 0 \rangle = 0. \quad (16)$$

Expression (16) is identically zero if both the orbitals  $\varphi_p$  and  $\varphi_q$  are either inactive or active. The action of the generators is then to produce a linear combination of states already present in  $|0\rangle$  (or zero), and the result then follows, as  $\hat{H}$  is diagonal in this space and the two generators in (16) are hermitian adjoint. This property of (16) is a manifestation of the above observation that the wavefunction is invariant to a unitary transformation among the active orbitals. In fact, only three types of singly-excited states interact with  $|0\rangle$  and thus need to be included in a super-CI calculation. These are

$$|i \rightarrow a\rangle = 2^{-1/2} \hat{E}_{ai} |0\rangle \quad (\varphi_i \text{ in } \{\text{Ii}\} \text{ and } \varphi_a \text{ in } \{\text{II}\}), \quad (17a)$$

$$|t \rightarrow a\rangle = n_t^{-1/2} \hat{E}_{ta} |0\rangle \quad (\varphi_t \text{ in } \{\text{Ia}\} \text{ and } \varphi_a \text{ in } \{\text{II}\}), \quad (17b)$$

and

$$|i \rightarrow t\rangle = m_t^{-1/2} \hat{E}_{ti} |0\rangle \quad (\varphi_i \text{ in } \{\text{Ii}\} \text{ and } \varphi_t \text{ in } \{\text{Ia}\}), \quad (17c)$$

where  $n_t = D_{tt}^{(0)}$  and  $m_t = 2 - D_{tt}^{(0)}$ . The numerical factors ensure that the SX-states, which have been constructed with the excitation operators  $\hat{E}_{pq}$ , are normalized. The hermitian adjoint of these operators give identically zero when operating on  $|0\rangle$  [cf. eqs. (6a) and (6b)], and the LHS of eq. (16) therefore reduces to one term.

The singly-excited states defined by (17a)–(17c) are normalized but in general not orthogonal. The overlap integrals are easily shown to be

$$\langle i \rightarrow a | j \rightarrow b \rangle = \delta_{ij} \delta_{ab}, \quad (18a)$$

$$\langle t \rightarrow a | u \rightarrow b \rangle = \delta_{ab} D_{tu}^{(0)} / (n_t n_u)^{1/2}, \quad (18b)$$

$$\langle i \rightarrow t | j \rightarrow u \rangle = \delta_{ij} (2\delta_{tu} - D_{tu}^{(0)}) / (m_t m_u)^{1/2}, \quad (18c)$$

using eq. (6) and the contraction formulae (7). Those overlap integrals not given above are identically zero. From the expressions (18) it follows that the SX-states are orthogonal if and only if the density matrix  $\mathbf{D}^{(0)}$  for the CAS wavefunction is diagonal – that is, if the *active orbitals are chosen to be natural orbitals*. If, at each iteration, the natural orbitals of

the super-CI wavefunction are chosen as improved orbitals, the active orbitals will be very close to natural orbitals for the CAS wavefunction after a few iterations in the SCF procedure. Therefore, the SX-states (17) will be assumed to be orthogonal throughout the calculation, although this situation does not obtain until convergence. It should be noted, however, that this assumption will only be used during the calculation of the super-CI expansion coefficients. The calculation of the density matrix (cf. appendix B) includes the overlap terms. It follows also from (18) that the SX-states are linearly independent except for pathological cases when the density matrix  $\mathbf{D}^{(0)}$  has eigenvalues equal to zero or two.

The super-CI wavefunction,  $|\text{SX}\rangle$ , can now be expressed in terms of the SX-states as

$$|\text{SX}\rangle = \hat{E}_{\text{SX}} |0\rangle, \quad (19)$$

where the super-CI generator is defined as a linear combination of the excitation operators used in (17a)–(17c),

$$\begin{aligned} \hat{E}_{\text{SX}} = & c_0 \hat{E}_0 + \sum_{i,a} c_{ia} 2^{-1/2} \hat{E}_{ai} + \sum_{i,a} c_{ia} n_i^{-1/2} \hat{E}_{ai} \\ & + \sum_{i,t} c_{it} m_i^{-1/2} \hat{E}_{it}. \end{aligned} \quad (20)$$

Here  $\hat{E}_0$  is the unit operator and  $c_0$  has been introduced to ensure a normalized SX wavefunction.

In order to calculate the expansion coefficients  $c$ , expressions for the matrix elements between an SX-state and  $|0\rangle$  (the BLB matrix elements) must first be obtained, as well as matrix elements between the different singly-excited states. This problem is discussed in the next section.

## 5. The BLB matrix elements and super-CI expansion coefficients

By inserting the explicit form of the hamiltonian (3) into (16), it may be noted that the BLB matrix elements can be expressed as a linear combination of matrix elements of generator products which involve at most three single generators. However, due to the definitions (17) of the SX-states, at least two of the six orbital indices involved will be either inactive or secondary. It is therefore possible to reduce these

generator products to contain at most two operators. This result is achieved by using the commutation relation (5) together with the contractions defined by eqs. (7). When this contraction has been performed the BLB matrix elements are expressible in terms of generator products containing at most two single generators, or, equivalently, in terms of the first- and second-order reduced density matrices (13). The required algebraic manipulations are straightforward, although somewhat tedious, and are omitted here. The resulting equations can be re-expressed, in part, in terms of effective one-electron operators. Two Fock-type operators  $\hat{F}^i$  and  $\hat{F}^a$  (the *inactive* and *active* one-electron operators) will therefore be defined, whose matrix elements are given by

$$F_{pq}^i = h_{pq} + 2 \sum_k^{(ii)} [(pq|kk) - \frac{1}{2}(pk|qk)], \quad (21a)$$

$$F_{pq}^a = \sum_{t,u}^{(ia)} D_{tu}^{(0)} [(pq|tu) - \frac{1}{2}(pt|qu)]. \quad (21b)$$

A matrix element of the *total* effective one-electron operator  $\hat{F}$  is given as the sum of these two contributions:

$$F_{pq} = F_{pq}^i + F_{pq}^a. \quad (22)$$

With the use of these operators, the BLB matrix elements corresponding to the three types of SX-states (17) are given as

$$\langle 0 | \hat{H} | i \rightarrow a \rangle = 2^{1/2} F_{ai}, \quad (23a)$$

$$\begin{aligned} \langle 0 | \hat{H} | t \rightarrow a \rangle = & n_t^{-1/2} \\ & \times \left\{ \sum_u F_{au}^i D_{tu}^{(0)} + \sum_{u,v,x} 2P_{uvx}^{(0)}(au|vx) \right\}, \end{aligned} \quad (23b)$$

$$\begin{aligned} \langle 0 | \hat{H} | i \rightarrow t \rangle = & m_i^{-1/2} \left\{ 2F_{it}^a + \sum_u (2\delta_{iu} - D_{iu}^{(0)}) F_{it}^i \right. \\ & \left. - \sum_{u,v,x} 2P_{uvx}^{(0)}(iu|vx) \right\}. \end{aligned} \quad (23c)$$

All summations in these expressions are over active orbitals only.

The matrix elements (23) have a very simple form, and are easily computed. The first step involves calculation of the one-electron matrices (21) and is most easily performed in a supermatrix formalism, that is, the matrices are first computed over the

AO basis using a list of supermatrix elements, and are then transformed to the MO-basis. The computational effort corresponds exactly to the setting-up of the Fock matrix in a conventional closed shell Hartree–Fock calculation [5]. In the second step, the remaining two-electron contributions to (23b) and (23c) are included. This involves multiplication of two-electron integrals by the second-order density matrix  $\mathbf{P}^{(0)}$ . Only two-electron integrals of types  $(au|vx)$  and  $(iu|vx)$  are needed, and thus only one of the integral indices runs over the entire orbital space. The most time-consuming part of this step is the calculation of the second term in (23b). The calculation of all such contributions is an  $n_{ia}n_{ia}^4$  process, and as  $n_{ia}$  is a small number the evaluation of the BLB matrix elements will constitute only a small fraction of the total computational effort. For the same reason the number of second-order density matrix elements is small enough to be kept in fast storage during the whole process.

However, in order to obtain an estimate of the expansion coefficients for the SX-states, it is not sufficient to calculate only the BLB matrix elements – the matrix elements between the SX-states are also needed when a super-CI method is used. These matrix elements will in general involve products of up to four single generators, in which at least two of the eight orbital indices will correspond to inactive or secondary orbitals. Six active indices then remain after contraction, and so the matrix elements are given in terms of first-, second- and third-order reduced density matrices for the active space\*. The explicit formulae for these matrix elements have been presented earlier [25]. For completeness they are, however, given again in appendix A.

Obtaining third-order reduced density matrices from a Shavitt graph is, however, a complicated procedure, and a simplification of this part of the calculation is therefore highly desirable. It may be noted that the iterative procedure used in an MCSCF calculation must obey one obvious condition: it should converge, that is, it should

\* Actually only a component of the third order density matrix is needed, according to the definition given in appendix A, but for the sake of simplicity the matrix  $Q$  defined in eq. (A.1) will in the text be referred to as the third order density matrix.

reduce the BLB matrix elements to zero – this condition is not only necessary but also sufficient. In practice, however, a second condition also has to be fulfilled: the convergence must be sufficiently rapid to keep the number of iterations within practical limits.

The above-mentioned condition can be fulfilled even if the super-CI secular problem is not solved exactly. Several levels of approximation are possible, and an optimal choice depends on two conflicting criteria – simplicity of evaluation and rapidity of convergence.

The simplest approximation which can be applied to the matrix elements between the SX-states is to assume that the third-order density matrix is given as a product of the first- and second-order density matrices, that is, to set

$$Q_{vxyztu}^{(0)} = P_{vxyz}^{(0)} \cdot D_{tu}^{(0)}. \quad (24)$$

This will affect only a small number of matrix elements (cf. appendix A). All matrix elements are then expressed in terms of first- and second-order density matrices. However, they still contain two-electron integrals with two orbital indices outside the active subspace.

The next level of approximation, which has been investigated in detail, is to assume that the first-order density matrix is diagonal

$$D_{tu}^{(0)} = n_t \delta_{tu}, \quad (25)$$

and to approximate the second-order density matrix by the occupation numbers  $n_i$ :

$$P_{tuu}^{(0)} = \frac{1}{2}n_t(n_u - \delta_{tu}), \quad (26a)$$

$$P_{tuu}^{(0)} = -\frac{1}{4}n_t n_u \quad (t \neq u), \quad (26b)$$

all other matrix elements being set to zero. Thus only the Coulomb and exchange parts of the second-order density matrix are kept. The matrix elements between the SX-states now take an especially simple form. They are given in eqs. (A.4a)–(A.4f) in appendix A. These matrix elements are all expressed in terms of the Fock operator (22) plus a few additional two-electron integrals. This level of approximation will be denoted A1. An obvious third level of approximation (denoted A2) is then to neglect these integrals and keep only the Fock matrix contributions. As will be illustrated in the next section this



approximation has, in fact, performed quite satisfactorily in a number of applications. It has the great advantage that the only transformed two-electron integrals necessary are those used to construct the BLB matrix elements (23), and they contain only one orbital index outside the small active subspace.

In the next section some examples will be given to illustrate the convergence properties of the approximate super-CI approach. It is clear from the timing data given that the choice between the different levels of approximation depend to a large extent on how the rate of convergence changes compared to the reduction in transformation time. Another approach which could be used in cases where the more approximate scheme does not converge satisfactorily is to use the same interaction matrix for several iterations and only change the BLB matrix elements.

## 6. Computational results

The first test applications of the CASSCF super-CI method have been directed towards examining the convergence behaviour of the method rather than the use of a very large number of configurations in  $|0\rangle$ . In this section, timing aspects of several calculations will be discussed; a fuller treatment of results obtained for several  $N_2$  potential curves is given in the next section.

Table 1 lists the convergence of CASSCF calculations on the ground state of  $N_2$  at  $R = 2.075$  au. Results are presented for both the A1 and A2 schemes described above. The atomic basis set used consisted of van Duijneveldt's (9s, 5p) basis [26], contracted as (5,1,1,1,1/3,1,1), and augmented with one d function, exponent 0.95. Orbitals from a closed-shell SCF calculation which gave  $E_{SCF} = -108.96898$  au, were used as a starting guess for both CASSCF calculations. The active space comprised the  $2\sigma_u$ ,  $3\sigma_g$ ,  $1\pi_u$ ,  $1\pi_g$  and  $3\sigma_u$  orbitals, giving a CASSCF wavefunction with 80 configurations. It is quite evident from table 1 that the convergence behaviour of the A2 scheme is in no way inferior to that of A1. Further the A2 scheme is more than three times faster (cf. table 2): as expected the major part of this difference results from the

need to transform integrals in which two indices span the full orbital space, and which results in a four-fold increase in CPU-time during the transformation step. For the A2 case this step requires around 60% of the time used in one iteration, while for A1 almost 85% of the iteration time is consumed by the transformation.

In both cases, the single excitation CI step requires a rather small fraction of the time. The difference between the A1 and A2 timings reflects mainly the different number of diagonalization iterations required during this step: it appears that the somewhat simpler form of the matrix elements used in the A2 scheme results in a matrix which is more diagonally dominant than is the case for A1. Owing to the appearance, during the iterations, of some diagonal super-CI matrix elements with values below that of  $|0\rangle$ , both schemes require the introduction of a level shift. A value of 2.0 au was used for the  $N_2$  calculations; although this is rather larger than the values commonly used in SCF calculations [27], our experience has been that

Table 1  
Convergence of the CASSCF super-CI method ( $N_2$  with 40 basis function,  $R = 2.075$  au)<sup>a)</sup>

Iteration	A1		A2	
	$E$	max (C) <sup>b)</sup>	$E$	max (C) <sup>b)</sup>
1	-107.07967	0.312	-107.07967	0.316
2	-108.88821	3.584	-108.88637	3.084
3	-109.03442	0.288	-109.03399	0.302
4	-109.05770	2.481	-109.05753	3.488
5	-109.06775	0.744	-109.06775	0.695
6	-109.09028	0.379	-109.08604	0.319
7	-109.10552	0.176	-109.10238	0.250
8	-109.11085	0.148	-109.11038	0.172
9	-109.11228	0.066	-109.11199	0.056
10	-109.11258	0.044	-109.11245	0.029
11	-109.11265	0.032	-109.11260	0.015
12	-109.11268	0.022	-109.11266	0.006
13	-109.11270	0.018	-109.11269	0.003
14	-109.11271	0.012	-109.11271	0.001
15	-109.11271	0.011	-109.11272 <sup>c)</sup>	$\approx 10^{-5}$

<sup>a)</sup> Starting vectors from an SCF calculation, on the closed shell ground state was used.

<sup>b)</sup> Maximum super-CI coefficient.

<sup>c)</sup> Better convergence than  $10^{-5}$  is not meaningful in a single precision calculation on a UNIVAC 1108 computer which has a word length of 36 bits.

Table 2  
Timing data for some CASSCF super-CI calculations

Molecule	Number of active orbitals	Number of active electrons	Number of configurations	Level of approximation	CPU time per iteration <sup>a)</sup>			Number of iterations <sup>b)</sup>	Basis set size
					transformation	CI	super-CI		
N <sub>2</sub>	7	8	80	A1	21	2	2	18	40
N <sub>2</sub>	7	8	80	A2	5	2	1	15	40
N <sub>2</sub>	6	6	32	A2	4	1	1	17	40
N <sub>2</sub>	9	8	726	A2	6	21 <sup>d)</sup>	1	5 <sup>c)</sup>	40
F <sub>2</sub>	7	6	80	A2	5	2	1	16	40
CuF <sub>2</sub>	12	21	86	A2	15	11	3	20	50

<sup>a)</sup> In seconds on a UNIVAC 1100/80 computer.

<sup>b)</sup> Number of iterations needed for convergence to  $10^{-5}$  au in energy and  $10^{-3}$  in the super-CI coefficients. In all cases starting vectors were obtained from SCF calculations.

<sup>c)</sup> Using orbitals from the seven active orbitals calculation.

<sup>d)</sup> First iteration. Decreases with the number of iterations (see text).

smaller values may not be adequate, and in any case seldom provide much improvement in convergence. As a final note on the timing for the super-CI step, it is worth reiterating that this step is expected to be quite rapid in general, as it involves obtaining only the lowest root from a small (dimension  $n_1 n_{II} + n_{I} n_{Ia}$ ) CI problem.

Some additional calculations are also given, as a further illustration of timing, in table 2. The first calculation is again the ground state of N<sub>2</sub>, but with the  $2\sigma_u$  orbital transferred from the active to the inactive subspace (see below). This gives six active orbitals and six active electrons, and is the smallest possible active subspace which properly describes the dissociation of N<sub>2</sub> into two nitrogen atoms. The dimension of the CI problem is now reduced from 80 to 32 ( $D_{2h}$  symmetry was used in all calculations presented in table 2). The timing for the CI step is correspondingly decreased, while the other times remain almost unchanged; the transformation time is a little smaller owing to the decrease in active subspace.

A preliminary calculation has also been performed on N<sub>2</sub>( $X^1\Sigma_g^+$ ) with eight electrons distributed among nine active orbitals; in addition to the calculation using seven active orbitals, the  $2\pi_u$  orbitals are also included in the active subspace. The CASCI wavefunction now comprises 726 configuration state functions. The timing is dominated by the CI step, which in the first iteration needs

21 s. In the direct CI scheme this time is proportional to the number of iterations necessary to converge on the CI energy and wavefunction. Six such CI iterations were needed in the first SCF iterations. Near SCF convergence fewer CI iterations are needed since the calculation can start from the CI coefficients of the preceding SCF iteration. Thus, in the last iteration only two CI iterations were needed and the CPU time used in the CI section dropped to 7 s.

Most noticeable in this calculation is that the time used in the orbital optimization step has not changed compared to the smaller N<sub>2</sub> calculations. This part of the calculation is indeed independent of the size of the CI wavefunction and varies only slowly with the size of the active subspace. This is a unique feature of the density matrix formulation of the super-CI approach as it has been presented in this work.

The two other results in table 2 represent calculation on the F<sub>2</sub> molecule, and, as a somewhat larger test problem, the CuF<sub>2</sub> molecule. The F<sub>2</sub> calculation was performed as part of an attempt to improve on previous, unsatisfactory MCSCF and CI calculations [28] which were felt to be in error due to the omission of configurations involving the  $2\pi_u$  orbital from the CI reference space. Accordingly, an active subspace of seven orbitals,  $3\sigma_g$ ,  $3\sigma_u$ ,  $1\pi_g$ ,  $4\sigma_g$  and  $2\pi_u$  was used. With six active electrons, this defines a similar CI problem to the larger N<sub>2</sub>

calculation. An atomic basis analogous to that used for  $N_2$  was used. Timings are generally similar to the  $N_2$  calculation; convergence was obtained in 16 iterations from a closed-shell SCF starting guess, despite the fact that the latter is a rather poor approximation for  $F_2$ .

Copper halides, such as  $CuF_2$ , and their electronic spectra form an area of current interest in this laboratory [29]. For the CASSCF calculations on the ground ( $^2\Sigma_g^+$ ) state, Wachter's atomic basis set [30] was used for Cu, contracted to (8,2,1,1,1,1/6,1,1,1/4,1) and the Dunning basis [31] contracted to (6,1,1,1/4,1) for F, giving 50 contracted functions. Molecular orbitals arising from the Cu 3d and 4s orbitals and the F 2p orbitals (in MO notation:  $6\sigma_g$ ,  $7\sigma_g$ ,  $8\sigma_g$ ,  $3\sigma_u$ ,  $1\delta_g$ ,  $4\sigma_u$ ,  $5\sigma_u$ ,  $1\pi_g$  and  $2\pi_g$ ) were included in the active subspace, giving 21 electrons in 12 orbitals. Timing data for the calculation are given in table 2. Even here, the transformation step needed for the A2 scheme comprises only half of the total time while the timing contribution for the super-CI step is essentially negligible. Convergence was achieved in 20 iterations, using orbitals from a closed-shell SCF calculation on  $CuF_2^-$  as a starting guess. This guess is a particularly poor one as the partitioning into occupied and virtual orbitals in  $CuF_2^-$  is rather different to that of  $CuF_2$ , which does not obey the aufbau principle at the RHF level.

Thus the general conclusions from the timing data discussed above provide strong support for the contention of earlier sections, that in calculations using large active subspaces and long CI expansions, the rate-determining step will be the CI calculation itself, or, for large atomic basis sets, the transformation step, while the super-CI step should require only a small fraction of the total time.

The rate of convergence achieved in the calculations presented in table 2 is acceptable, but not comparable to that which can be obtained in a full second-order process (see e.g. ref. [32]). However, at least in the  $N_2$  case in table 2, a second-order process would have to converge in less than six iterations to be competitive with the A2 scheme in timing, since the iteration time increases with a factor of three. The calculations of the potential curves for  $N_2$ , to be discussed in the next section, show even better convergence behaviour, since in this case starting vectors from nearby points on

the curves could be used. Generally, these calculations converge in about 10 iterations.

## 7. Potential curves for $N_2$

As an example of the capabilities of the CASSCF method, potential curves for the three lowest lying states ( $X^1\Sigma_g^+$ ,  $A^3\Sigma_u^+$ ,  $B^3\Pi_g$ ) of the  $N_2$  molecule have been calculated. Although MCSCF calculations on excited states have appeared [33], there has been surprisingly little work published on the lower states (although, see ref. [34] for details of several unpublished studies). The three states listed provide a variety of conditions under which the CASSCF method can be tested, as the ground state is dominantly a closed shell near equilibrium, the  $A^3\Sigma_u^+$  state is open-shell, and the  $B^3\Pi_g$  state requires density matrix averaging if full axial symmetry is to be retained. Further, while the X and A states dissociate into two ground ( $^4S^o$ ) state nitrogen atoms, the B state dissociates into the  $^4S^o + ^2D^o$  channel [35], and hence calculation of these curves also provides a test of the ability of the CASSCF method in reproducing the splitting between different asymptotic limits.

An obvious choice of active subspace for such calculations would comprise the  $3\sigma_g$ ,  $1\pi_u$ ,  $1\pi_g$  and  $3\sigma_u$  orbitals, denoted CASSCF-6; such a choice may equivalently be regarded as including the atom 2p shells in the active subspace, and automatically includes all configurations required for dissociation into the proper atomic states. Further extension of this active subspace is discussed below.

The atomic basis set used was that given in the previous section. The A2 super-CI scheme was used for all calculations. The number of configurations included in  $|0\rangle$  was 32, 27 and 24 for the three states, respectively. ( $D_{2h}$  symmetry was used in the calculations instead of  $D_{\infty h}$  for computational reasons.) A grid size of 0.05 au was used around the energy minimum region for each of the states.

The resulting potential curves are plotted in fig. 1, and spectroscopic constants obtained from these curves by means of a Dunham analysis [36], together with experimental values [34] are given in table 3. A fifth-order polynomial, fitted to seven points around the minimum, was used in this analysis.

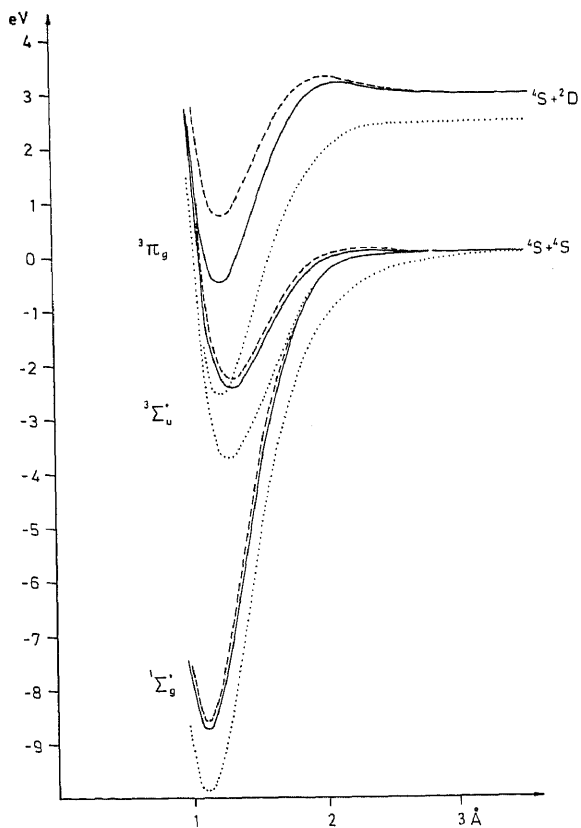


Fig. 1. Calculated and experimental potential curves for the three lowest states of the  $N_2$  molecule: (---) give the CASSCF-6 results; (—) the CASSCF-7 results and (...) the experimental curves [34].

For the ground state, the basis set used gives a  $R_e$  value of 1.069 Å at the SCF level [37], and so there is a substantial improvement over the SCF value at the CASSCF level. The dissociation energy of the ground state, a notorious failure of the SCF method, which gives an error of 4.5 eV, or almost 50% of  $D_e$  [38], is also brought into more satisfactory agreement with experiment using the CASSCF method. For comparison, Lie and Clementi, in an MCSCF calculation using only the configurations needed for proper dissociation, obtained a  $D_e$  value of 7.27 eV [39] in spite of the fact that the calculations were performed with a much larger basis set (7s, 4p and 2d).

The dissociation energy results for the excited states are less satisfactory. Although the absolute error for the A state  $D_e$  is approximately the same as that for the ground state, the relative error is disappointingly large. The entire curve for the  $B^3\Pi_g$  state is in rather poor agreement with experiment, displaying a maximum at 2.0 Å, and a minimum which is much too shallow. The asymptotic limit, however, is in fair agreement with experiment, and is actually very close to the HF limit of 2.85 eV (no correlation is included in the asymptotic limit wavefunction).

A rationalization of some of these difficulties, and their resolution, is presented in the following discussion, and involves enlargement of the active subspace.

Table 3  
Spectroscopic constants for  $N_2$ <sup>a)</sup>

	$X^1\Sigma_g^+$			$A^3\Sigma_u^+$			$B^3\Pi_g$		
	$n_{1a} = 6$	$n_{1a} = 7$	exp <sup>b)</sup>	$n_{1a} = 6$	$n_{1a} = 7$	exp <sup>b)</sup>	$n_{1a} = 6$	$n_{1a} = 7$	exp <sup>b)</sup>
$R_e$ (Å)	1.105	1.108	1.098	1.307	1.309	1.287	1.243	1.230	1.213
$D_e$ (eV)	8.574	8.759	9.905	2.268	2.432	3.680	2.171	3.386	4.896
$\omega_e$ (cm <sup>-1</sup> )	2366.3	2332.6	2358.6	1410.1	1385.0	1460.6	1543.0	1679.6	1733.4
$\omega_e x_e$ (cm <sup>-1</sup> )	16.93	14.55	14.32	19.59	15.15	13.87	17.75	14.39	14.12
$B_e$ (cm <sup>-1</sup> )	1.970	1.962	1.998	1.410	1.405	1.455	1.557	1.592	1.637
$\alpha_e$ (cm <sup>-1</sup> )	0.0164	0.0175	0.0173	0.0203	0.0195	0.0180	0.0224	0.0179	0.0179
$d_e^{c)}$ (cm <sup>-1</sup> × 10 <sup>-6</sup> )	5.46	5.55	5.76 <sup>d)</sup>	5.64	5.78	6.1 <sup>d)</sup>	6.42	5.72	5.9 <sup>d)</sup>

<sup>a)</sup> First and second columns give calculated results with six and seven active orbitals, respectively.

<sup>b)</sup> Ref. [34].

<sup>c)</sup> Centrifugal distortion constant.

<sup>d)</sup> Ref. [51].

Most discussion of MCSCF configuration selection (or, in this context, the choice of the active subspace) has concentrated on the need to describe the separated atom channels properly [40, 41]. However, recent experience in calculations on  $F_2$  and  $Li_2$  [28] has suggested that consideration of the united atom channel can be an important requirement in the choice of active subspace, and a detailed discussion of this problem will be given elsewhere [42].

For the present case, the X, A and B states correlate adiabatically with the lowest  $^1D$ ,  $^3P^o$  and  $^3P$  states of Si [34, 43]. With the above choice of a six-orbital active subspace, it is straightforward to determine from the correlation diagram [34] that the ground state actually correlates with a mixture of Si  $^1D$  ( $M_L = 0$ ) and  $^1S$ , while the A and B states correlate with a complicated mixture of excited states involving the Si 3d orbitals. One reason for these problems is the inclusion of the  $2\sigma_u$  MO in the inactive subspace, as this orbital tends to Si  $3p_z$ , which is actually higher in energy than the united atom limits of  $3\sigma_g$  (Si 3s) and  $1\pi_u$  (Si 2p). The Si  $3p_z$  orbital has an occupation number of less than two in all three united atom limits, which suggests that the  $2\sigma_u$  orbital should be transferred from the inactive to the active subspace.

CASSCF calculations have been performed with an active subspace of seven orbitals, obtained by augmenting the list given above with  $2\sigma_u$  (CASSCF-7). The potential curves obtained are also plotted in fig. 1 and the spectroscopic constants are given in table 3. For all three states the inclusion of  $2\sigma_u$  into the active subspace leads to improved results. The improvement in the  $^3\Pi_g$  results is, however, particularly noteworthy. The error in the  $R_e$  value decreases from 0.030 Å to 0.017 Å and the binding energy ( $D_e$ ) which in the CASSCF-6 calculation was only 44% of the experimental value is now 68%. Substantial improvements are also evident from the last columns of table 3. That the  $2\sigma_u$  orbital is especially important in the B state can also be inferred from the occupation numbers. The  $2\sigma_u$  orbital has an occupation number of 1.99 at  $R_e$  for the X and A states. At the equilibrium bond distance for the B state this number is 1.95; 0.05 electrons have been transferred to the  $3\sigma_g$  orbital.

However, even if the active  $2\sigma_u$  orbital leads to a considerable improvement in the results for the B state, some of the spectroscopic constants are still far from the experimental values. The error in the  $D_e$  value is, for example, 1.15 eV for the X state. Even if this can be attributed to a large extent to the lack of dynamical correlation in the present calculations, it is not unlikely that the  $2\pi_u$  orbital also plays an important role in the electronic structure of the  $N_2$  molecule. This orbital was found to be of particular importance in recent studies on  $F_2$  [28]. Calculations of a number of  $N_2$  potential curves, where  $2\pi_u$  is included into the active subspace are presently under way. These calculations employ a more extended basis set than that used here\*.

The vertical excitation energies,  $X^1\Sigma_g^+ \rightarrow A^3\Sigma_u^+$  and  $X^1\Sigma_g^+ \rightarrow B^3\Pi_g$  can be obtained from the potential curves. The calculated values are 6.33 and 8.99 eV, to be compared with the experimental values 6.2 and 8.1 eV, respectively [34]. The larger error for the X  $\rightarrow$  B excitation is partly explained by the error in the asymptotic energy differences. The  $^2D-^4S$  energy difference obtained here is 2.88 eV, which is close to the HF value 2.85 eV [44]. The experimental splitting is only 2.39 eV [34], and the difference is due to dynamical correlation effects which are not included in the present treatment. If, however, the  $^3\Pi_g$  potential curve is shifted to the correct asymptotic limit the vertical X  $\rightarrow$  B energy difference is lowered to 8.50 eV, in much closer agreement with experiment.

In addition, the CASSCF-7 calculations result in a potential maximum for the B state at about 2.0 Å (a small maximum is actually also obtained for the A state). The appearance of such maxima has been noted before for  $O_2$  [41] and CH [45], they are caused by a breakdown of spherical symmetry in the separated atoms – the calculations at the asymptotic limit are performed in linear rather than spherical symmetry. This results in an artificial lowering of the atomic energies caused by a contamination of  $d_\sigma$  into  $s_\sigma$  and different radial dependencies in  $2p_\sigma$  and  $2p_\pi$ . This is especially serious for atomic states with

\* A preliminary calculation on the X state using 9 active orbitals (including  $2\pi_u$ ) and the same basis set as before gives a  $D_e$  value of 9.05 eV.

$L$  quantum number different from zero, where the different  $M_L$  and  $M_S$  components are found to be non-degenerate [41]. There does not seem to be any simple way of avoiding this at the MCSCF level, even though the inclusion of singly-excited configurations (first-order configuration interaction) seems to reduce the splitting to very small values, and removes the maxima from the potential curves [41]. This problem has, however, to be borne in mind when the MCSCF method is used to calculate energy surfaces for chemical reactions, for which either the reactants or the products possess higher symmetry than the compound system.

## 8. Concluding remarks

The density matrix formulation of the super-CI MCSCF method presented in this report, reveals some of the difficulties inherent in this method. The matrix elements between the SX-states involve the third order density matrix of the reference function and therefore cannot be evaluated without encountering computational difficulties. On the other hand, the conventional technique where all singly-excited states from each of the configurations comprising the reference function are considered explicitly severely limits the number of such configurations. In this work an attempt has been made to resolve this problem by approximating the matrix elements. It has been shown that an approximation is possible which also, to a large extent, removes one of the other bottlenecks of MCSCF calculations, the four index transformation of molecular integrals. The price to be paid for this simplification is a decreased convergence rate. In the examples given here, however, the convergence is satisfactory and it is doubtful whether a full second-order process would be competitive. This may not always be the case, of course.

The super-CI technique is naturally not the only orbital optimization method which can be used in the CASSCF method. In a forthcoming paper it will be shown how it can be applied using a full second-order process: the Newton–Raphson method [25, 46].

The most important result obtained from the present formulation of the MCSCF method is

probably the negligible amount of time needed for the orbital optimization step and the application of the graphical unitary group approach to solution of the CI problem and the construction of the density matrices. As a result it is now possible to treat long MC expansions in a computationally efficient way.

Another important aspect is the use of a complete expansion in an active orbital subspace. This offers certain computational advantages, in that the energy is unaffected by unitary mixing of these orbitals. Further, this form of expansion has additional benefits when it is desired to treat states which are not the lowest of their symmetry type. Thus, assuming that the active space is chosen so that states lower in energy than the one in question are properly described, not only is the CASCI wavefunction orthogonal to the lower roots, but the SX wavefunction for the state in question is also orthogonal to these lower roots. Hence modification of the wavefunction using the SX natural orbitals should not improve any lower state at the expense of the desired state – in principle, no “constraining” of the orbital variations need be imposed. This feature is unique to the case for which the MC expansion is complete in the active space.

Finally, the benefits of such a complete expansion are not only computational. As in the SCF method the orbital concept again occupies a central position and it is possible to discuss the chemical problem entirely in terms of the molecular orbitals and their occupation numbers. Obviously it is not possible to include dynamical correlation effects into the CASSCF wavefunction, since this would require a very large active subspace. The CASSCF wavefunction is, however, an excellent zeroth-order approximation, and it can be expected that the remaining correlation effects will vary smoothly with the parameters which determine an energy hypersurface. An approximate treatment of these effects might therefore be sufficient. It is, for example, straightforward to include, to second order, the correlation energy arising from all primary to secondary double replacements into the CASSCF energy. A large fraction of the remaining correlation energy can probably be recovered by such a calculation.

## Acknowledgement

One of us (P.R.T.) wishes to acknowledge the support of a Commonwealth Scientific and Industrial Research Organization Postdoctoral

Fellowship during part of this work. This work was supported by a grant from the Swedish Natural Science Research Council.

## Appendix A: Expressions for the SX-state matrix elements

The matrix elements between the SX-states (17) given below are expressed in terms of the Fock operators defined in expressions (21) and (22), the density matrices over the active space, as defined in (13), and the two-electron integrals (4b). In addition a component of the third order density matrix, according to

$$Q_{rxyztu} = \frac{1}{2} \langle 0 | (\hat{E}_{r,x} \hat{E}_{y,z} - \delta_{xy} \hat{E}_{r,z}) \hat{E}_{tu} | 0 \rangle, \quad (\text{A.1})$$

is needed. The matrix elements are obtained using the definition (3) of the hamiltonian and the ‘‘contraction’’ formulae (7) for the generators. Defining

$$\hat{H}' = \hat{H} - \langle 0 | \hat{H} | 0 \rangle \quad (\text{A.2})$$

we now obtain,

$$\begin{aligned} \langle i \rightarrow t | \hat{H}' | j \rightarrow u \rangle (m_t m_u)^{1/2} &= \delta_{ij} \sum_r [2\delta_{tr} F_{uv}^i - D_{tr} F_{uv}^i - D_{ur} F_{tr}^i] + \delta_{ij} \sum_{v,x} [D_{tu} D_{vx} - 2P_{tuvx}] F_{vx}^i \\ &+ 2\delta_{ij} F_{tu}^a - (2\delta_{tu} - D_{tu}) F_{ij}^i - \sum_r (\delta_{ur} - D_{ur}) [(vt|ij) - 2(vj|ti)] - \sum_v (\delta_{tr} - D_{tr}) [(vu|ij) - 2(vi|uj)] \\ &- \sum_{r,x} \{2[\delta_{tu} D_{rx} - P_{tvr,x}] (vx|ij) - [\delta_{tu} D_{rx} + 2P_{tvr,x}] (vj|xi)\} \\ &- 2\delta_{ij} \sum_{r,x,y} P_{rxy} (vx|uy) + \delta_{ij} \sum_{r,x,y,z} (P_{rxyz} D_{tu} - Q_{rxyztu}) (vx|yz), \end{aligned} \quad (\text{A.3a})$$

$$\begin{aligned} \langle i \rightarrow t | \hat{H}' | j \rightarrow a \rangle (2m_t)^{1/2} &= 2\delta_{ij} F_{at} - \delta_{ij} \sum_r D_{tr} F_{ar}^i + \sum_r D_{tr} [(av|ij) - 2(aj|vi)] \\ &- 2(at|ij) - 2(aj|ti) - 2\delta_{ij} \sum_{r,x,y} P_{rxy} (av|xy), \end{aligned} \quad (\text{A.3b})$$

$$\langle i \rightarrow t | \hat{H}' | u \rightarrow a \rangle (m_t m_u)^{1/2} = - \sum_r D_{ur} [(at|vi) - 2(av|ti)] - 2 \sum_{r,x} P_{tvr,x} (ax|vi), \quad (\text{A.3c})$$

$$\langle i \rightarrow a | \hat{H}' | j \rightarrow b \rangle = \delta_{ij} F_{ab} - \delta_{ab} F_{ij} + 2(ai|bj) - (ab|ij), \quad (\text{A.3d})$$

$$\langle i \rightarrow a | \hat{H}' | t \rightarrow b \rangle (2n_t)^{1/2} = \sum_r D_{tr} [2(ai|bv) - (ab|iv)] - \delta_{ab} \sum_r D_{tr} F_{ri}^i - 2\delta_{ab} \sum_{r,x,y} P_{rxy} (vi|xy), \quad (\text{A.3e})$$

$$\begin{aligned} \langle t \rightarrow a | \hat{H}' | u \rightarrow b \rangle (n_t n_u)^{1/2} &= D_{tu} F_{ab}^i + \delta_{ab} \sum_{r,x} [2P_{tvr,x} - D_{tr} D_{rx}] F_{vx}^i \\ &+ \sum_{r,x} [2P_{tvr,x} (ab|vx) + 2P_{tvr,x} (ax|bv)] - 2\delta_{ab} \sum_{r,x,y} P_{urxy} (tv|xy) \\ &+ \delta_{ab} \sum_{r,x,y,z} [Q_{rxyztu} - D_{tu} P_{rxyz}] (vx|yz). \end{aligned} \quad (\text{A.3f})$$

These matrix elements can be approximated by assuming the first-order density matrix to be diagonal, and by using an occupation number approximation of the second- and third-order density matrices according to (26). The following simplified formulae are then obtained.

$$\begin{aligned} \langle i \rightarrow t | \hat{H}' | j \rightarrow u \rangle (m_t m_u)^{1/2} &= \delta_{ij} [\frac{1}{2} m_t m_u (1 - \delta_{tu}) + m_t \delta_{tu}] F_{tu} - \delta_{tu} m_t F_{ij} + \frac{1}{2} m_t m_u [2(ti|uj) - (tu|ij)] \\ &- \frac{1}{2} \delta_{tu} m_t n_t [(tt|ij) + (ti|tj)] + \frac{1}{4} \delta_{ij} [m_t m_u (1 - \delta_{tu}) + m_t \delta_{tu}] [n_t (tt|tu) + n_u (tu|uu)], \end{aligned} \quad (\text{A.4a})$$

$$\langle i \rightarrow t | \hat{H} | j \rightarrow a \rangle (2m_t)^{1/2} = \delta_{ij} m_t F_{at} + \frac{1}{2} n_t m_t (at|tt) + m_t [2(aj|ti) - (at|ij)], \quad (\text{A.4b})$$

$$\langle i \rightarrow t | \hat{H} | u \rightarrow a \rangle (m_t n_u)^{1/2} = \frac{1}{2} n_u m_t [2(au|ti) - (at|ui) + \delta_{tu} (at|ti)], \quad (\text{A.4c})$$

$$\langle i \rightarrow a | \hat{H} | j \rightarrow b \rangle = \delta_{ij} F_{ab} - \delta_{ab} F_{ij} + 2(ai|bj) - (ab|ij), \quad (\text{A.4d})$$

$$\langle i \rightarrow a | \hat{H} | t \rightarrow b \rangle (2n_t)^{1/2} = n_t [2(ai|bt) - (ab|ti) - \delta_{ab} F_{it}] + \frac{1}{2} \delta_{ab} m_t (tt|ti), \quad (\text{A.4e})$$

$$\begin{aligned} \langle t \rightarrow a | \hat{H}' | u \rightarrow b \rangle (n_t n_u)^{1/2} &= \delta_{tu} n_t F_{ab} - \delta_{ab} [\frac{1}{2} n_t n_u (1 - \delta_{tu}) + n_t \delta_{tu}] F_{tu} + \frac{1}{2} n_t n_u [2(at|bu) - (ab|tu)] \\ &\quad - \frac{1}{2} m_t n_t \delta_{tu} [(at|bt) + (ab|tt)] + \frac{1}{4} \delta_{ab} [n_t n_u (1 - \delta_{tu}) + n_t \delta_{tu}] [m_t (tt|tu) + m_u (tu|uu)]. \end{aligned} \quad (\text{A.4f})$$

In all these expressions  $n_t = D_{tt}$  and  $m_t = 2 - D_{tt}$ , and the index (0) has been omitted from the density matrix elements in order to simplify the notation.

## Appendix B: The super-CI density matrix elements

The elements of the super-CI density matrix are defined as

$$D_{pq}^{\text{SX}} = \langle 0 | \hat{E}_{\text{SX}}^\dagger \hat{E}_{pq} \hat{E}_{\text{SX}} | 0 \rangle \quad (\text{B.1})$$

with the super-CI generator  $\hat{E}_{\text{SX}}$  defined in (20). In order to simplify the notations the following unnormalized super-CI coefficients are used:

$$a_0 = c_0, \quad a_{it} = c_{it} m_t^{-1/2}, \quad a_{ia} = c_{ia} 2^{-1/2}, \quad a_{ta} = c_{ta} n_t^{-1/2}. \quad (\text{B.2})$$

The super-CI density matrix elements are obtained using the same technique as was used for the SX state matrix elements. Expressed in terms of the coefficients (B.2) and the first- and second-order density matrices (**D** and **P**) for the CAS wavefunction we obtain:

$$D_{ij}^{\text{SX}} = 2\delta_{ij} - 2 \sum_a a_{ia} a_{ja} - \sum_{t,u} a_{it} a_{ju} (2\delta_{tu} - D_{tu}), \quad (\text{B.3a})$$

$$D_{it}^{\text{SX}} = a_0 \sum_u a_{iu} (2\delta_{tu} - D_{tu}) - \sum_{a,u} a_{ia} a_{ua} D_{tu}, \quad (\text{B.3b})$$

$$D_{ia}^{\text{SX}} = 2a_0 a_{ia}, \quad (\text{B.3c})$$

$$D_{tu}^{\text{SX}} = D_{tu} + \sum_{v,x} (2P_{vxtu} - D_{vx} D_{tu}) \left[ \sum_a a_{va} a_{xa} + \sum_i a_{iv} a_{ix} \right] + \sum_{i,v} [a_{iv} a_{iu} (\delta_{tv} - D_{tv}) + a_{iv} a_{it} (\delta_{uv} - D_{uv})], \quad (\text{B.3d})$$

$$D_{ia}^{\text{SX}} = a_0 \sum_u a_{ua} D_{tu} + \sum_{i,u} a_{ia} a_{iu} (2\delta_{tu} - D_{tu}), \quad (\text{B.3e})$$

$$D_{ab}^{\text{SX}} = 2 \sum_i a_{ia} a_{ib} + \sum_{t,u} a_{ta} a_{ub} D_{tu}. \quad (\text{B.3f})$$

## References

- [1] F. Hund, A. Physik 51 (1928) 759.
- [2] R.S. Mulliken, Phys. Rev. 32 (1928) 168.
- [3] D.R. Hartree, Cambridge Phil. Soc. 24 (1928) 89.
- [4] V. Fock, Z. Physik 61 (1930) 126.
- [5] C.C.J. Roothaan, Rev. Mod. Phys. 23 (1951) 69; 32 (1960) 179.
- [6] P.O. Löwdin, Phys. Rev. 97 (1955) 1474.
- [7] C.F. Bender and E.R. Davidson, Phys. Rev. 139 (1969) 23.
- [8] W. Meyer, Intern. J. Quantum Chem. S5 (1971) 341.
- [9] R. Ahlrichs, H. Lischka, V. Staemmler and W. Kutzelnigg, J. Chem. Phys. 62 (1975) 1225.
- [10] D.R. Hartree, W. Hartree and B. Swirles, Phil. Trans. R. Soc. London Ser. A238 (1939) 299.
- [11] D.W. Smith and O.W. Day, J. Chem. Phys. 62 (1975) 113.
- [12] J. Hinze, J. Chem. Phys. 59 (1973) 6424.
- [13] A.C. Wahl and G. Das in: Methods of electronic



- structure theory, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
- [14] E. Dalgaard and P. Jørgensen, *J. Chem. Phys.* 69 (1978) 3833.
- [15] F. Grein and T.C. Chang, *Chem. Phys. Letters* 12 (1971) 44.
- [16] F. Grein and A. Banerjee, *Intern. J. Quantum Chem.* S9 (1975) 147; *J. Chem. Phys.* 66 (1977) 1054.
- [17] L. Brillouin, *Act. Sci. Ind.* 71 (1933) 159.
- [18] B. Levy and G. Berthier, *Intern. J. Quantum Chem.* 2 (1968) 307.
- [19] L.M. Cheung, S.T. Elbert and K. Ruedenberg, *Intern. J. Quantum Chem.* 14 (1979) 1069.
- [20] S.T. Elbert, private communication.
- [21] K. Ruedenberg and K.R. Sundberg, in: *Quantum science, methods and structure*, eds. J.-C. Calais, O. Goscinski, J. Linderberg and Y. Öhrn (Plenum Press, New York, 1976).
- [22] I. Shavitt, *Intern. J. Quantum Chem.* S11 (1977) 131; S12 (1978) 5.
- [23] J. Paldus, in: *Theoretical chemistry: advances and perspectives*, eds. H. Eyring and D. Henderson, Vol. 2 (Academic Press, New York, 1976) p. 131.
- [24] B. Roos, *Chem. Phys. Letters* 15 (1972) 153; B. Roos and P. Siegbahn, in: *Methods of electronic structure theory*, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
- [25] P. Siegbahn, A. Heiberg, B. Roos and B. Levy, *Phys. Scr.* 21 (1980) 323.
- [26] F.B. Van Duijneveldt, IBM Research Report RJ945 (1971).
- [27] M.F. Guest and V.R. Saunders, *Mol. Phys.* 28 (1974) 819.
- [28] B. Jönsson, B. Roos, P. Siegbahn and P. Taylor, to be published.
- [29] P.C. de Mello, M. Hehenberger, S. Larsson and M. Zerner, *Inorg. Chem.*, to be published.
- [30] A.J.H. Wachter, *J. Chem. Phys.* 52 (1970) 1033.
- [31] T.H. Dunning, *J. Chem. Phys.* 53 (1970) 2823.
- [32] D.L. Yeager and P. Jørgensen, *J. Chem. Phys.* 71 (1979) 755.
- [33] M. Krauss and D.B. Neumann, *Mol. Phys.* 32 (1976) 101.
- [34] A. Lofthus and P.H. Krupenie, *J. Phys. Chem. Ref. Data* 6 (1977) 113.
- [35] G. Herzberg, *Spectra of diatomic molecules* (Van Nostrand, Princeton, 1950).
- [36] J.L. Dunham, *Phys. Rev.* 41 (1932) 721.
- [37] R. Ahlrichs, H. Lischka, B. Zurawski and W. Kutzelnigg, *J. Chem. Phys.* 63 (1975) 4685.
- [38] W. Butscher, S.-K. Shih, R.J. Buenker and S.D. Peyerimhoff, *Chem. Phys. Letters* 52 (1978) 457.
- [39] G.C. Lie and E. Clementi, *J. Chem. Phys.* 60 (1974) 1288.
- [40] F.P. Billingsley II and M. Krauss, *J. Chem. Phys.* 60 (1974) 4130.
- [41] R.P. Saxon and B. Liu, *J. Chem. Phys.* 67 (1977) 5432.
- [42] P.R. Taylor, to be published.
- [43] A.C. Hurley, *Introduction to the electron theory of small molecules* (Academic Press, New York, 1976).
- [44] E. Clementi and C. Roetti, *At. Data Nucl. Data Tables* 14 (1974) 177.
- [45] G.C. Lie, J. Hinze and B. Liu, *J. Chem. Phys.* 59 (1973) 1872.
- [46] P. Siegbahn, A. Heiberg and B. Roos, to be published
- [47] M. Moshinsky, *Group theory and the many-body problem* (Gordon and Breach, New York, 1968).
- [48] I. Shavitt, Technical Report (Batelle Columbus Laboratories, 1979).
- [49] B.R. Brooks and H.F. Schaefer III, *J. Chem. Phys.* 70 (1979) 5092.
- [50] P.E.M. Siegbahn, *J. Chem. Phys.* 72 (1980) 1647.
- [51] K.P. Huber and G. Herzberg, *Constants of diatomic molecules* (Van Nostrand, Princeton, 1979).