

Molecular Hessians for large-scale MCSCF wave functions

Trygve U. Helgaker and Jan Almlöf^{a)}

Department of Chemistry, University of Oslo, Blindern, N-0315 Oslo 3, Norway

Hans Jørgen Aa. Jensen and Poul Jørgensen

Department of Chemistry, Aarhus University, DK-8000 Århus C, Denmark

(Received 7 August 1985; accepted 19 February 1986)

The calculation of molecular Hessians for large-scale multiconfiguration self-consistent-field (MCSCF) functions is described. The formalism is based on exponential parametrization of the wave function and symmetric orthonormalization of the molecular orbitals. Extensive use is made of one-index transformations of the molecular integrals, both to construct the gradient vectors that appear in the linear MCSCF response equations, and to perform the multiplication of the trial vectors on the electronic Hessian in the iterative, direct solution of the response equations. No element of the electronic Hessian is ever calculated explicitly, allowing for use of large configuration expansions. Efficient methods are developed for obtaining the solution vectors of the linear response equations. The accuracy of the molecular Hessian is analyzed in terms of the accuracy of these solution vectors. To allow for large basis sets Fock matrices are used to minimize transformations and integrals are recalculated to minimize storage requirements. Integral derivatives are calculated following the McMurchie–Davidson scheme. A simplified algorithm for calculation of derivatives of integrals involving one-center overlap distributions is described. Sample calculations involving several thousand configurations are reported.

I. INTRODUCTION

The study of Born–Oppenheimer potential energy surfaces is an important and challenging task in quantum chemistry. The systematic study of such surfaces requires the capability to locate and characterize stationary points efficiently. Much effort has therefore been directed towards the development of efficient algorithms for the evaluation of molecular gradients and Hessians using quantum-chemical approximate wave functions.

Early work in this field focused on the calculation of gradients¹ (and later Hessians^{2–4}) for single-configuration self-consistent-field (SCF) wave functions. However, due to the limitations inherent in the Hartree–Fock approximation, such techniques are mainly useful for studying potential surfaces at or close to the equilibrium conformation. An adequate description of the electronic structure far from equilibrium normally requires the use of correlated wave functions. Such wave functions are now routinely calculated in many laboratories.^{5–7} Consequently there is a need to develop efficient techniques for the calculation of molecular gradients as well as Hessians for state-of-the-art correlated wave functions.

In the pioneering work on molecular gradients by Pulay^{1,8} the theory was worked out in the language of Fock operators. This approach was later used in deriving general expressions for the molecular Hessian of a multiconfiguration (MC) SCF wave function.⁹ A Fock-operator based algorithm was also used by Yamaguchi *et al.*¹⁰ who carried out the first MCSCF molecular Hessian calculation using a two-configuration wave function.

A disadvantage of the Fock-operator technique is that it

requires the MO coefficients to be optimized subject to orthonormality conditions, a requirement that makes the formalism rather nontransparent. To avoid this, an exponential parametrization^{11–13} of the orbital rotations has often been used. In this formalism the allowed orbital variations are expressed in terms of independent variables and the need for a complicated handling of Lagrangian multipliers is eliminated.

The Fock-operator formalism may furthermore be cumbersome in the sense that each type of approximate wave function must be treated separately. This problem may be overcome in the exponential parametrization scheme by transferring the geometry dependence of the basis set to the Hamiltonian, expressing the geometrical expansion of the Hamiltonian in terms of symmetrically orthonormalized orbitals.^{14,15} The explicit form of the geometry dependence then becomes independent of the wave function, and standard response function theory may be applied to evaluate molecular gradients and Hessians¹⁴ as well as higher derivatives¹⁶ for any approximate wave function.

Only a few calculations of MCSCF Hessians, using either of the above-mentioned approaches, have so far been reported.^{17–19} Our approach differs from earlier attempts mainly in the handling of the linear response equations and in the treatment of the integral derivatives.

In our implementation the response equations are solved iteratively in a way that permits the use of large basis sets as well as large configuration expansions. This is accomplished by performing the multiplication of trial vectors on the electronic Hessian in a direct fashion (“direct MCSCF”²⁰), bypassing the calculation of individual elements of the electronic Hessian. The response equations are furthermore solved simultaneously, i.e., the solution vectors for all independent Cartesian displacements are expanded in

^{a)} Present address: Department of Chemistry, University of Minnesota, 207 Pleasant Street S. E., Minneapolis, MN 55455.

a common set of trial vectors. As demonstrated in this paper, this implies that the convergence error in the molecular Hessian is quadratic in the error of the individual solution vectors, greatly reducing the number of iterations needed to calculate the molecular Hessian to a given numerical accuracy. This is especially important for correlated wave functions, as in these cases the time spent on the solution of the response equations is likely to be considerable.

The atomic integrals are calculated following the scheme devised by McMurchie and Davidson,²¹ based on the earlier work of Boys.²² Certain simplifications have been introduced in the calculation of derivative integrals involving one-center overlap distributions, leading to reductions in the computer time for small and medium-sized molecules. To reduce the requirements on external storage the first derivatives of the two-electron integrals are calculated, written on disk and processed with respect to three Cartesian displacements (one atom) at a time. In this way the disk requirements for molecular Hessian calculations are not significantly larger than for the calculation of the wave function itself.

Sample calculations presented in this paper indicate that the time and effort needed to calculate MCSCF molecular Hessians are comparable to that of obtaining the MCSCF wave function in the first place. In other words, the MCSCF molecular Hessian may be calculated whenever the electronic structure calculation itself is practical.

In the following section we outline the mathematical formalism and give general expressions for the MCSCF molecular gradient and Hessian. Section III contains a detailed description of the actual implementation of those expressions and in Sec. IV we report numerical results to demonstrate the potential of the response function approach. The last section contains some concluding remarks.

II. FORMALISM

Let us first briefly state the purpose of this section. We wish to define a reference electronic wave function which is variationally optimal at some reference geometry but which is also defined at other geometries as an explicit function of the geometry. In terms of this wave function we may generate any wave function within the given orbital and configuration space by unitary transformations of molecular orbitals and configuration states. In particular, we are able to define the optimal MCSCF wave function at each molecular geometry and hence the optimal electronic energy as an explicit function of the nuclear coordinates.

Consider a molecular system in the neighborhood of some reference geometry \mathbf{X}_0 . At the reference geometry the molecular electronic state is described by the reference wave function (RWF):

$$|\text{RWF}\rangle = \sum_{\mu} C_{\mu} |\mu\rangle, \quad (1)$$

where each electronic configuration $|\mu\rangle$ is a Slater determinant,

$$|\mu\rangle = \prod_p \hat{a}_{\mu_p}^+ |\text{vac}\rangle \quad (2)$$

which may be projected to ensure correct spin and/or spatial symmetry.

Each determinant is constructed from a finite set of orthonormal molecular orbitals $\{\phi_p\}$ which are expanded in a set of atomic orbitals fixed on the nuclei

$$\phi_p(\mathbf{X}_0) = \sum_{\mu} C_{p\mu}^0 \chi_{\mu}(\mathbf{X}_0). \quad (3)$$

These orbitals may be divided into three groups—inactive, active, and secondary. The inactive orbitals are doubly occupied in all configurations, while the remaining electrons are distributed among the active orbitals, leaving the secondary orbitals empty in all configurations. We follow the notation of Siegbahn *et al.*²³ and denote inactive orbitals by $\{i, j, k, l\}$, active by $\{t, u, v, x\}$, and secondary by $\{a, b, c, d\}$. Orbitals of a general or unspecified type are referred to by $\{p, q, r, s\}$ or $\{\alpha, \beta, \gamma, \delta\}$. Finally we use $\{\mu, \nu, \lambda, \sigma\}$ to denote atomic orbitals.

The RWF is determined to give the variationally best approximation to the exact wave function at the reference geometry. We may extend its definition to other geometries by introducing the symmetrically orthonormalized molecular orbitals¹⁵ (OMO's)

$$\psi_p(\mathbf{X}) = \sum_q T_{pq}(\mathbf{X}) \phi_q(\mathbf{X}), \quad (4)$$

where

$$T_{pq} \equiv (\mathbf{S}^{-1/2})_{pq} \quad (5)$$

and

$$\begin{aligned} S_{pq}(\mathbf{X}) &= \langle \phi_p(\mathbf{X}) | \phi_q(\mathbf{X}) \rangle \\ &= \sum_{\mu\nu} C_{p\mu}^0 C_{q\nu}^0 \langle \chi_{\mu}(\mathbf{X}) | \chi_{\nu}(\mathbf{X}) \rangle \end{aligned} \quad (6)$$

is the overlap between the unmodified molecular orbitals (UMO's) evaluated at \mathbf{X} . In terms of these orthonormalized orbitals the creation operators and configurations of Eq. (2) are defined at all geometries:

$$|\mu(\mathbf{X})\rangle = \prod_p \hat{a}_{\mu_p}^+(\mathbf{X}) |\text{vac}\rangle \quad (7)$$

and their geometry dependence is such that all inner products $\langle \mu(\mathbf{X}) | \nu(\mathbf{X}) \rangle$ are preserved. By the same token all (transition) expectation values of strings of creation and annihilation operators are preserved.^{14,15} In particular the configuration interaction (CI) coupling coefficients

$$\begin{aligned} A_{pq}^{\mu\nu} &= \langle \mu(\mathbf{X}) | \hat{E}_{pq}(\mathbf{X}) | \nu(\mathbf{X}) \rangle, \\ A_{pqrs}^{\mu\nu} &= \langle \mu(\mathbf{X}) | \hat{E}_{pq}(\mathbf{X}) \hat{E}_{rs}(\mathbf{X}) - \delta_{qr} \hat{E}_{ps}(\mathbf{X}) | \nu(\mathbf{X}) \rangle \end{aligned} \quad (8)$$

are geometry independent. In Eq. (8) we have introduced the generators of the unitary group

$$\hat{E}_{pq}(\mathbf{X}) = \sum_{\sigma} \hat{a}_{p\sigma}^+(\mathbf{X}) \hat{a}_{q\sigma}(\mathbf{X}), \quad (9)$$

where σ labels spin. As the configurations [Eq. (7)] and operators [Eq. (9)] in the final equations always occur in (transition) expectation value expressions, the geometry dependence of the configurations and the operators does not enter the computational expressions and may therefore be disregarded. We may now extend the definition of the reference wave function [Eq. (1)] to all geometries

$$|\text{RWF}(\mathbf{X})\rangle = \sum_{\mu} C_{\mu} |\mu(\mathbf{X})\rangle. \quad (10)$$

Note that although the RWF is defined at all geometries, it is variationally optimal at the reference geometry only. At any given geometry we are now in a position to express an arbitrary wave function in terms of unitary operators working on the RWF²⁴:

$$|\text{WF}\rangle = \exp(-\hat{\kappa})\exp(-\hat{S})|\text{RWF}(\mathbf{X})\rangle. \quad (11)$$

In this expression we have introduced the antisymmetric operator

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq} \hat{E}_{pq}^-, \quad (12)$$

$$\hat{E}_{pq}^- = \hat{E}_{pq} - \hat{E}_{qp}$$

which describes orthogonal rotations in the orbital space and the antisymmetric operator

$$\hat{S} = \sum_{K>0} S_K \hat{P}_K, \quad (13)$$

$$\hat{P}_K = |\text{RWF}\rangle\langle K| - |K\rangle\langle\text{RWF}|$$

which describes orthogonal rotations in the configuration space. In Eq. (13) $\{|K(\mathbf{X})\rangle\}$ is some basis for the orthogonal complement to $|\text{RWF}(\mathbf{X})\rangle$. The operators in Eqs. (12) and (13) should only contain nonredundant variations. For CASSCF wave functions (full CI within the active space) this is easily achieved by including only inactive-active, inactive-secondary, and active-secondary rotations in the orbital rotation operator. For general MCSCF wave functions the redundancy problem is more complicated.^{18,24}

The Hamiltonian may be expressed in the OMO representation as

$$\hat{H}(\mathbf{X}) = \sum_{pq} h_{pq}^{\text{OMO}}(\mathbf{X}) \hat{E}_{pq}^- + \frac{1}{2} \sum_{pqrs} g_{pqrs}^{\text{OMO}}(\mathbf{X}) (\hat{E}_{pq}^- \hat{E}_{rs}^- - \delta_{rq} \hat{E}_{ps}^-). \quad (14)$$

The integrals in this Hamiltonian may further be expressed in the unmodified molecular orbital basis

$$h_{pq}^{\text{OMO}}(\mathbf{X}) = \sum_{\alpha\beta} h_{\alpha\beta}(\mathbf{X}) T_{p\alpha}(\mathbf{X}) T_{q\beta}(\mathbf{X}), \quad (15a)$$

$$g_{pqrs}^{\text{OMO}}(\mathbf{X}) = \sum_{\alpha\beta\gamma\delta} g_{\alpha\beta\gamma\delta}(\mathbf{X}) T_{p\alpha}(\mathbf{X}) T_{q\beta}(\mathbf{X}) T_{r\gamma}(\mathbf{X}) T_{s\delta}(\mathbf{X}). \quad (15b)$$

By differentiating Eqs. (15a) and (15b) with respect to coordinate X_a we obtain (denoting differentiation by superscripts):

$$\frac{d}{dX_a} (h_{pq}^{\text{OMO}}) = (h_{pq}^{\text{OMO}})^a = h_{pq}^a + \{\text{T}^a, h\}_{pq}, \quad (16a)$$

$$\frac{d}{dX_a} (g_{pqrs}^{\text{OMO}}) = (g_{pqrs}^{\text{OMO}})^a = g_{pqrs}^a + \{\text{T}^a, g\}_{pqrs}, \quad (16b)$$

where we have introduced the one-index transformations¹⁵:

$$\{\text{T}^a, h\}_{pq} = \sum_{\alpha} (T_{p\alpha}^a h_{\alpha q} + T_{q\alpha}^a h_{p\alpha}), \quad (17a)$$

$$\{\text{T}^a, g\}_{pqrs} = \sum_{\alpha} (T_{p\alpha}^a g_{\alpha qrs} + T_{q\alpha}^a g_{p\alpha rs} + T_{r\alpha}^a g_{pq\alpha s} + T_{s\alpha}^a g_{pqra}). \quad (17b)$$

By generalizing this procedure we may also express higher derivatives in terms of UMO-integral derivatives and one-index transformations of these integrals.¹⁶

We have now defined the wave function and the Hamiltonian at all geometries and are therefore equipped to study the molecular energy at an arbitrary geometry using standard methods. The total energy

$$\mathcal{E}(\mathbf{X}, \kappa, \mathbf{S}) = \langle \text{WF} | \hat{H}(\mathbf{X}) | \text{WF} \rangle = \langle \text{RWF} | \exp(\hat{S}) \exp(\hat{\kappa}) \hat{H}(\mathbf{X}) \times \exp(-\hat{\kappa}) \exp(-\hat{S}) | \text{RWF} \rangle \quad (18)$$

is written as a Baker–Campbell–Hausdorff expansion in the antisymmetric operators $\hat{\kappa}$ and \hat{S} . After collecting the orbital and state variational parameters into a single column vector λ , we obtain the Taylor expansion²⁴

$$\mathcal{E}(\mathbf{X}, \lambda) = E(\mathbf{X}) + \tilde{\mathbf{f}}(\mathbf{X}) \lambda + \frac{1}{2} \tilde{\mathbf{G}}(\mathbf{X}) \lambda + O(\lambda^3). \quad (19)$$

We have here introduced the electronic energy of the RWF (the “static” energy)

$$E(\mathbf{X}) = \langle \text{RWF} | \hat{H}(\mathbf{X}) | \text{RWF} \rangle, \quad (20)$$

the gradient $\mathbf{f}(\mathbf{X})$, and the Hessian $\mathbf{G}(\mathbf{X})$. Explicit expressions for the MCSCF gradient and the MCSCF Hessian will be given later.

For a fixed geometry \mathbf{X} we notice that Eq. (19) is the standard Taylor expansion used in second-order MCSCF optimization. In accordance with the Born–Oppenheimer approximation we require the electronic energy to be stationary with respect to all orbital and state variations at \mathbf{X} :

$$\left(\frac{\partial}{\partial \lambda} \right) \mathcal{E}(\mathbf{X}, \lambda) = \mathbf{f}(\mathbf{X}) + \mathbf{G}(\mathbf{X}) \lambda + O(\lambda^2) = \mathbf{0}. \quad (21)$$

Recall that the RWF was defined to be the optimal MCSCF wave function describing the electronic state of interest at the reference geometry \mathbf{X}_0 , that is

$$\mathbf{f}(\mathbf{X}_0) = \mathbf{0}, \quad (22)$$

and $\lambda = \mathbf{0}$ is the desired solution of Eq. (21) for $\mathbf{X} = \mathbf{X}_0$. Equation (21) thus determines the MCSCF wave function and consequently the MCSCF parameters λ as functions of the geometry \mathbf{X} . Using Eq. (21) we can eliminate the dependent parameters λ from Eq. (19) and obtain the potential energy as a function of the free (independent) parameters:

$$\mathcal{E}(\mathbf{X}) = E(\mathbf{X}) - \frac{1}{2} \tilde{\mathbf{f}}(\mathbf{X}) \mathbf{G}^{-1}(\mathbf{X}) \mathbf{f}(\mathbf{X}) + O(\mathbf{f}^3). \quad (23)$$

All geometry dependence is now explicit and, as discussed previously, confined to the Hamiltonian integrals $h_{pq}^{\text{OMO}}(\mathbf{X})$ and $g_{pqrs}^{\text{OMO}}(\mathbf{X})$ of Eq. (14). Because of this explicit dependence we can now obtain the molecular gradient and Hessian by straightforward differentiation of Eq. (23). At the reference geometry simplifications occur due to Eq. (22), and we obtain

$$\mathcal{E}^a = E^a, \quad (24a)$$

$$\mathcal{E}^{ab} = E^{ab} - \tilde{\mathbf{f}}^a \mathbf{G}^{-1} \mathbf{f}^b. \quad (24b)$$

Note that the gradient contains only a static contribution while the Hessian contains both a static contribution and a

contribution due to the relaxation of the electronic wave function, i.e., due to the response of the electrons to changes in the nuclear coordinates in accordance with the Born–Oppenheimer approximation.

III. IMPLEMENTATION

In the previous section we gave expressions for the molecular gradient and Hessian in a symmetrically orthonormalized molecular orbital representation. We now consider the practical evaluation of these expressions. This section is divided into five parts. We first describe the calculation of the static contribution to the molecular gradient and Hessian. After we have discussed the evaluation of the derivatives of the Fock matrix, we treat the relaxation part of the Hessian with emphasis on techniques for solving the linear equations involved. We then consider the integral derivative evaluation in some detail and finally give a short overview of the computer code ABACUS written on the basis of the algorithms and formulas given in this work.

A. The molecular gradient and the static contribution to the molecular Hessian

In the OMO representation the static energy is

$$E(\mathbf{X}) = \langle \text{RWF} | \hat{H}(\mathbf{X}) | \text{RWF} \rangle \\ = \sum_{pq} D_{pq} h_{pq}^{\text{OMO}}(\mathbf{X}) + \sum_{pqrs} P_{pqrs} g_{pqrs}^{\text{OMO}}(\mathbf{X}), \quad (25)$$

where the integrals are given by Eq. (15) and the geometry-independent density matrices are

$$D_{pq} = \langle \text{RWF} | \hat{E}_{pq} | \text{RWF} \rangle, \quad (26a)$$

$$P_{pqrs} = \frac{1}{2} \langle \text{RWF} | \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} | \text{RWF} \rangle. \quad (26b)$$

Expressions for the molecular gradient and the static contribution to the molecular Hessian are obtained by straightforward differentiation of Eq. (25), keeping in mind that the density elements are constant. In Appendix A a detailed derivation of these static contributions is carried out. The final results may be written as

$$E^a = E^{ua} + E^{sa}, \quad (27a)$$

$$E^{ab} = E^{uab} + E^{sab}. \quad (27b)$$

Here the superscript *u* denotes differentiation of the unmodified Hamiltonian integrals while *s* denotes the remaining terms involving the derivatives of the overlap matrices, see Eq. (15).

The first contributions to the static part of the derivatives Eq. (27) are

$$E^{ua} = \sum_{pq} D_{pq} h_{pq}^a + \sum_{pqrs} P_{pqrs} g_{pqrs}^a, \quad (28a)$$

$$E^{uab} = \sum_{pq} D_{pq} h_{pq}^{ab} + \sum_{pqrs} P_{pqrs} g_{pqrs}^{ab}. \quad (28b)$$

These expressions are best evaluated in the AO basis by contracting the differentiated AO integrals with AO density elements as soon as the integrals have been calculated.¹ In this way one may avoid storing the integrals on disk making this

part of the program very efficient in terms of I/O. The first derivatives of the two-electron integrals will later be recalculated to construct other contributions to the molecular Hessian, as discussed in Sec. III E.

The second contribution to the static part of the molecular derivatives may be written as¹⁵

$$E^{sa} = -\text{tr } S^a F, \quad (29a)$$

$$E^{sab} = -\text{tr } S^{ab} F - \text{tr}(S^a Y^b + S^b Y^a), \quad (29b)$$

where the generalized Fock matrix²³

$$F_{pq} = \sum_{\beta} D_{p\beta} h_{q\beta}^{\text{OMO}} + 2 \sum_{\beta\gamma\delta} P_{p\beta\gamma\delta} g_{q\beta\gamma\delta}^{\text{OMO}} \quad (30)$$

and the auxiliary matrix

$$Y^a = F^{ua} + \frac{1}{2} F^{sa} - \frac{1}{2} S^a F \quad (31)$$

have been introduced. The calculation of the individual contributions to the derivative of the Fock matrix

$$F^a = F^{ua} + F^{sa} \quad (32)$$

will be discussed in Sec. III B. Returning to Eq. (29) we note that E^{sa} and the first part of E^{sab} may be calculated in the AO basis at the same time as the one-electron contributions to E^{ua} and E^{uab} are calculated, with overlaps replacing the one-electron Hamiltonian integrals and the Fock matrix replacing the one-electron density matrix. The second contribution to E^{sab} involves the matrices S^a and Y^a , both of which are readily obtained in the MO basis as by-products from the calculation of the relaxation part of the Hessian.

B. The Fock matrix and its derivatives

The construction of the Fock matrix and its derivatives may be simplified considerably by partitioning the orbital space. As shown by Siegbahn *et al.*²³ the Fock matrix may be written as

$$F_{iq} = 2({}^c F_{iq} + {}^v F_{iq}), \\ F_{iq} = \sum_u D_{iu} {}^c F_{qu} + Q_{iq}, \\ F_{aq} = 0, \quad (33)$$

where we have introduced the inactive Fock matrix

$${}^c F_{pq} = h_{pq}^{\text{OMO}} + \sum_i (2g_{pqii}^{\text{OMO}} - g_{piqi}^{\text{OMO}}), \quad (34)$$

the active Fock matrix

$${}^v F_{pq} = \sum_{tu} D_{tu} (g_{pqtu}^{\text{OMO}} - \frac{1}{2} g_{ptiqu}^{\text{OMO}}), \quad (35)$$

and the auxiliary matrix

$$Q_{iq} = 2 \sum_{uwx} P_{iwx} g_{quwx}^{\text{OMO}}. \quad (36)$$

In terms of the \mathcal{P} supermatrix²⁵

$$\mathcal{P}_{pqrs}^{\text{OMO}} = g_{pqrs}^{\text{OMO}} - \frac{1}{4} g_{prqs}^{\text{OMO}} - \frac{1}{4} g_{psrq}^{\text{OMO}}, \quad (37)$$

the inactive and active Fock matrices may be written

$${}^c F_{pq} = h_{pq}^{\text{OMO}} + 2 \sum_i \mathcal{P}_{pqii}^{\text{OMO}}, \\ {}^v F_{pq} = \sum_{tu} D_{tu} \mathcal{P}_{pqtu}^{\text{OMO}}. \quad (38)$$

These matrices may be evaluated very efficiently in the AO basis:

$$\begin{aligned} {}^cF_{\mu\nu} &= h_{\mu\nu} + \sum_{\lambda\sigma} \mathcal{P}_{\mu\nu\lambda\sigma} {}^cD_{\lambda\sigma}, \\ {}^vF_{\mu\nu} &= \sum_{\lambda\sigma} \mathcal{P}_{\mu\nu\lambda\sigma} {}^vD_{\lambda\sigma}, \end{aligned} \quad (39)$$

where the AO density matrices are

$$\begin{aligned} {}^cD_{\lambda\sigma} &= 2 \sum_I C_{i\lambda}^0 C_{i\sigma}^0, \\ {}^vD_{\lambda\sigma} &= \sum_{tu} D_{tu} C_{i\lambda}^0 C_{u\sigma}^0, \end{aligned} \quad (40)$$

and the MO coefficients $\{C_{i\lambda}^0\}$ are defined in Eq. (3). The final MO matrices are obtained by transforming the Fock matrices to the MO basis. Note that the construction of the \mathcal{P} supermatrix is advantageous only when the Fock matrices are needed several times with different density elements, for example in iterative optimization procedures. The Q matrix involves the two-electron density matrix and is therefore best constructed in the MO basis according to Eq. (36). This, of course, requires the transformation of the two-electron integrals to MO basis with one general and three active indices.

We now turn to the evaluation of the derivatives of the matrices ${}^cF, {}^vF$, and Q in Eqs. (34)–(36). By differentiating we obtain

$$\begin{aligned} {}^cF^a &= {}^cF^{ua} + {}^cF^{sa}, \\ {}^vF^a &= {}^vF^{ua} + {}^vF^{sa}, \\ Q^a &= Q^{ua} + Q^{sa}, \end{aligned} \quad (41)$$

where the u terms involve the derivatives of the UMO Hamiltonian integrals, and the s terms the derivatives of the overlap matrix. The u contributions to these matrices may be evaluated in exactly the same way as the undifferentiated matrices. However, the use of the \mathcal{P} supermatrix [Eq. (37)] is no longer advisable as we only need to construct the derivative matrices once. For this reason we use a code that sets these matrices up in AO basis directly from the derivative AO integrals.

By differentiating the overlap matrices we find that the s contributions to the inactive and active Fock matrices are

$$\begin{aligned} {}^cF_{pq}^{sa} &= \sum_{\alpha} ({}^cF_{\alpha q} T_{p\alpha}^a + {}^cF_{p\alpha} T_{q\alpha}^a) \\ &+ {}^cW_{pq}^a = \{T^a, {}^cF\}_{pq} + {}^cW_{pq}^a, \\ {}^vF_{pq}^{sa} &= \sum_{\alpha} ({}^vF_{\alpha q} T_{p\alpha}^a + {}^vF_{p\alpha} T_{q\alpha}^a) \\ &+ {}^vW_{pq}^a = \{T^a, {}^vF\}_{pq} + {}^vW_{pq}^a, \end{aligned} \quad (42)$$

where we have introduced

$$\begin{aligned} {}^cW_{pq}^a &= \sum_{ri} \mathcal{P}_{pqri}^{\text{OMO}} V_{ri}^a, \\ {}^vW_{pq}^a &= \sum_{ri} \mathcal{P}_{pqri}^{\text{OMO}} V_{ri}^a, \end{aligned} \quad (43)$$

and used

$$V_{pq}^a = - \sum_{\alpha} S_{p\alpha}^a D_{\alpha q} \quad (44)$$

to simplify the expressions. The first contributions to the derivative Fock matrices $\{T^a, {}^cF\}$ and $\{T^a, {}^vF\}$ are easily calculated in the MO basis by one-index transforming the corresponding undifferentiated matrices. The last contributions are most easily calculated in AO basis:

$$\begin{aligned} {}^cW_{\mu\nu}^a &= \sum_{\lambda\sigma} \mathcal{P}_{\mu\nu\lambda\sigma} {}^cV_{\lambda\sigma}^a, \\ {}^vW_{\mu\nu}^a &= \sum_{\lambda\sigma} \mathcal{P}_{\mu\nu\lambda\sigma} {}^vV_{\lambda\sigma}^a, \end{aligned} \quad (45)$$

where

$$\begin{aligned} {}^cV_{\lambda\sigma}^a &= \sum_{ri} C_{r\lambda}^0 C_{i\sigma}^0 V_{ri}^a, \\ {}^vV_{\lambda\sigma}^a &= \sum_{ri} C_{r\lambda}^0 C_{i\sigma}^0 V_{ri}^a, \end{aligned} \quad (46)$$

and then transformed to MO basis. We thus see that it is possible to construct the derivatives of the inactive and active Fock matrices without transforming the two-electron integrals to MO basis. The derivatives of the Q matrix, on the other hand, are most efficiently calculated in the MO basis. Following Eq. (16b) we first calculate the derivatives of the UMO integrals g^a and the one-index transformed integrals $\{T^a, g\}$. Once these elements are available we may construct Q^{ua} and Q^{sa} in exactly the same way as we construct the original matrix Q from the undifferentiated integrals g , see Eq. (36).

C. The relaxation contribution to the molecular Hessian

The relaxation contribution to the molecular Hessian is the second term of Eq. (24b),

$$R^{ab} = -\tilde{f}^a G^{-1} f^b. \quad (47)$$

This expression may be written as

$$R^{ab} = \tilde{f}^b \lambda^a, \quad (48)$$

where λ^a is the solution vector to the MCSCF response equations

$$G \lambda^a = -\tilde{f}^a. \quad (49)$$

These equations are expressed in the basis of the orthogonal complement set of states $\{|K\rangle\}$ to the MCSCF state $|\text{RWF}\rangle$. A more convenient representation for large CSF spaces is the CSF basis $\{|\mu\rangle\}$. Following Lengsfeld and Liu²⁶ we write

$$\begin{pmatrix} {}^{cc}G & {}^{co}G \\ {}^{oc}G & {}^{oo}G \end{pmatrix} \begin{pmatrix} c\lambda^a \\ o\lambda^a \end{pmatrix} = - \begin{pmatrix} c f^a \\ o f^a \end{pmatrix}, \quad (50)$$

where¹⁵

$$c f_{\mu}^a = 2(\langle \text{RWF} | \hat{H}^a | \mu \rangle - C_{\mu} \mathcal{E}^a), \quad (51a)$$

$$o f_{pq}^a = 2(F_{pq}^a - F_{qp}^a), \quad (51b)$$

and

$${}^{cc}G_{\mu,\nu} = 2(\langle \mu | \hat{H} | \nu \rangle - \delta_{\mu\nu} \mathcal{E}), \quad (52a)$$

$${}^{co}G_{\mu,pq} = 2\langle \text{RWF} | [\hat{E}_{pq}^-, \hat{H}] | \mu \rangle, \quad (52b)$$

$${}^{oo}G_{pq,rs} = \langle \text{RWF} | [\hat{E}_{pq}^-, [\hat{E}_{rs}^-, \hat{H}]] | \text{RWF} \rangle. \quad (52c)$$

The superscripts c and o are used to denote the CSF and orbital parts, respectively, of the differentiated gradient f^a and the electronic Hessian G . Although not obvious from the expression (52c), ${}^{oo}G$ is symmetric. This follows from the fact that the MCSCF orbital gradient ${}^o f$ vanishes at the reference geometry [Eq. (22)].

The above linear equations (49) and (50) are identical in structure to the Newton–Raphson equations used for second-order optimization of MCSCF wave functions.²⁷ In fact, the Newton–Raphson equations may be obtained from the response equations simply by replacing the differentiated gradient f^a with the MCSCF gradient f . This means that we may use the same techniques—indeed the same computer code—to solve both the Newton–Raphson equations and the response equations. Accordingly, whenever it is possible to calculate the wave function we may also calculate at approximately the same cost the relaxation part of the molecular Hessian.

The remainder of this section is divided into four parts. We first consider the construction of the differentiated gradients f^a , and then go on to describe how direct MCSCF techniques may be used to perform multiplications of trial vectors on the electronic Hessian G without constructing the individual elements of G . Such techniques are essential whenever the dimension of the electronic Hessian precludes its explicit construction. After this we discuss the modification of the conjugate gradient (CG) method²⁸ we have used for solving the MCSCF response equations. This point requires careful consideration as a straightforward implementation of the CG method often leads to slow convergence in the orbital part of the solution vectors ${}^o \lambda^a$ due to large off-diagonal elements in the electronic Hessian. Finally in this section we discuss the relationship between the convergence error in the individual solution vectors λ^a and the error in the relaxation part of the molecular Hessian R^{ab} . In particular we demonstrate that the error in R^{ab} is quadratic in the error of the individual vectors λ^a if these vectors are all expanded in a common set of trial vectors.

1. Construction of the differentiated gradients

The configuration part of the differentiated gradient ${}^c f^a$ [Eq. (51a)] is constructed in an integral-driven scheme^{29,30} according to the equation

$$\langle \text{RWF} | \hat{H}^a | \mu \rangle = \sum_{tu} B_{tu}^\mu {}^c F_{tu}^a + \sum_{tuwx} B_{tuwx}^\mu (g_{tuwx}^{\text{OMO}})^a + C_\mu I^a, \quad (53)$$

where

$$I^a = \sum_i [(h_{ii}^{\text{OMO}})^a + {}^c F_{ii}^a]. \quad (54)$$

The transition density elements in Eq. (53) are obtained by contracting the CI coupling coefficients of Eq. (8) with the CSF expansion coefficients

$$B_{tu}^\mu = \sum_\nu A_{tu}^{\mu\nu} C_\nu, \quad (55a)$$

$$B_{tuwx}^\mu = \sum_\nu A_{tuwx}^{\mu\nu} C_\nu. \quad (55b)$$

Equation (53) is identical to the one used in wave function optimization except that the derivative of the inactive Fock matrix replaces the undifferentiated inactive Fock matrix, and the derivatives of the OMO Hamiltonian integrals replace the undifferentiated integrals. The derivatives of these integrals are calculated according to Eq. (16b), i.e., by adding one-index transformed undifferentiated integrals to the differentiated UMO integrals. The derivative of the inactive Fock matrix is calculated as described in Sec. III B. Note that by contracting $\langle \text{RWF} | \hat{H}^a | \mu \rangle$ [Eq. (53)] with the CSF coefficients of the MCSCF state, we recover the molecular gradient \mathcal{E}^a . This is a very useful and quick test on the calculation.

The calculation of the orbital part of the differentiated gradient ${}^o f^a$ [Eq. (51b)] is trivial once the derivative of the total Fock matrix F^a is available, see Sec. III B.

2. Direct MCSCF techniques

The key step in the solution of the response equation is the multiplication of trial vectors b on the electronic Hessian²⁸:

$$\begin{pmatrix} {}^c \sigma \\ {}^o \sigma \end{pmatrix} = \begin{pmatrix} {}^{cc}G & {}^{co}G \\ {}^{oc}G & {}^{oo}G \end{pmatrix} \begin{pmatrix} {}^c b \\ {}^o b \end{pmatrix}. \quad (56)$$

In the direct MCSCF scheme this multiplication is performed without calculating the elements of the electronic Hessian explicitly. After some simple algebra the working equations are seen to be^{5,20}

$$\sum_\nu {}^{cc}G_{\mu,\nu} {}^c b_\nu = 2(\langle \mu | \hat{H} | B \rangle - {}^c b_\mu \mathcal{E}), \quad (57a)$$

$$\begin{aligned} \sum_\nu {}^{oc}G_{pq,\nu} {}^c b_\nu &= 2(\langle \text{RWF} | [\hat{E}_{pq}, \hat{H}] | B \rangle \\ &+ \langle B | [\hat{E}_{pq}, \hat{H}] | \text{RWF} \rangle), \end{aligned} \quad (57b)$$

$$\sum_{r>s} {}^{oo}G_{\mu,rs} {}^o b_{rs} = 2\langle \mu | \hat{K} | \text{RWF} \rangle, \quad (57c)$$

$$\sum_{r>s} {}^{oo}G_{pq,rs} {}^o b_{rs} = 2\langle \text{RWF} | [\hat{E}_{pq}, \hat{K}] | \text{RWF} \rangle, \quad (57d)$$

where $|B\rangle$ is the state vector

$$|B\rangle = \sum_\mu {}^c b_\mu |\mu\rangle \quad (58)$$

and \hat{K} is the transformed Hamiltonian operator³¹

$$\hat{K} = \sum_{p>q} {}^o b_{pq} [\hat{E}_{pq}, \hat{H}]. \quad (59)$$

Note that \hat{K} is Hermitian and has the same general structure as the original Hamiltonian \hat{H} . The Hamiltonian integrals of \hat{K} are obtained by one-index transforming the \hat{H} integrals, using ${}^o b_{pq}$ (ordered as an antisymmetric matrix) as transformation matrix.

We would like to emphasize the simplicity and usefulness of the above formulas (57). To summarize, a linear transformation on a CSF vector requires the construction of a CSF gradient with $|B\rangle$ as the reference state [Eq. (57a)], and the construction of an orbital gradient with a symmetric

transition density matrix [Eq. (57b)]. Similarly a linear transformation on an orbital vector requires the construction of a CSF gradient [Eq. (57c)] and an orbital gradient [Eq. (57d)] with one-index transformed integrals [Eq. (59)]. The calculation of the above expressions in terms of Fock matrices has been discussed in detail by Jensen and Ågren.³²

3. Implementation of a modified conjugate gradient algorithm

As previously mentioned a straightforward implementation of the generalized conjugate gradient (GCG) method,²⁸ where new trial vectors are generated by dividing the residual vectors by the diagonal Hessian elements, will often lead to slow convergence in the orbital part of the solution vector. A more efficient algorithm is obtained by dividing the trial vectors into two classes: Configuration and orbital trial vectors. A configuration trial vector [${}^c\tilde{\mathbf{b}}, \tilde{\mathbf{0}}$] is obtained by discarding the orbital part of the GCG trial vector [$\tilde{\mathbf{b}}, \tilde{\mathbf{b}}$], while an orbital trial vector [$\tilde{\mathbf{0}}, {}^o\tilde{\mathbf{b}}$] is obtained by discarding the configuration part of the GCG trial vector. This split trial vector procedure adds flexibility to the GCG method as in each iteration we may generate either a configuration or an orbital trial vector depending on which part of the current solution vector has the largest residual. Slow convergence in the orbital part of the solution vector (a typical situation) means that more orbital than configuration trial vectors must be added. When the residual is mainly confined to the orbital space the number of trial vectors in a GCG calculation will be roughly equal to the number of orbital trial vectors in a split GCG calculation. The split GCG procedure therefore becomes more economical as we can save a number of expensive linear transformations of configuration trial vectors. We also note that the linear transformation of a full trial vector is only slightly faster than the separate transformations of a configuration and an orbital trial vector. Furthermore, as we shall see later, the splitting also enables us to determine orbital trial vectors using a better algorithm than the GCG algorithm and at low cost.

Let us assume that we have performed n split GCG iterations. In the course of these iterations we have generated a set of orthonormal trial vectors orthogonal to the MCSCF state vector

$$\{\mathbf{b}_i\} = \left\{ \begin{pmatrix} {}^c\mathbf{b} \\ \mathbf{0} \end{pmatrix}_1, \begin{pmatrix} \mathbf{0} \\ {}^o\mathbf{b} \end{pmatrix}_2, \begin{pmatrix} \mathbf{0} \\ {}^o\mathbf{b} \end{pmatrix}_3, \begin{pmatrix} {}^c\mathbf{b} \\ \mathbf{0} \end{pmatrix}_4, \dots, \begin{pmatrix} \mathbf{0} \\ {}^o\mathbf{b} \end{pmatrix}_n \right\} \quad (60)$$

and carried out the corresponding linear transformations according to Eq. (56):

$$\{\sigma_i\} = \left\{ \begin{pmatrix} {}^c\sigma \\ {}^o\sigma \end{pmatrix}_1, \begin{pmatrix} {}^c\sigma \\ {}^o\sigma \end{pmatrix}_2, \begin{pmatrix} {}^c\sigma \\ {}^o\sigma \end{pmatrix}_3, \begin{pmatrix} {}^c\sigma \\ {}^o\sigma \end{pmatrix}_4, \dots, \begin{pmatrix} {}^c\sigma \\ {}^o\sigma \end{pmatrix}_n \right\}. \quad (61)$$

We may then set up the reduced set of linear equations

$$\begin{pmatrix} {}^c\mathbf{f}^{Ra} \\ {}^o\mathbf{f}^{Ra} \end{pmatrix} + \begin{pmatrix} {}^{cc}\mathbf{G}^R & {}^{co}\mathbf{G}^R \\ {}^{oc}\mathbf{G}^R & {}^{oo}\mathbf{G}^R \end{pmatrix} \begin{pmatrix} {}^c\lambda^{Ra} \\ {}^o\lambda^{Ra} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \quad (62)$$

in terms of the above vectors (60) and (61):

$$\mathbf{f}_i^{Ra} = \tilde{\mathbf{b}}_i \mathbf{f}^a, \quad (63)$$

$$\mathbf{G}_{ij}^R = \tilde{\mathbf{b}}_i \sigma_j. \quad (64)$$

The reduced response equations (62) determine an optimal solution vector [${}^c\tilde{\lambda}^{Ra}, {}^o\tilde{\lambda}^{Ra}$] in the basis of the trial vectors (60). In the split GCG approach outlined above we may now calculate either an improved set of configuration coefficients by applying the GCG algorithm to the [${}^c\tilde{\lambda}^{Ra}, \tilde{\mathbf{0}}$] vector, or an improved set of orbital rotation parameters by applying the GCG algorithm to the [$\tilde{\mathbf{0}}, {}^o\tilde{\lambda}^{Ra}$] vector. The new trial vector (configuration or orbital) is added to the basis (60), and the iterative procedure is continued until the norm of the residual vector is smaller than the requested threshold.

To illustrate the convergence characteristics of the split GCG approach we report test calculations on ONF (8029 CSF and 342 orbital parameters) and NH₃ (490 CSF and 191 orbital parameters). The configuration and orbital residuals obtained in each iteration are listed in Tables I and II (left columns). In these calculations we have generated a configuration trial vector whenever the configuration residual is larger than the orbital residual and vice versa, and each calculation was converged to a tolerance of 10⁻².

In the ONF calculation little is gained by using the split GCG algorithm as configuration and orbital trial vectors are added alternately to the trial basis. In each iteration the residual for the type of basis vector added drops by a factor 0.2–0.5, while the residual for the other basis type is either unchanged or slightly increased. Nine trial vectors of either type are needed for convergence.

The calculation on NH₃ presents a case where it is essential to use the split GCG approach. After each configuration iteration we must typically generate five orbital trial vectors to bring the orbital residual down below the configuration residual. In the GCG approach these five orbital trial vectors

TABLE I. Norm of orbital and configuration residual vector when converging one set of linear equations for ONF to an accuracy of 10⁻².

Iteration number	Split generalized conjugate gradient		Optimal orbital trial vectors	
	Configuration residual	Orbital residual	Configuration residual	Orbital residual
1	1.346	3.852 ^a	1.346	3.852 ^a
2	1.373 ^a	0.936	1.408 ^a	0.124
3	0.674	1.081 ^a	0.693 ^a	0.452
4	0.724 ^a	0.443	0.347	0.625 ^a
5	0.377	0.567 ^a	0.456 ^a	0.028
6	0.449 ^a	0.160	0.257 ^a	0.146
7	0.234	0.244 ^a	0.121	0.197 ^a
8	0.266 ^a	0.082	0.155 ^a	0.010
9	0.130 ^a	0.118	0.075 ^a	0.036
10	0.063	0.139 ^a	0.040	0.043 ^a
11	0.073 ^a	0.058	0.045 ^a	0.004
12	0.045	0.067 ^a	0.013 ^a	0.013
13	0.050 ^a	0.029	0.006	0.014 ^a
14	0.024	0.032 ^a	0.008	0.001
15	0.026 ^a	0.012		
16	0.012	0.014 ^a		
17	0.013 ^a	0.007		
18	0.006	0.010 ^a		
19	0.008	0.006		

^a A trial vector of this type is added.

TABLE II. Norm of orbital and configuration residual vector when converging one set of linear equations for NH₃ to an accuracy of 10⁻².

Iteration number	Split generalized conjugate gradient		Optimal orbital trial vectors	
	Configuration residual	Orbital residual	Configuration residual	Orbital residual
1	0.256	2.964 ^a	0.256	2.964 ^a
2	0.177	1.161 ^a	0.273	0.310 ^a
3	0.236	0.999 ^a	0.294 ^a	0.111
4	0.283	0.869 ^a	0.166	0.194 ^a
5	0.297	0.456 ^a	0.191 ^a	0.058
6	0.296 ^a	0.228	0.109 ^a	0.078
7	0.169	0.293 ^a	0.045	0.087 ^a
8	0.186	0.268 ^a	0.057 ^a	0.013
9	0.192	0.315 ^a	0.029 ^a	0.016
10	0.200	0.308 ^a	0.011	0.019 ^a
11	0.196	0.338 ^a	0.015 ^a	0.005
12	0.192	0.295 ^a	0.006	0.005
13	0.190 ^a	0.158		
14	0.104	0.158 ^a		
15	0.106	0.111 ^a		
16	0.104 ^a	0.077		
17	0.041	0.077 ^a		
18	0.042	0.060 ^a		
19	0.042	0.060 ^a		
20	0.042	0.073 ^a		
21	0.043	0.074 ^a		
22	0.042	0.089 ^a		
23	0.045	0.056 ^a		
24	0.046 ^a	0.032		
25	0.023	0.037 ^a		
26	0.026 ^a	0.022		
27	0.012	0.025 ^a		
28	0.013	0.022 ^a		
29	0.015	0.034 ^a		
30	0.017	0.041 ^a		
31	0.019	0.019 ^a		
32	0.019 ^a	0.010		
33	0.009	0.010 ^a		
34	0.009	0.010 ^a		
35	0.009	0.014 ^a		
36	0.010	0.014 ^a		
37	0.010 ^a	0.007		
38	0.004	0.008		

^a A trial vector of this type is added.

would be replaced by five trial vectors containing both orbital and configuration components. As many as 7 configuration and 30 orbital iterations are needed to arrive at the converged vector.

Although the NH₃ calculation illustrates the importance of splitting the trial vectors into two classes, the large number of orbital trial vectors compared to the number of configuration trial vectors is disturbing as each of the orbital iterations requires the construction of a CSF gradient [Eq. (57c)] with one-index transformed integrals. For large configuration expansions construction of CSF gradients is very expensive. For this reason we suggest the following alternative procedure: Rather than generating each new orbital trial vector ${}^o\mathbf{b}$ by applying the GCG algorithm to the current orbital solution ${}^o\lambda^{Ra}$, we generate ${}^o\mathbf{b}$ by solving the orbital part of the response equations (50) exactly, keeping the current approximate configuration solution ${}^c\lambda^{Ra}$ fixed. The determination of this "optimal"

orbital trial vector does not require the construction of any CSF gradients or any density matrices. To see this we substitute [${}^c\tilde{\lambda}^{Ra}, {}^o\mathbf{b}$] in the response equations (50). The second component of Eq. (50),

$${}^o\mathbf{f}^a + {}^{oc}\mathbf{G}^c\lambda^{Ra} + {}^{oo}\mathbf{G}^o\mathbf{b} = 0 \quad (65)$$

which determines ${}^o\mathbf{b}$ for a given ${}^c\lambda^{Ra}$, then becomes a set of linear equations

$${}^o\mathbf{F}^a + {}^{oo}\mathbf{G}^o\mathbf{b} = 0 \quad (66)$$

with a modified gradient

$${}^o\mathbf{F}^a = {}^o\mathbf{F}^a + {}^{oc}\mathbf{G}^c\lambda^{Ra}. \quad (67)$$

Equation (66) is easily solved as its dimension is equal to the number of orbital rotations. Note that ${}^{oc}\mathbf{G}^c\lambda^{Ra}$ in Eq. (67) is straightforwardly constructed from the linearly transformed vectors (61), which are already available, and thus ${}^o\mathbf{b}$ may be determined without having to go through the list of CI coupling coefficients. Once ${}^o\mathbf{b}$ has been determined from Eq. (66) it is orthonormalized to the basis (60) and appended to the basis.

To illustrate the efficiency of this scheme we have repeated the above calculations on ONF and NH₃ using the optimal orbital trial vectors rather than the GCG orbital vectors. In the ONF case (Table I) only a slight improvement is observed. This is not surprising as the orbital and configuration components converge equally well and very little can therefore be gained by improving the convergence in just the orbital part of the solution vector. The total of nine configuration and nine orbital split GCG iterations is reduced to eight and five, respectively, when optimal orbital trial vectors are used. Note that whenever a new optimal orbital trial vector is generated the orbital residual drops significantly below the configuration residual and occasionally two configuration iterations have to be carried out before the configuration residual is smaller than the orbital residual. In the NH₃ case (Table II) we observe a drastic improvement when the optimal orbital trial vectors are used. The total of 7 configuration and 30 orbital vectors is reduced to 6 and 5, respectively, clearly a major improvement on the original scheme.

In conclusion we note that the split GCG algorithm and the optimal orbital algorithm both require approximately the same number of configuration trial vectors—they differ mainly in the number of orbital trial vectors. The optimal algorithm depends mainly on the relative timings for construction of the CSF gradient Eq. (57c) and for solving the linear equations (66). Note that it will often be advantageous to calculate the orbital Hessian ${}^{oo}\mathbf{G}$ (or even its inverse ${}^{oo}\mathbf{G}^{-1}$) explicitly in optimal orbital trial vector calculations as the same matrix ${}^{oo}\mathbf{G}$ is used in all iterations and for all atomic displacements.

4. Simultaneous solution of the linear equations

To calculate the molecular Hessian we must solve

$$\mathbf{G}\lambda^b = -\mathbf{f}^b \quad (49)$$

for each of the $3N - 6$ ($3N - 5$) independent nuclear distortions. As iterative methods are used, the solution vectors λ^b are only determined to a specified accuracy. However, we are not

interested in the accuracy of the individual solution vectors λ^b themselves. Rather it is the accuracy of the relaxation contribution written as

$$R^{ab} = \tilde{\lambda}^a f^b = \tilde{f}^a \lambda^b \quad (68)$$

which determines the overall accuracy of the calculation. In Appendix B we demonstrate that if we solve all Eqs. (49) simultaneously (expanding the approximate solutions λ^b in the same set of trial vectors), the error in R^{ab} becomes quadratic in the errors of the individual solution vectors λ^b . In addition the approximate relaxation matrix R^{ab} is automatically symmetric. Conversely, when each perturbation is treated separately the matrix R^{ab} becomes nonsymmetric and the errors in the off-diagonal elements are linear in the errors of the solution vectors λ^b .

To illustrate the importance of simultaneous solutions we have carried out some test calculations on the water molecule. Tables III and IV give the absolute errors in R^{ab} obtained by one-by-one solution and simultaneous solution, respectively, (both converged to 10^{-2} in λ^b). The errors in R^{ab} obtained by simultaneous solution is 10^{-5} , which is less than 10^{-2} squared, and the matrix is symmetric to machine accuracy. Conversely, the one-by-one solution gives errors of the order 10^{-5} in diagonal elements only, while the off-diagonal elements are correct to 10^{-3} . Also the matrix is symmetric to 10^{-3} only. Note that all diagonal deviations are positive, reflecting the fact that G is positive definite. [See Eq. (B9).] Additional benefits provided by simultaneous solutions are reduction of overhead and possibly faster convergence to a given accuracy in λ^b . The reason for this faster convergence is simply that the reduced space generated by a given vector f^b is shared by all other vectors as well.

D. Integral evaluation

A new code has been written to calculate first and second derivative Gaussian integrals. The program follows the McMurchie–Davidson scheme^{21,33} by expanding the Cartesian integrals in Hermitian integrals, which are evaluated following the scheme originally devised by Boys.²² For a description of the McMurchie–Davidson scheme and for comparisons with other algorithms see Refs. 33 and 34.

The ABACUS integral code has been vectorized with inner loops over primitive Gaussian functions. To maximize vector lengths we have chosen to treat all primitive functions on the same atom and with the same angular quantum number simultaneously, whether or not these primitive functions contribute to the same final contracted orbital. In the final step of the calculation the primitive Cartesian integrals are transformed to

TABLE III. The deviation from the exact relaxation contribution in a H_2O calculation when the linear set of equations are solved one by one to a tolerance of 10^{-2} .

	1	2	3
1	0.000 031	– 0.000 015	0.000 224
2	– 0.001 185	0.000 034	– 0.000 112
3	0.000 349	– 0.000 797	0.000 055

TABLE IV. The deviation from the exact relaxation contribution in a H_2O calculation when the linear set of equations are solved simultaneously to a tolerance of 10^{-2} .

	1	2	3
1	0.000 026	– 0.000 013	0.000 016
2	– 0.000 013	0.000 014	– 0.000 007
3	0.000 016	– 0.000 007	0.000 026

contracted Cartesian integrals. This approach is well suited for general contraction schemes³⁵ but at the moment we have only implemented the simpler segmented contractions.

An important feature of the McMurchie–Davidson scheme is that it allows us to treat the two overlap distributions in the electron repulsion integrals separately as illustrated below:

$$\langle \text{Herm.} | \text{Herm.} \rangle \rightarrow \langle \text{Herm.} | \text{Cart.} \rangle \rightarrow \langle \text{Cart.} | \text{Cart.} \rangle .$$

Following McMurchie and Davidson²¹ we take advantage of this fact by transforming the potentially large number of primitive Cartesian components from the first “half-calculation” to the less numerous contracted Cartesian elements before the final integrals are calculated in the second half-calculation. As noted by Saunders,³³ for asymmetric overlap distributions it is not unimportant (in terms of efficiency) which distribution is treated first. Especially in derivative calculations where the differentiation of the orbitals often leads to highly asymmetric overlap distributions care must be taken as to which distribution is treated first. ABACUS has the ability to treat either overlap distribution first, yielding considerable savings in computer time.

It is often advantageous to calculate the integral derivatives directly, without recourse to intermediate undifferentiated integrals.³⁶ This is achieved by combining the appropriate Hermitian-to-Cartesian expansion coefficients before these coefficients are contracted with the Hermitian integrals. However, we shall see below that in many cases the derivative integrals may be calculated both simpler and more efficiently by taking advantage of the derivative properties of the Hermitian functions used to expand the overlap distributions. To show this we restrict ourselves to the one-dimensional case and consider the overlap distribution of two Cartesian orbitals fixed on nuclei with coordinates A and B ,

$$\Omega_{AB}^i = X_A^i X_B^i \exp(-\alpha X_A^2) \exp(-\beta X_B^2), \quad (69)$$

where we have used the notation $X_A = X - A$. Introducing the coordinates:

$$P = (\alpha A + \beta B) / (\alpha + \beta),$$

$$Q = A - B \quad (70)$$

this overlap distribution may be written as a finite linear combination of Hermitian functions

$$\Lambda_i(P) = \left(\frac{\partial}{\partial P} \right)^i \exp[-(\alpha + \beta) X_p^2] \quad (71)$$

as described by McMurchie and Davidson.²¹ This gives us

$$\Omega_{AB}^{ij} = K(Q) \sum_{t=0}^{i+j} E_{ij}^t(Q) \Lambda_t(P), \quad (72)$$

where we have introduced the factor

$$K(Q) = \exp[-\alpha\beta Q^2/(\alpha + \beta)] \quad (73)$$

and the expansion coefficients $E_{ij}^t(Q)$ which may be evaluated recursively. For later use we note that for a fixed set of exponents α and β the coefficients $E_{ij}^t(Q)$ are functions of Q only.

We now determine the derivatives of the overlap distributions and note that these may be obtained by differentiating Eq. (69) directly or alternatively by differentiating the Hermitian expansion (72). By direct differentiation of Eq. (69) with respect to A we find

$$\left(\frac{\partial}{\partial A}\right) \Omega_{AB}^{ij} = 2\alpha \Omega_{AB}^{i+j+1,j} - i\Omega_{AB}^{i-1,j}. \quad (74)$$

If we substitute the Hermitian expansion (72) in this expression and collect terms to the same order in t , we obtain the final result

$$\left(\frac{\partial}{\partial A}\right) \Omega_{AB}^{ij} = K(Q) \sum_{t=0}^{i+j+1} F_{ij}^t(Q) \Lambda_t(P), \quad (75)$$

where

$$F_{ij}^t(Q) = 2\alpha E_{i+1,j}^t(Q) - iE_{i-1,j}^t(Q). \quad (76)$$

The expansion (75) is similar to Eq. (72) except that the summation of Hermitian functions now contains an extra term $t = i + j + 1$. This is computationally significant as this summation is the time-consuming step in the McMurchie–Davidson algorithm.

If we now turn to the sum of the derivatives with respect to A and B , it would seem that we have to calculate the individual derivatives first using Eq. (75) and then add the results. However, this is not necessary. This can be seen by noting that

$$\begin{aligned} \left(\frac{\partial}{\partial A} + \frac{\partial}{\partial B}\right) K(Q) &= 0, \\ \left(\frac{\partial}{\partial A} + \frac{\partial}{\partial B}\right) E_{ij}^t(Q) &= 0 \end{aligned} \quad (77)$$

because $K(Q)$ and $E_{ij}^t(Q)$ are functions of Q only, and further that

$$\left(\frac{\partial}{\partial A} + \frac{\partial}{\partial B}\right) \Lambda_t(P) = \Lambda_{t+1}(P) \quad (78)$$

from the definition Eq. (71). Differentiating expansion (72) instead of the expression (69), we find

$$\left(\frac{\partial}{\partial A} + \frac{\partial}{\partial B}\right)^n \Omega_{AB}^{ij} = K(Q) \sum_{t=0}^{i+j} E_{ij}^t(Q) \Lambda_{t+n}(P) \quad (79)$$

which is identical to the original expansion (72) except that the index t has been incremented by n in the Hermitian function. This implies that the sum of the derivatives may be calculated almost as fast as the undifferentiated integrals themselves. This is very useful when both orbitals are centered on the same nucleus (a common situation in small and medium sized molecules) as we are then only interested in the sum of the derivatives. This feature has been implemented in the ABACUS code. Even in cases where the individual

derivatives are needed (i.e., two-center overlap distributions) it might be faster to calculate the sum of those derivatives and only one of the individual derivatives, especially in highly vectorized codes where overhead has been minimized.

E. Program overview

Before closing this section we will discuss some general features of the ABACUS program. For this purpose it is useful to describe the main points of the program structure:

(A) Calculate contributions from the nuclear–nuclear part of the Hamiltonian.

(B) Calculate the static contributions

$$E^{ua}(1) - \text{tr } S^a F,$$

$$E^{uab}(1) - \text{tr } S^{ab} F,$$

where $E(1)$ denotes contributions from the one-electron part of the Hamiltonian. Write first derivatives of the one-electron Hamiltonian and overlap integrals on disk.

(C) Transform the two-electron density matrix to AO basis. Calculate the two-electron static contributions $E^{ua}(2)$ and $E^{uab}(2)$.

(D) To calculate Y^a and f^a loop over atoms a .

(1) Recalculate first derivatives of two-electron integrals with respect to the coordinates of atom a and write these on disk.

(2) Construct ${}^c F^{ua}$ and ${}^v F^{ua}$.

(3) Transform the differentiated two-electron integrals to MO basis (1 general and 3 active indices) and construct Q^{ua} .

(4) Construct reorthonormalization matrices ${}^c F^{sa}$ and ${}^v F^{sa}$.

(5) One-index transform the undifferentiated two-electron MO integrals. Add these to the differentiated two-electron integrals and construct Q^{sa} .

(6) Construct matrices Y^a .

(7) Construct gradients f^a .

(E) Calculate the remaining static reorthonormalization contribution $-\text{tr}(S^a Y^b + S^b Y^a)$.

(F) Solve the response equations and calculate the relaxation contributions $\hat{f}^a \lambda^b$.

(G) Use translational and rotational symmetries to determine the remaining elements in the molecular Hessian.

As can be seen from the above scheme all first derivatives of the two-electron integrals are calculated twice. First we calculate them in step (C) simultaneously with the second derivatives, taking full advantage of translational symmetry.³⁷ (Rotational symmetry has not been implemented in this part of the program.) The calculated integrals are immediately combined with AO density elements and added to the appropriate components of E^{ua} and E^{uab} . No integrals are written on disk in this step. The first derivatives are later needed in step (D) to calculate the Fock matrix contributions ${}^c F^{ua}$, ${}^v F^{ua}$, and Q^{ua} and also the differentiated two-electron MO integrals. These derivatives are therefore recalculated in this step, one atom at a time, written on disk and processed completely before the next atom is considered. In this way the disk requirements are reduced.

The response equations are solved in step (F). Note that at this point all necessary gradients f^a have been calculated. This means that we may treat simultaneously as many gradients as the computer memory allows in this part of the calculation. As described above, this greatly increases the efficiency of this time-consuming step.

In steps (D), (E), and (F) we make full use of translational and rotational symmetries to reduce the computational costs. First the program determines the maximum number of independent external directions (5 or 6), then it picks out 5(6) nuclear coordinates that describe these external directions. Finally in step (G) the as yet undetermined elements of the Hessian matrix are calculated using the procedure of Page *et al.*³⁸

IV. SAMPLE CALCULATIONS

To demonstrate the potential of the presented formalism as implemented in the ABACUS program we report sample calculations on the water and nitrosyl fluoride (ONF) molecules. All calculations were carried out at the experimental equilibrium geometries. The wave functions were determined using the SIRIUS MCSCF program³² together with the MOLECULE integral code³⁹ and Siegbahn's CASSCF GUGA code.^{30(a)} In calculating the molecular Hessians the response equations were converged to 10^{-2} or better in the solution vectors which implies that the molecular Hessians are accurate to 10^{-4} a.u.

A. Application to the water molecule

We used Tatewaki's MAXI-3 basis set⁴⁰ on oxygen augmented with a diffuse p shell⁴¹ with exponent 0.059 and two d -type polarization functions⁴² with exponents 2.704 and 0.535. On hydrogen Huzinaga's $4s$ basis⁴³ was contracted to $3s$ ⁴⁴ and augmented with one diffuse s function and two p shells (basis set D of Ref. 45). The entire basis set—(O/10,7,2/4,5,2) and (H/5,3/4,2)—thus consisted of 71 primitive functions contracted to 51. The active space was chosen to give a balanced treatment of the correlation effects as described by Roos⁴⁶. Each of the four valence orbitals (excluding $1s$) was correlated with a weakly occupied orbital of same symmetry, and an additional weakly occupied orbital of A_2 symmetry was included to give angular correlation to the $1b_1$ lone-pair orbital. The distribution of 8 valence electrons among 9 active orbitals gives rise to 2688 singlet CSF's belonging to the A' representation of the C_s group (5292 CSF's without symmetry).

The calculated total energy is -76.191725 a.u. and the calculated forces and force constants are presented in Table V. Assuming a perfectly harmonic energy surface the predicted MCSCF equilibrium geometry is 0.9637 Å and 104.50° , which may be compared with the experimental values^{47(a)} 0.9578 Å and 104.48° . The MCSCF force constants are in close agreement with the experimental data with a maximum deviation from the experiment of 5%. The average deviation from experiment is less than 3%.

B. Application to the nitrosyl fluoride molecule

In order to investigate the program performance for a fairly large configuration space we carried out a calculation on ONF, using Dunning's double zeta ($9s5p/4s2p$) basis⁴⁴ aug-

TABLE V. Forces and force constants for H_2O . Units are $aJ \text{ \AA}^{-n}$. Theoretical values are calculated at the experimental geometry using the coordinates given in Ref. 47(b). The calculated force constants are corrected for nonequilibrium geometry effects as described by Pulay (Ref. 1). Experimental force constants are from Ref. 47(a).

	MCSCF	Experimental
φ_r	0.0544	0.0
φ_α	0.0034	0.0
F_r	8.555	8.454
F_α	0.741	0.761
$F_{r\alpha}$	-0.099	-0.101
F_{rr}	0.259	0.248

mented with two d -type polarization functions with exponents⁴² (2.704, 0.535), (1.986, 0.412), and (3.559, 0.682) on oxygen, nitrogen, and fluorine, respectively. The basis thus contained a total of 108 primitive functions contracted to 66. The configuration space chosen consisted of a full-valence active set. Distributing the 18 valence electrons among the 12 active orbitals resulted in 8029 configurations of A' symmetry in the C_s point group (15 730 without symmetry).

The total MCSCF energy obtained for ONF is -228.849630 a.u. and the forces and force constants are summarized in Table VI. The overall agreement with experiment is not as good as one might expect for a calculation of this size, the harmonic force constants being off by 6%–28% from the experimental values.⁴⁸ The theoretical equilibrium geometry predicted assuming a perfectly harmonic energy surface is 1.1536 Å (ON), 1.5165 Å (NF), and 110.04° . The corresponding experimental values are 1.1316 Å, 1.5166 Å, and 109.92° and the predicted ON distance is thus seen to be 0.02 Å longer than the experimental value. The above discrepancies may partly be a basis set effect as the basis set used is known to perform poorly for polar molecules.⁴⁹ An improved agreement with experimental values may also be obtained with a better choice of active space where each strongly occupied orbital has a weakly occupied counterpart as in the water calculation. In our ONF

TABLE VI. Forces and force constants for ONF. Units are $aJ \text{ \AA}^{-n}$. The calculation is performed at experimental geometry (Ref. 48) and force constants are corrected for nonequilibrium geometry as described by Pulay (Ref. 1). Experimental force constants are from Ref. 48.

	MCSCF	Experimental
φ_{ON}	0.3952	0.0
φ_{NF}	0.0365	0.0
φ_α	0.0119	0.0
F_{ON}	17.934	15.912
F_{NF}	2.441	2.133
F_α	1.958	1.841
$F_{ON/NF}$	1.627	1.902
$F_{ON/\alpha}$	0.414	0.323
$F_{NF/\alpha}$	0.283	0.236

calculation the active space is too crowded and the weakly occupied orbitals are used to correlate lone pairs at the expense of left-right correlation of bond orbitals.

lucular systems involving four-center integrals the proportion of the time spent on the calculation of derivative integrals will increase.

C. Timings

The timings on a VAX 11/780 for the H₂O and ONF MCSCF calculations are given in Table VII. The time-consuming part of the calculations is the solution of the linear equations, which is seen to account for 60%–80% of the total computer time used in the calculation of the Hessian. However, as much as 40% of the time spent on these equations was used to obtain optimal trial vectors according to Eq. (66). If we had constructed the orbital part of the Hessian explicitly the time spent on obtaining the orbital trial vectors would be drastically reduced. In fact, the total time for the molecular Hessian may be reduced by almost 25%.

It is interesting to compare the time spent on the calculation of the differentiated integrals with the time it takes to calculate the undifferentiated integrals. Using the MOLECULE integral code,³⁹ 66 min were spent on the calculation of the two-electron integrals for ONF (excluding the construction of the \mathcal{P} -supermatrix integrals). This may be compared to the 298 min used to calculate both the first and the second derivative integrals using the ABACUS integral code. Although this comparison is not strictly applicable as different codes are used in the two calculations it is nevertheless indicative of the efficiency of the McMurchie–Davidson scheme.²¹ For larger mo-

V. DISCUSSION

We have shown how MCSCF molecular gradients and Hessians may be calculated for state of the art MCSCF wave functions using a response function approach. By transferring the geometry dependence of the basis set to the Hamiltonian integrals the evaluation of the matrix elements appearing in the molecular Hessian is simplified. Reorthonormalization effects are, for example, straightforwardly accounted for by means of one-index transformations on the molecular integrals.

To evaluate the relaxation contribution to the molecular Hessian $3N - 6$ sets of linear equations must be solved. These equations have the same structure as the Newton–Raphson equations used for second-order MCSCF wave function optimization. The same computer code can therefore be used both for wave function optimization and for calculation of relaxation contributions to second-order properties. We have used a direct MCSCF computer code,³² which makes calculations with highly correlated MCSCF wave functions practical. Another important feature of our relaxation code is the simultaneous solution of wave function responses for all independent nuclear distortions. In this way the errors in the relaxation part of the Hessian become quadratic in the errors of the solutions to the individual wave function responses, reducing the computa-

TABLE VII. Timings for the MCSCF sample calculations on a VAX 11/780.

	H ₂ O	ONF
Number of basis functions primitive/contracted	71/51	108/66
Number of configurations	2688	8029
MCSCF CALCULATIONS		
Integrals	13 min	66 min
MCSCF/No. of iterations	370 min/10	1100 min/8
Sum MCSCF:	383 min	1166 min
MOLECULAR HESSIAN CALCULATIONS		
A. Static Hessian		
Integrals (1 and 2 el.)	2 + 64 min	4 + 298 min
Density transformation	7 min	23 min
Sum static part:	73 min	325 min
B. Relaxation Hessian (three directions)		
Integrals	24 min	121 min
Integral transformations	36 min	209 min ^a
Construction of $3 \mathbf{F}^a$	12 min	59 min
Linear equations	543 min	1217 min
Sum relaxation part:	615 min	1606 min
Sum molecular Hessian:	688 min	1931 min

^a 140 min of the 209 min used in AO sort because of inefficient sort in the program when the ONF calculation was done.

tional effort needed for the calculation of the relaxation contributions to a given accuracy.

To further reduce the computational burden in molecular Hessian calculations we have made extensive use of the Fock matrix formulation of Siegbahn *et al.*²³ This reduces the time spent in integral transformations—in fact only differentiated molecular orbital integrals involving one general and three active orbital indices need to be calculated. In the limit of single-configuration SCF calculations no integrals at all (differentiated or undifferentiated) are transformed to the molecular orbital basis. This corresponds to the SCF formulation of Bacskay,⁵⁰ where the coupled Hartree–Fock equations are solved completely in the AO basis.

The use of Fock matrices also makes the program more flexible with respect to the size of the molecular basis sets. To further reduce the constraints on basis sets the full set of differentiated two-electron integrals is never stored on disk simultaneously. Rather the differentiated integrals are calculated with respect to three perturbations (one atom) at a time, making the external storage requirements not substantially higher than for the wave function determination itself.

The time required for evaluating the MCSCF molecular Hessian is in general of the same order of magnitude as the time required for carrying out a second-order optimization of the wave function. We may therefore expect that in the near future MCSCF molecular Hessians will be routinely used to study molecular potential surfaces and serve as an important tool for improving our understanding of the electronic structure of molecular systems.

ACKNOWLEDGMENTS

We wish to thank Per Siegbahn for the use of his GUGA program and Peter Taylor for the use of his two-electron density matrix transformation program. Peter Pulay has kindly supplied us with his code for constructing the B matrix used for transformation from Cartesian to internal coordinates. This work was supported by the Danish Natural Science Research Council.

APPENDIX A: STATIC ENERGY DERIVATIVES IN TERMS OF FOCK MATRICES

The static energy derivatives E^a and E^{ab} are in Sec. III A expressed in terms of Fock matrices. We now show how these expressions may be derived. From Eq. (16) we see that

$$\begin{aligned} (h^{\text{OMO}})^a &= h^a - \frac{1}{2} \{S^a, h\}, \\ (g^{\text{OMO}})^a &= g^a - \frac{1}{2} \{S^a, g\}, \end{aligned} \quad (\text{A1})$$

where we have used the notation of Eq. (17) for one-index transformations. The molecular gradient E^a is obtained by contracting the Hamiltonian integrals (A1) with the one- and two-electron density matrices. This is a straightforward task once we note that for any matrix A,

$$\sum_{pq} D_{pq} \{A, h\}_{pq} + \sum_{pqrs} P_{pqrs} \{A, g\}_{pqrs} = 2 \text{tr} AF, \quad (\text{A2})$$

where F is the Fock matrix (30). This means that contracting the density matrices with one-index transformed integrals is equivalent to contracting the transformation matrix with the

Fock matrix. Using Eqs. (A1) and (A2) we immediately see that

$$E^a = E^{ua} - \text{tr} S^a F \quad (\text{A3})$$

which is the expression for the molecular gradient.

To obtain the second derivatives E^{ab} we first express the second derivatives of the OMO integrals in terms of one-index transformations, then contract these derivatives with the density matrices and finally rewrite the expressions in terms of Fock matrices using Eq. (A2). The second (and higher) derivatives of the OMO integrals may be determined using the techniques given by Simons *et al.*¹⁶ We find

$$\begin{aligned} (h^{\text{OMO}})^{ab} &= h^{ab} - \frac{1}{2} \{S^{ab}, h\} \\ &\quad - \frac{1}{2} \{S^a, h^b\} - \frac{1}{4} \{S^a, \{-\frac{1}{2} S^b, h\}\} + \frac{1}{4} \{S^a S^b, h\} \\ &\quad - \frac{1}{2} \{S^b, h^a\} - \frac{1}{4} \{S^b, \{-\frac{1}{2} S^a, h\}\} + \frac{1}{4} \{S^b S^a, h\} \end{aligned} \quad (\text{A4})$$

and similarly in the two-electron case. Contracting these expressions with the density matrices we obtain

$$\begin{aligned} E^{ab} &= E^{uab} - \text{tr} S^{ab} F - \text{tr} S^a F^{ub} - \frac{1}{2} \text{tr} S^a F^{sb} + \frac{1}{2} \text{tr} S^a S^b F \\ &\quad - \text{tr} S^b F^{ua} - \frac{1}{2} \text{tr} S^b F^{sa} + \frac{1}{2} \text{tr} S^b S^a F \\ &= E^{uab} - \text{tr} S^{ab} F - \text{tr} (S^a Y^b + S^b Y^a), \end{aligned} \quad (\text{A5})$$

where we have used the matrix

$$Y^a = F^{ua} + \frac{1}{2} F^{sa} - \frac{1}{2} S^a F, \quad (\text{A6})$$

introduced in Sec. III A. Higher order static contributions such as E^{abc} may of course be determined using the same technique.

APPENDIX B: ERROR ANALYSIS FOR SIMULTANEOUS SOLUTION OF LINEAR EQUATIONS

An important reason for expanding all $3N - 6$ ($3N - 5$) independent solution vectors of Eq. (49) in the *same* set of trial vectors is that the errors in the relaxation contribution R^{ab} to the molecular Hessian then become quadratic in the errors in the individual solution vectors. A proof is given in this Appendix.

The response equations (49) are solved by projecting them onto the space spanned by the orthonormal trial vectors $B = \{\mathbf{b}_i | i = 1, n\}$:

$$G^R \lambda^{Ra} = -\mathbf{f}^{Ra}, \quad (\text{B1})$$

where

$$G^R = QGQ, \quad (\text{B2})$$

$$\mathbf{f}^{Ra} = Q\mathbf{f}^a, \quad (\text{B3})$$

and Q is the projector onto B,

$$Q = \sum_{i=1}^n \mathbf{b}_i \tilde{\mathbf{b}}_i. \quad (\text{B4})$$

As a measure of the error in λ^{Ra} we use the residual

$$\delta^{Ra} = G\lambda^{Ra} + \mathbf{f}^a \quad (\text{B5})$$

which is orthogonal to B since

$$Q\delta^{Ra} = Q(G\lambda^{Ra} + \mathbf{f}^a) = G^R \lambda^{Ra} + \mathbf{f}^{Ra} = 0. \quad (\text{B6})$$

This means that δ^{Ra} is orthogonal to all approximate solution

vectors λ^{Rb} as $\lambda^{Rb} \in B$ for all b .

To obtain an expression for the exact solution λ^a in terms of the approximate solution λ^{Ra} and the residual δ^{Ra} , we multiply Eq. (B5) by G^{-1} from the left and use Eq. (49):

$$\lambda^a = \lambda^{Ra} - G^{-1}\delta^{Ra}. \quad (\text{B7})$$

We may now determine the error Δ^{ab} in the relaxation part of the molecular Hessian [see Eq. (47)] by using Eq. (B7) twice:

$$\begin{aligned} \Delta^{ab} &= R^{Rab} - R^{ab} = \tilde{f}^a(\lambda^{Rb} - \lambda^b) \\ &= \tilde{f}^a G^{-1} \delta^{Rb} = -\tilde{\lambda}^a \delta^{Rb} \\ &= -\tilde{\lambda}^{Ra} \delta^{Rb} + \tilde{\delta}^{Ra} G^{-1} \delta^{Rb}. \quad (\text{B8}) \end{aligned}$$

Since the approximate solution vectors λ^{Ra} and the residual vectors δ^{Rb} are orthogonal, we obtain the final result

$$\Delta^{ab} = \tilde{\delta}^{Ra} G^{-1} \delta^{Rb}. \quad (\text{B9})$$

This shows that Δ^{ab} is quadratic in the residuals, for example if all residuals are brought below 10^{-2} Δ^{ab} will be of order 10^{-4} and we have quadratic accuracy in the approximate R^{Rab} compared to the accuracy of the individual response vectors λ^{Ra} . Furthermore, Δ^{ab} in Eq. (B9) is symmetric in a and b and we gain the additional benefit that the approximate solution has the correct symmetric structure.

If we had used different trial vectors for λ^{Ra} and λ^{Rb} then $\tilde{\lambda}^{Ra} \delta^{Rb}$ in Eq. (B8) would be different from zero; the error Δ^{ab} would be linear in the residuals and the approximate solution would only be symmetric to the same order.

¹P. Pulay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III, (Plenum, New York, 1977), Vol. 4, and references therein.

²J. Gerratt and I. M. Mills, *J. Chem. Phys.* **49**, 1719 (1968).

³K. Thomsen and P. Swanström, *Mol. Phys.* **26**, 735 (1973).

⁴J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem. Symp.* **13**, 225 (1979).

⁵J. Olsen, D. L. Yeager, and P. Jørgensen, *Adv. Chem. Phys.* **54**, 1 (1983).

⁶B. O. Roos, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson (Reidel, Dordrecht, 1983).

⁷P. E. M. Siegbahn, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson (Reidel, Dordrecht, 1983).

⁸P. Pulay, *Mol. Phys.* **17**, 197 (1969).

⁹P. Pulay, *J. Chem. Phys.* **78**, 5043 (1983).

¹⁰Y. Yamaguchi, Y. Osamura, G. Fitzgerald, and H. F. Schaefer III, *J. Chem. Phys.* **78**, 1607 (1983).

¹¹P. Jørgensen and J. Linderberg, *Int. J. Quantum Chem.* **4**, 587 (1970).

¹²B. Levy, *Int. J. Quantum Chem.* **4**, 297 (1970).

¹³E. Dalgaard and P. Jørgensen, *J. Chem. Phys.* **69**, 3833 (1978).

¹⁴P. Jørgensen and J. Simons, *J. Chem. Phys.* **79**, 334 (1983).

¹⁵T. U. Helgaker and J. Almlöf, *Int. J. Quantum Chem.* **26**, 275 (1984).

¹⁶J. Simons, P. Jørgensen, and T. U. Helgaker, *Chem. Phys.* **86**, 413 (1984).

¹⁷R. N. Camp, H. F. King, J. W. McIver, Jr., and D. Mullally, *J. Chem. Phys.* **79**, 1088 (1983).

¹⁸M. R. Hoffmann, D. J. Fox, J. F. Gaw, Y. Osamura, Y. Yamaguchi, R. S. Grev, G. Fitzgerald, H. F. Schaefer III, P. J. Knowles, and N. C. Handy, *J. Chem. Phys.* **80**, 2660 (1984).

¹⁹M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfeld III, *J. Chem. Phys.* **81**, 434 (1984).

²⁰H. J. Aa. Jensen and P. Jørgensen, *J. Chem. Phys.* **80**, 1204 (1984).

²¹L. E. McMurchie and E. R. Davidson, *J. Comput. Phys.* **26**, 218 (1978).

²²S. F. Boys, *Proc. R. Soc. London Ser. A* **200**, 542 (1950).

²³P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, *J. Chem. Phys.* **74**, 2384 (1981).

²⁴P. Jørgensen and J. Simons, *Second Quantization-Based Methods in Quantum Chemistry* (Academic, New York, 1981).

²⁵C. C. J. Roothaan and P. S. Bagus, *Methods Comput. Phys.* **2**, 47 (1963).

²⁶B. H. Lengsfeld III and B. Liu, *J. Chem. Phys.* **75**, 478 (1981); B. H. Lengsfeld III, *ibid.* **77**, 4073 (1982).

²⁷D. L. Yeager and P. Jørgensen, *J. Chem. Phys.* **71**, 755 (1979); E. Dalgaard, *Chem. Phys. Lett.* **65**, 559 (1979).

²⁸M. Hestenes, *Conjugate Direction Methods in Optimization* (Springer, Berlin, 1980).

²⁹B. Roos, *Chem. Phys. Lett.* **15**, 153 (1972).

³⁰(a) P. E. M. Siegbahn, *Chem. Phys. Lett.* **109**, 417 (1984); (b) P. J. Knowles and N. C. Handy, *ibid.* **111**, 315 (1984).

³¹P. Jørgensen, J. Olsen, and D. L. Yeager, *J. Chem. Phys.* **75**, 5802 (1981).

³²H. J. Aa. Jensen and H. Ågren, *Chem. Phys. Lett.* **110**, 140 (1984).

³³V. R. Saunders, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson (Reidel, Dordrecht, 1983).

³⁴D. Hegarty and G. van der Velde, *Int. J. Quantum Chem.* **23**, 1135 (1983).

³⁵R. C. Raffanetti, *J. Chem. Phys.* **58**, 4452 (1973).

³⁶P. Saxe, Y. Yamaguchi, and H. F. Schaefer III, *J. Chem. Phys.* **77**, 5647 (1982).

³⁷A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield, and M. Conrad, *Chem. Phys. Lett.* **45**, 595 (1977).

³⁸M. Page, P. Saxe, G. F. Adams, and B. H. Lengsfeld III, *Chem. Phys. Lett.* **104**, 587 (1984).

³⁹J. Almlöf, USIP Report 74-29, University of Stockholm (1974).

⁴⁰H. Tatewaki, *J. Comput. Chem.* **6**, 237 (1985).

⁴¹Th. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 3.

⁴²*Gaussian Basis Sets for Molecular Calculations*, edited by S. Huzinaga (Elsevier, Amsterdam, 1984).

⁴³S. Huzinaga, *J. Chem. Phys.* **42**, 1293 (1965).

⁴⁴Th. H. Dunning, Jr., *J. Chem. Phys.* **53**, 2823 (1970).

⁴⁵G. D. Purvis III and R. J. Bartlett, *Phys. Rev. A* **23**, 1594 (1981).

⁴⁶B. O. Roos, *Int. J. Quantum Chem. Symp.* **14**, 175 (1980).

⁴⁷(a) A. R. Hoy, I. M. Mills, and G. Strey, *Mol. Phys.* **24**, 1265 (1972); A. R. Hoy and P. R. Bunker, *J. Mol. Spectrosc.* **74**, 1 (1979); (b) R. D. Amos, *Chem. Phys. Lett.* **108**, 185 (1984).

⁴⁸C. Degli Esposti, G. Cazzoli, and P. G. Favero, *J. Mol. Spectrosc.* **109**, 229 (1985), and references therein.

⁴⁹J. Almlöf and K. Fægri, Jr., *J. Chem. Phys.* **79**, 2284 (1983).

⁵⁰G. B. Bacskay, *Chem. Phys.* **61**, 385 (1981).