

OPTIMIZATION AND CHARACTERIZATION OF A MULTICONFIGURATIONAL SELF-CONSISTENT FIELD (MCSCF) STATE

JEPPE OLSEN AND DANNY L. YEAGER

*Chemistry Department
Texas A & M University
College Station, Texas*

AND

POUL JØRGENSEN

*Chemistry Department
Aarhus University
Aarhus, Denmark*

CONTENTS

I.	Introduction	2
II.	MCSCF Optimization Using Exponential Unitary Operators	6
	A. Unitary Transformations of the Multiconfigurational Hartree–Fock State	6
	B. Variations in the Total Energy	8
III.	The Newton–Raphson Iterative Function	10
	A. The Linear Newton–Raphson Iterative Function	11
	B. The Nonlinear Newton–Raphson Iterative Function	14
	C. Implementation of the Nonlinear Newton–Raphson Approach	21
	D. Matrix Elements in a Nonlinear Newton–Raphson Approach	23
IV.	Characterization of an MCSCF State	25
	A. n th State of the CI Using the MCSCF Configuration State Functions (CSFs)	27
	B. Eigenvalues of the Hessian	28
	C. Excitation Energies in the Multiconfigurational Time-Dependent Hartree–Fock Approximation	30
	D. Stability Condition for a Multiconfigurational Hartree–Fock State	32

V.	Calculations with the Newton–Raphson Approach	34
A.	Step Size and Sign Control in the Newton–Raphson Approach	35
B.	Small Hessian Eigenvalues	40
C.	Step Size and Sign Controlled Nonlinear Newton–Raphson Calculations	41
D.	Initial Guess of Orbitals and States	62
E.	Conditions for Improved Iterative Methods	65
VI.	Generalized and Fixed Hessian Approaches	66
A.	The Generalized Newton–Raphson Perturbative Approach	66
B.	The Linear Fixed Hessian Approach	67
C.	Optimal Use of Fixed Hessian Approaches	69
D.	Numerical Results	72
VII.	Update Methods	78
A.	Broyden Unsymmetric Rank-1 Update and Its Symmetrization	79
B.	The Broyden Family (DFP and BFGS Updates)	82
C.	A Variational Approach to Update Procedures	86
D.	Inverse Updates	87
E.	Convergence Characteristics of Hessian Update Methods	88
F.	Numerical Results	90
G.	Other Applications of Update Methods	100
VIII.	Cubic Contributions in MCSCF Optimization	101
A.	Theory	101
B.	Numerical Results	107
C.	Summary and Conclusions	112
IX.	Effective Implementation and Combination of MCSCF Procedures	114
A.	Simplifications of Transformations	116
B.	Effective Treatment of Large CI Expansions	121
C.	Combination and Overview of Methods	126
D.	Higher (Infinite) Order Procedures	134
X.	Summary and Conclusions	141
Appendix A.	Redundant Variables	143
Appendix B.	The First Partial Derivative of the Total Energy	145
Appendix C.	Formulas for $(\mathbf{P}^T \mathbf{F}(0))$ and $(\mathbf{P}^T \mathbf{G}(0) \mathbf{P})$	151
Appendix D.	Eigenvalues of the Full Hessian and the Reduced Hessian Matrix	156
Appendix E.	Construction of Cubic Contributions for the Iterative Cubic and Perturbative (Chebyshev) Cubic Procedures	158
Appendix F.	Two-Electron Integral Transformations in MCSCF	164
References		173

I. INTRODUCTION

In multiconfigurational self-consistent field (MCSCF) calculations a simultaneous optimization is performed for the orbitals and configuration expansion coefficients. Until recently, MCSCF optimization procedures considered the orbital and state optimization problems separately.^{1, 2} The state optimization problem was treated by performing a configuration interaction (CI) calculation within the MCSCF configuration space, whereas the

orbital optimization was carried out using multiconfigurational extensions of the Hartree–Fock (HF) iterative procedure.^{1,2} The Hartree–Fock iterative procedure^{3,4} has been used very successfully to determine an optimal set of orbitals for a single configuration state. The success of the HF iterative scheme largely relies, however, on the physical justification (an independent set of particles moving in an average potential) and interpretation (Koopmans's theorem⁵) that the orbitals can be given in the HF approximation. The physical justification and interpretation serve in a sense as the “driving force” in getting the HF iterative scheme to converge. When several configurations are used to describe a molecular system, orbitals cannot be given such a direct physical justification and interpretation and MCSCF iterative procedures that represent extensions of the HF iterative procedure often lose their ability to converge reliably.

The MCSCF orbital optimization problem is a standard optimization problem of numerical analysis and has to be attacked with the efficient optimization procedures of numerical analysis to converge reliably.^{6–9} It is the purpose of the present review to discuss how the MCSCF optimization problem may be parameterized to a form for which the standard optimization procedures of numerical analysis can be applied and to show how convergence problems essentially disappear when these optimization procedures are correctly used.

In the following discussion (Section II) we will describe how the MCSCF optimization problem can be formulated in terms of unitary operators that are exponentials of nonredundant operators for orbital^{10–12} and for state optimization.^{11,12} The optimized MCSCF state may then be described by the effect of these unitary operators acting upon an unoptimized state. By applying the variational principle, different optimization procedures may be easily derived.

A widely used optimization procedure in numerical analysis is the Newton–Raphson technique.^{6–9} Newton–Raphson^{10–37} and other complete second-order^{2,19,28–33} and approximate second-order^{2,11,13,15,35,37} techniques are now used fairly extensively for SCF and MCSCF optimization. In Section III we demonstrate how the Newton–Raphson technique may be implemented into MCSCF. The Newton–Raphson iterative function is derived from a quadratic expansion of the energy function. When the Newton–Raphson approach is applied in numerical analysis, a linear transformation among the variables is carried out in between each step of the iterative procedure. In Section III, we discuss the Newton–Raphson technique and describe how it is most straightforwardly applied in MCSCF if a nonlinear transformation of the variables is carried out between each step of the iterative procedure.^{10–17}

Before proceeding further with an analysis of various optimization procedures, we briefly discuss characteristics of an MCSCF state.^{13,15–17} In a CI calculation, the n th root in energy of the CI secular problem is an upper bound to the n th state. The upper bound property is commonly used as a criterion for determining whether a CI state is a valid representation of the desired state. The upper bound property can be established because the configuration expansion coefficients may be used as *linear* variation parameters. In addition to configuration expansion coefficients, an MCSCF calculation also optimizes nonlinear orbital parameters. An alternative to the upper bound criterion therefore has to be established to decide whether an MCSCF state is a proper representation of the state of consideration. In this review we discuss the criteria that an MCSCF state may be required to fulfill in order to be a valid representation of the state of consideration. We consider requirements that may be imposed to ensure that the response of the MCSCF state to an external one-electron perturbation is physically feasible, for example; that the response calculation is stable and gives the appropriate number of negative excitation energies. We also discuss requirements that may be justified based on the MCSCF states' ability to simulate properties of a full CI solution; for instance, the Hessian matrix for an MCSCF calculation on the n th state may be required to have $n - 1$ negative eigenvalues as does the full CI Hessian matrix.

The MCSCF state should, of course, desirably fulfill all the discussed criteria to be a proper representation of the state of consideration. In many cases, however, it is not possible to get the MCSCF state to satisfy all criteria when a finite basis set is used and only a limited number of configurations are included in the MCSCF calculation. The criteria that may need to be imposed on the MCSCF state as the minimal requirement may then depend on the physical property that the MCSCF state is to calculate. For example, if the MCSCF state is used to evaluate second-order properties it becomes a requirement that the linear MCSCF response calculation give the desired number of negative excitation energies, whereas if the MCSCF state is used only to evaluate dipole moments the requirement that the response calculation give the correct number of negative excitation energies may be less important. A more mathematical and detailed discussion of these points is presented in Section IV.

When the Newton–Raphson approach is applied far from convergence (global convergence), step size control algorithms have been applied in numerical analysis to get the Newton–Raphson approach to converge reliably.⁹ We discuss in Section V how step size and sign control algorithms may be implemented in an MCSCF calculation^{13,15,17} and report a series of nonlinear step size and sign controlled Newton–Raphson MCSCF calculations to illustrate the convergence problems that may appear when optimizing an

MCSCF state using a Newton–Raphson technique. The Newton–Raphson approach is found to converge rapidly and reliably in nearly all cases provided the step sizes and signs are controlled when far from convergence. When impediments to convergence are observed, they can usually be associated to an incorrect structure or small eigenvalues of the Hessian matrix.^{9,17}

While the most important characteristic for global convergence is the reliability of the approach in bringing the calculation into the local region, the most important property for obtaining local convergence is the (computational) efficiency at which the approach is able to bring the calculation from a point relatively near the stationary point to the stationary point. In the local region the convergence characteristics of an iterative procedure may be expressed in terms of a total order of convergence in an error vector that describes the distance from the point of consideration to the desired stationary point.^{8,9} When n steps of such a Newton–Raphson iterative procedure are applied, the Newton–Raphson approach is shown to give a total order of convergence of 2^n in the error vector of the initial iteration. If the Hessian matrix is kept fixed during a sequence of n iterations (the initial fixed Hessian iteration is a Newton–Raphson iteration), the total order of convergence of such a procedure is $n + 1$. Since each iteration in a fixed Hessian sequence requires only construction of the first derivative matrix (the gradient) of the total energy whereas the Newton–Raphson approach requires construction of a Hessian matrix, fixed Hessian procedures may efficiently be used to speed up the local convergence of an MCSCF calculation.^{16,22} This is further discussed in Section VI.

When a fixed Hessian series of iterations is performed, a new energy gradient is calculated at each step of the iterative procedure. Since the difference between the new energy gradient and the energy gradient of the previous iteration contains information about the Hessian matrix at the new point, it is desirable also to incorporate this information into the iterative scheme. Such an incorporation is obtained by updating the Hessian matrix during a sequence of iterations. In the numerical analysis literature Hessian update techniques are considered to be the most efficient method for optimizing a function of many variables.⁹ In Section VII we demonstrate how Hessian update methods may be implemented in MCSCF and how the Hessian update methods constitute an efficient iterative approach for obtaining local convergence of an MCSCF calculation.²¹ Our Hessian update results indicate that the Hessian update methods may also have very promising global convergence characteristics when implemented with a step size control algorithm.

The convergence problem may alternatively be approached using a cubic expansion of the energy.^{16,18,20} In Section VIII we derive and study three cubic optimization procedures. These cubic approaches have desirable local

convergence properties, whereas the iterative cubic approach is particularly useful when far from convergence. In fact, with the iterative cubic technique, step size and sign constraint procedures are often not required.

Since certain of the optimization procedures to be considered herein are designed specifically for the global problem and others for efficiencies in the local region, in Section IX we discuss the effective implementation and combination of MCSCF procedures. In Section X we conclude and summarize.

II. MCSCF OPTIMIZATION USING EXPONENTIAL UNITARY OPERATORS

The initial step of an MCSCF calculation uses an initial guess of a set of orbitals $\langle a_r^+ \rangle$. For example, such a guess may be a set of single configuration Hartree–Fock orbitals with regular virtual orbitals or improved virtual orbitals^{38–39} or grand canonical orbitals.⁴⁰ In Section V we discuss how to get an initial set of orthogonal orbitals of sufficient quality to ensure that an MCSCF iterative procedure will converge. The initial “guess” of orbitals may then be used to calculate an initial multiconfiguration reference state $|0\rangle$, say, by performing a CI calculation within the configurations of the MCSCF calculation. The orthogonal complement set of states $\langle |k\rangle \rangle$ to $|0\rangle$ may also be determined, for example, through a CI calculation (alternatively, see Section IX.B), and MCSCF iterative procedures can then be established that carry out simultaneous unitary transformations within the configuration space $\langle |0\rangle, |k\rangle \rangle$ and orbital space $\langle a_r^+ \rangle$. In this next section we describe how to parameterize a calculation to carry out simultaneous transformations in the orbital and configurational spaces and how to implement MCSCF iterative procedures based on simultaneous orbital and state transformations.

A. Unitary Transformations of the Multiconfigurational Hartree–Fock State

The multiconfigurational Hartree–Fock (MCSCF) reference state $|0\rangle$ may be regarded as a member of the set of states $\langle |j\rangle \rangle$

$$|0\rangle = \sum_g |\Phi_g\rangle C_{g0} \quad (1)$$

$$|k\rangle = \sum_g |\Phi_g\rangle C_{gk} \quad (2)$$

where the coefficient matrix \mathbf{C} forms a unitary matrix. The configuration state functions $|\Phi_g\rangle$ are composed of simple linear combinations of de-

terminants. The determinants $|\Phi_f^D\rangle$ are given as

$$|\Phi_f^D\rangle = \prod_{r \in f} a_r^+ |\text{vac}\rangle \quad (3)$$

where $\prod_{r \in f} a_r^+$ refers to an ordered product of creation operators of electrons in spin orbitals. We assume in the following that all orbitals and states are real. A detailed discussion of how a simultaneous unitary transformation of the orbitals and the state expansion coefficients of the MCSCF reference state may be carried out is given in Refs. 10–12. Below we summarize the results of this derivation.

A unitary transformation among the states $\{|j\rangle\}$ may be described as^{11,12}

$$\exp(i\hat{S})|j\rangle = \sum |k\rangle (\exp(-\mathbf{S}))_{kj} = \sum |k\rangle T_{kj} \quad (4)$$

where

$$\hat{S} = i \sum_{k \neq 0} S_{k0} (|k\rangle\langle 0| - |0\rangle\langle k|) \quad (5)$$

and $\mathbf{T} = \exp(-\mathbf{S})$ is a unitary matrix and \mathbf{S} a real antisymmetric matrix with elements S_{k0} and S_{0k} and zero elsewhere.

A unitary transformation of the orbitals may similarly be described^{10–12} as

$$\tilde{a}_r^+ = \exp(i\hat{\kappa}) a_r^+ \exp(-i\hat{\kappa}) \quad (6)$$

where

$$\hat{\kappa} = i \sum_{r > s} \kappa_{rs} (a_r^+ a_s - a_s^+ a_r) \quad (7)$$

From Eqs. (6) and (7) we get

$$\tilde{a}_r^+ = \sum_s a_s^+ (\exp - \kappa)_{sr} = \sum_s a_s^+ X_{sr} \quad (8)$$

where $\mathbf{X} = \exp(-\kappa)$ is a unitary matrix and κ an antisymmetric matrix with the elements κ_{rs} and κ_{sr} of Eq. (7) and zero elsewhere.

A unitary transformation of the reference state $|0\rangle$ that considers simultaneously a unitary transformation in the orbital and in the configurational space may therefore be described^{11–13} as

$$|\tilde{0}\rangle = \exp(i\hat{\kappa}) \exp(i\hat{S}) |0\rangle \quad (9)$$

Thus the states

$$|\tilde{k}\rangle = \exp(i\kappa)\exp(i\hat{S})|k\rangle \quad (10)$$

together with $|\tilde{0}\rangle$ will form an orthonormal set. That is, through specifying a set of parameters κ and S we may generate an arbitrary state $|\tilde{0}\rangle$ and ensure that this state together with the states $|\tilde{k}\rangle$ form an orthonormal set of states $\{|\tilde{j}\rangle\}$. In what follows, for convenience we sometimes use $\underline{\lambda}$ to denote the set of rotational parameters arranged as a column vector

$$\begin{pmatrix} \kappa \\ \mathbf{s} \end{pmatrix}$$

and occasionally refer to $\underline{\lambda}$ as the step length vector. The specific choice of step length vector $\underline{\lambda} = \mathbf{0}$ represents, of course, the untransformed state $|0\rangle$.

The set of vectors $(a_r^+ a_s - a_s^+ a_r)|0\rangle$, $(|k\rangle\langle 0| - |0\rangle\langle k|)|0\rangle$ may be linearly dependent, and some variables may be redundant when an optimization is carried out. The elimination of linear dependencies and redundant operators for a specific choice of configurations in the reference state is discussed in detail in Appendix A.^{10,17}

B. Variations in the Total Energy

The total energy corresponding to the unitary transformation of the reference state given in Eq. (9) may be written¹⁰⁻¹² as

$$E(\underline{\lambda}) = E(\underline{\kappa}, \mathbf{s}) = \langle \tilde{0} | H | \tilde{0} \rangle = \langle 0 | \exp(-i\hat{S}) \exp(-i\kappa) H \exp(i\kappa) \exp(i\hat{S}) | 0 \rangle \quad (11)$$

The total energy is thus defined in terms of a set of rotational parameters, $\underline{\lambda}$. A Taylor series expansion may be carried out in these rotational parameters at $\underline{\lambda} = {}^0\underline{\lambda}$:

$$\begin{aligned} E(\underline{\lambda}) &= E({}^0\underline{\lambda}) + \left. \frac{\partial E(\underline{\lambda})}{\partial \lambda_i} \right|_{\underline{\lambda} = {}^0\underline{\lambda}} [\lambda_i - {}^0\lambda_i] \\ &\quad + \frac{1}{2} \left. \frac{\partial^2 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j} \right|_{\underline{\lambda} = {}^0\underline{\lambda}} [\lambda_i - {}^0\lambda_i][\lambda_j - {}^0\lambda_j] \\ &\quad + \frac{1}{6} \left. \frac{\partial^3 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j \partial \lambda_k} \right|_{\underline{\lambda} = {}^0\underline{\lambda}} [\lambda_i - {}^0\lambda_i][\lambda_j - {}^0\lambda_j][\lambda_k - {}^0\lambda_k] \\ &\quad + \frac{1}{24} \left. \frac{\partial^4 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j \partial \lambda_k \partial \lambda_l} \right|_{\underline{\lambda} = {}^0\underline{\lambda}} [\lambda_i - {}^0\lambda_i][\lambda_j - {}^0\lambda_j] \\ &\quad \times [\lambda_k - {}^0\lambda_k][\lambda_l - {}^0\lambda_l] \\ &\quad + \dots \end{aligned} \quad (12)$$

Note that here and in all subsequent equations, unless noted otherwise, for convenience we use the Einstein summation convention.

Introducing the supermatrix notation for the partial derivatives

$$\frac{\partial E(\underline{\lambda})}{\partial \lambda_i} \Big|_{\underline{\lambda} = {}^0\underline{\lambda}} = F_i({}^0\underline{\lambda}) \quad (13)$$

$$\frac{\partial^2 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j} \Big|_{\underline{\lambda} = {}^0\underline{\lambda}} = G_{ij}({}^0\underline{\lambda}) \quad (14)$$

$$\frac{\partial^3 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j \partial \lambda_k} \Big|_{\underline{\lambda} = {}^0\underline{\lambda}} = K_{ijk}({}^0\underline{\lambda}) \quad (15)$$

$$\frac{\partial^4 E(\underline{\lambda})}{\partial \lambda_i \partial \lambda_j \partial \lambda_k \partial \lambda_l} \Big|_{\underline{\lambda} = {}^0\underline{\lambda}} = M_{ijkl}({}^0\underline{\lambda}) \quad (16)$$

allows us to express the Taylor expansion in a compact form:

$$\begin{aligned} E(\underline{\lambda}) = & E({}^0\underline{\lambda}) + F_i({}^0\underline{\lambda}) [\lambda_i - {}^0\lambda_i] + \tfrac{1}{2} G_{ij}({}^0\underline{\lambda}) [\lambda_i - {}^0\lambda_i] [\lambda_j - {}^0\lambda_j] \\ & + \tfrac{1}{6} K_{ijk}({}^0\underline{\lambda}) [\lambda_i - {}^0\lambda_i] [\lambda_j - {}^0\lambda_j] [\lambda_k - {}^0\lambda_k] \\ & + \tfrac{1}{24} M_{ijkl}({}^0\underline{\lambda}) [\lambda_i - {}^0\lambda_i] [\lambda_j - {}^0\lambda_j] [\lambda_k - {}^0\lambda_k] [\lambda_l - {}^0\lambda_l] + \dots \end{aligned} \quad (17)$$

The partial derivatives in Eqs. (13–16) are cumbersome to evaluate at a general point ${}^0\underline{\lambda}$. In Appendix B we have derived an explicit expression for the first partial derivative $F_i({}^0\underline{\lambda})$. The second partial derivative is approximately one order of magnitude more difficult to evaluate than the first partial derivative; the third partial derivative is approximately an order of magnitude more complicated to evaluate than the second partial derivative; and so on. General expressions for the second and higher partial derivatives evaluated at an arbitrary point ${}^0\underline{\lambda}$ will not be given.

The vector \mathbf{F} and matrix \mathbf{G} represent the slope and curvature of the energy hypersurface, respectively. \mathbf{F} is often referred to as the energy gradient vector, while \mathbf{G} is often denoted the Hessian matrix. The MCSCF reference state represents a stationary point on the energy hypersurface. At such a point, the first order variation in the total energy is zero, that is, $\delta E(\underline{\lambda}) = 0$. The condition that $\delta E(\underline{\lambda}) = 0$ may be used to determine the step length to take on the energy hypersurface in order to reach a stationary point. From Eq. (17), the first-order variation of the total energy may be used to

obtain $\underline{\lambda}$:

$$F_i({}^0\underline{\lambda}) + G_{ij}({}^0\underline{\lambda})[\lambda_j - {}^0\lambda_j] + \frac{1}{2}K_{Ijk}({}^0\underline{\lambda})[\lambda_j - {}^0\lambda_j][\lambda_k - {}^0\lambda_k] + \dots = 0 \quad (18)$$

At the stationary point $\underline{\lambda} = \underline{\alpha}$,

$$\mathbf{F}(\underline{\alpha}) = \mathbf{0} \quad (18a)$$

If we could determine a set of $\underline{\lambda}$ parameters from Eq. (18) without truncating this equation at any order of $\underline{\lambda}$, we would reach a stationary point on the energy hypersurface in one iteration. The iterative nature of MCSCF approaches originates from using truncated forms of Eq. (18) to determine the stationary points on the energy hypersurface. These truncated forms require successive applications to determine a stationary point on the energy hypersurface.

If Eq. (18) is truncated after the second term, which corresponds to truncating Eq. (12) after terms quadratic in $[\underline{\lambda} - {}^0\underline{\lambda}]$, second-order approaches are derived. These are discussed in detail in Section III. If Eq. (17) is truncated after the third terms [cubic in $[\underline{\lambda} - {}^0\underline{\lambda}]$ in Eq. (12)] cubic approaches may be developed. These are studied in Section VIII.

III. THE NEWTON-RAPHSON ITERATIVE FUNCTION

As a first example, we consider the Newton-Raphson iterative approach. The Newton-Raphson iterative function is derived by truncating the energy function [Eq. (12)] through the second power in the rotational parameters.¹⁰⁻¹² The Newton-Raphson approach straightforwardly illustrates how the local convergence of an iterative procedure may be rationalized in terms of a total order of convergence in an error vector.^{16, 22} When the Newton-Raphson approach is implemented in numerical analysis, a linear transformation of the variables is performed between each step of the iterative procedure.⁹ We demonstrate how it often may be more convenient when using the Newton-Raphson MCSCF approach to perform a nonlinear transformation of the variables between each step of the iterative procedure when orbital optimizations are carried out. The local convergence of the Newton-Raphson approach is then discussed, and we demonstrate that a sequence of n linear or nonlinear Newton-Raphson MCSCF iterations will have a total order of convergence of 2^n in the error vector of the initial iteration.

A. The Linear Newton–Raphson Iterative Function

The Newton–Raphson iterative function is derived from Eq. (18) by neglecting the nonlinear terms:

$$F_i({}^0\lambda) + G_{ij}({}^0\lambda)[\lambda_j - {}^0\lambda_j] = 0 \quad (19)$$

The Newton–Raphson iterative function therefore becomes

$${}^k\lambda_i = {}^{k-1}\lambda_i - G_{ij}^{-1}({}^{k-1}\lambda) F_j({}^{k-1}\lambda) \quad (20)$$

where ${}^k\lambda$ denotes the set of λ parameters determined from the k th iteration of Eq. (20). The \mathbf{F} and \mathbf{G} matrices are evaluated at the $(k-1)$ st iteration point, ${}^{k-1}\lambda$. A sequence of iterations which uses Eq. (20) as an iterative function will result in a linear transformation of the variables λ in each step of the iterative sequence.

In the following we discuss in a little more detail how a sequence of iterations may be carried out which uses Eq. (20) as an iterative function. The first step of a Newton–Raphson sequence of iterations from an initial guess of orbitals and states (${}^0\lambda = \mathbf{0}$) is to apply Eq. (20) to determine a set of parameters

$${}^1\lambda - {}^0\lambda = {}^1\lambda = \begin{pmatrix} {}^1\kappa \\ {}^1\mathbf{S} \end{pmatrix}.$$

This parameter set describes how rotations of the orbitals and states are carried out relative to the parameter point ${}^0\lambda = \mathbf{0}$. The partial derivatives in Eq. (20) are evaluated at ${}^0\lambda = \mathbf{0}$ in the initial iteration only. The second step of the Newton–Raphson sequence is then to determine a parameter set

$${}^2\lambda - {}^1\lambda = \begin{pmatrix} {}^2\kappa - {}^1\kappa \\ {}^2\mathbf{S} - {}^1\mathbf{S} \end{pmatrix}$$

that describes how rotations are carried out relative to the point ${}^1\lambda - {}^0\lambda = {}^1\lambda$. In the second step of the iterative procedure, the partial derivatives in Eq. (20) need to be evaluated at ${}^1\lambda$ (no orbital or state transformations are carried out between each step of the iterative procedure). The third step of the Newton–Raphson iterative procedure determines ${}^3\lambda - {}^2\lambda$, which requires partial derivatives to be evaluated at ${}^2\lambda$. And so on. Thus when n steps of Eq. (20) are applied, the transformed orbitals after the n th iteration are de-

terminated as

$${}^n\tilde{a}_r^+ = \exp(i\hat{\kappa}(n)) a_s^+ \exp(-i\hat{\kappa}(n)) \quad (21)$$

where

$$\hat{\kappa}(n) = i \sum_{r > s} {}^n\kappa_{rs} (a_r^+ a_s - a_s^+ a_r) \quad (22)$$

Defining the matrix

$${}^n\mathbf{X} = \exp({}^n\kappa) \quad (23)$$

Eq. (21) may be rewritten using Eq. (8):

$${}^n\tilde{a}_r^+ = \sum_s a_s^+ ({}^n\mathbf{X})_{sr} \quad (24)$$

The transformed MCSCF reference state after n steps of such an iterative procedure becomes

$$|{}^n0\rangle = \exp(i\hat{\kappa}(n)) \exp(i\hat{S}(n)) |0\rangle \quad (25)$$

and the coefficients of the states $\{|{}^n0\rangle, |{}^n k\rangle\}$ thus are determined as

$${}^n\mathbf{C} = {}^0\mathbf{C} \exp(-{}^n\mathbf{S}) \quad (26)$$

where ${}^0\mathbf{C}$ is the matrix of the initial guess (note $|{}^00\rangle = |0\rangle$). Equations (24)–(26) demonstrate that a straightforward application of the Newton–Raphson iterative function results in an iterative procedure in which a linear transformation of the vector amplitudes

$$\underline{\lambda} = \begin{pmatrix} \underline{\kappa} \\ \underline{\mathbf{S}} \end{pmatrix}$$

is carried out between each step of the iterative procedure.

We now examine the local convergence properties of a sequence of n Newton–Raphson iterations. Consider initially how the $\mathbf{F}(\underline{\lambda})$ and $\mathbf{G}(\underline{\lambda})$ matrix elements may be expanded around the stationary point $\underline{\alpha}$

$$F_i(\underline{\lambda}) = G_{ij}(\underline{\alpha}) [\lambda_j - \alpha_j] + \frac{1}{2} K_{ijk}(\underline{\alpha}) [\lambda_j - \alpha_j][\lambda_k - \alpha_k] + \dots \quad (27)$$

$$G_{ij}(\underline{\lambda}) = G_{ij}(\underline{\alpha}) + K_{ijk}(\underline{\alpha}) [\lambda_k - \alpha_k] + \dots \quad (28)$$

since $F_i(\underline{\alpha}) = 0$. The k th step of a Newton–Raphson sequence of iterations in Eq. (20) may now be rewritten using Eqs. (27) and (28)

$$\begin{aligned} {}^k\lambda_i &= {}^{k-1}\lambda_i - \left(G_{ij}(\underline{\alpha}) + K_{ijm}(\underline{\alpha}) \left[{}^{k-1}\lambda_m - \alpha_m \right] + \dots \right)^{-1} \\ &\quad \times \left(G_{ij}(\underline{\alpha}) \left[{}^{k-1}\lambda_j - \alpha_j \right] \right. \\ &\quad \left. + \frac{1}{2} K_{ijm}(\underline{\alpha}) \left[{}^{k-1}\lambda_j - \alpha_j \right] \left[{}^{k-1}\lambda_m - \alpha_m \right] + \dots \right) \\ &= \alpha_i + \frac{1}{2} G_{in}^{-1}(\underline{\alpha}) K_{njm}(\underline{\alpha}) \left[{}^{k-1}\lambda_j - \alpha_j \right] \left[{}^{k-1}\lambda_m - \alpha_m \right] + O({}^{k-1}\underline{\lambda} - \underline{\alpha})^3 \end{aligned} \quad (29)$$

where we have expanded the inverse matrix. The k th Newton–Raphson iteration thus contains errors of order 2 in the error vector

$${}^{k-1}\mathbf{e} = {}^{k-1}\underline{\lambda} - \underline{\alpha} \quad (30)$$

of the $(k-1)$ st iteration. The error term of the k th iteration may be written as

$${}^k e_i = \frac{1}{2} G_{in}^{-1}(\underline{\alpha}) K_{njl}(\underline{\alpha}) {}^{k-1} e_j {}^{k-1} e_l \quad (31)$$

When a sequence of two Newton–Raphson iterations are carried out after an initial guess of orbitals, we obtain a total order of convergence of 4 in ${}^0\mathbf{e}$ (the error vector of the initial iteration) with an error term

$${}^2 e_i = \frac{1}{2} L_{ijk} {}^1 e_j {}^1 e_k = \frac{1}{8} L_{ijk} L_{jmn} {}^0 e_m {}^0 e_n L_{kpq} {}^0 e_p {}^0 e_q \quad (32)$$

where

$$L_{ijk} = G_{in}^{-1}(\underline{\alpha}) K_{njk}(\underline{\alpha}) \quad (33)$$

After n Newton–Raphson iterations are performed, the total order of convergence is 2^n in ${}^0\mathbf{e}$ and the error term becomes

$${}^n e_i = \frac{1}{2} L_{ijk} {}^{n-1} e_j {}^{n-1} e_k = 2^{-(2^n-1)} L_{ik_1 k_2 \dots k_{(2^n)}}^{(n)} {}^0 e_{k_1} {}^0 e_{k_2} \dots {}^0 e_{k_{(2^n)}} \quad (34)$$

where L is defined recursively

$$L_{ik_1 k_2 \dots k_{(2^n)}}^{(n)} = L_{il_1 l_2} L_{l_1 k_1 k_2 \dots k_{(2^{n-1})}}^{(n-1)} L_{l_2 k_{(2^{n-1}+1)} k_{(2^{n-1}+2)} \dots k_{(2^n)}}^{(n-1)} \quad (35)$$

and

$$L_{ik_1k_2}^{(1)} = L_{ik_1k_2} \quad (36)$$

We have now described how a sequence of linear Newton-Raphson iterations may be performed and that a sequence consisting of n iterations has a total order of convergence of 2^n in the error vector of the initial iteration.

B. The Nonlinear Newton-Raphson Iterative Function

The Newton-Raphson MCSCF approach of the previous section is just a standard application of the Newton-Raphson technique discussed in elementary textbooks of numerical analysis. The application of the Newton-Raphson approach requires evaluation of the first and second derivatives [Eqs. (13) and (14)] of the energy function at a general point. The second partial derivative is rather cumbersome to evaluate at a general point in MCSCF. Partial derivatives of the energy function are in general much easier to evaluate at the single point $\underline{\lambda} = \mathbf{0}$ than at a general point. This fact may be used to modify the Newton-Raphson approach to a form that is practically applicable in MCSCF.

According to Eq. (9) a transformed state $|\tilde{0}\rangle$ is defined in terms of a set of orbital operators $\langle a_r^+ \rangle$ and states $\langle |j\rangle \rangle$ and a set of rotational parameters $\underline{\lambda}$. As previously discussed, the parameter set $\underline{\lambda} = \mathbf{0}$ represents the initial guess of the reference state $|0\rangle$ and orbitals. The initial Newton-Raphson iteration requires only evaluation of partial derivatives at $\underline{\lambda} = \mathbf{0}$. The initial Newton-Raphson iteration determines a parameter set ${}^1\underline{\lambda}$ which through Eqs. (8) and (9) defines the rotations that have to be carried out to determine transformed sets of orbital operators $\langle {}^1\tilde{a}_r^+ \rangle$ and states $\langle |{}^1\tilde{j}\rangle \rangle$. These transformed orbitals and states may then be considered as a new initial guess. The Newton-Raphson iterative procedure may now be applied with this new "initial" guess of orbitals and states. The parameter set ${}^2\underline{\lambda}$ determined in the second Newton-Raphson iteration then describes how the rotations of the orbital operators $\langle {}^1\tilde{a}_r^+ \rangle$ and states $\langle |{}^1\tilde{j}\rangle \rangle$ have to be carried out [use Eqs. (8) and (9)] to determine a new transformed set of orbitals $\langle {}^2\tilde{a}_r^+ \rangle$ and states $\langle |{}^2\tilde{j}\rangle \rangle$, and so on. We note that $\langle {}^n\tilde{a}_r^+ \rangle$ and $\langle |{}^n\tilde{j}\rangle \rangle$ are used in iteration $n+1$ (at iteration point n) to determine ${}^{n+1}\underline{\lambda}$ since $n=0$ is the initial guess of orbitals and states (i.e., $\langle {}^0\tilde{a}_r^+ \rangle = \langle a_r^+ \rangle$ and $\langle |{}^0\tilde{j}\rangle \rangle = \langle |j\rangle \rangle$).

After n applications of such a sequence of Newton-Raphson iterations, n unitary transformations have been carried out on the orbitals and the states. The orbital operators after n such unitary transformations may be described as

$$\begin{aligned} {}^n\tilde{a}_r^+ &= \exp[i^n\hat{k}(n)] \cdots \exp[i^2\hat{k}(2)] \exp[i^1\hat{k}(1)] a_r^+ \\ &\times \exp[-i^1\hat{k}(1)] \exp[-i^2\hat{k}(2)] \cdots \exp[-{}^n\hat{k}(n)] \end{aligned} \quad (37)$$

where ${}^p\hat{\kappa}(m)$ is defined as

$${}^p\hat{\kappa}(m) = i \sum {}^m \kappa_{rs} [{}^p\tilde{a}_r^+ {}^p\tilde{a}_s - {}^p\tilde{a}_s^+ {}^p\tilde{a}_r] \quad (38)$$

Henceforth, the tilde~ and parentheses () will generally be omitted when reference is made to the initial basis. Note that in Eq. (38) m and p are not necessarily equal; that is, p refers to the number of orbital operator transformations and m refers to the parameter set determined in the m th step of the iterative procedure.

The ${}^p\tilde{a}_r^+$ in Eq. (37) may also be expressed in terms of the initial set of creation and annihilation operators. Using the relation

$$\exp[i({}^2\hat{\kappa}(2))] = \exp[i\hat{\kappa}(1)] \exp[i\hat{\kappa}(2)] \exp[-i\hat{\kappa}(1)], \quad (39)$$

Eq. (37) may be written as¹⁴

$${}^p\tilde{a}_r^+ = \exp[i\hat{\kappa}(1)] \cdots \exp[i\hat{\kappa}(n)] a_r^+ \exp[-i\hat{\kappa}(n)] \cdots \exp[-i\hat{\kappa}(1)], \quad (40)$$

where $\hat{\kappa}(m)$ is expressed in terms of the initial set of creation and annihilation operators:

$$\hat{\kappa}(m) = i \sum {}^m \kappa_{rs} [a_r^+ a_s - a_s^+ a_r] \quad (41)$$

The evaluation of the transformed set of creation operators may be performed as

$${}^p\tilde{a}_r^+ = \sum_s a_s^+ [\exp(-{}^1\kappa) \exp(-{}^2\kappa) \cdots \exp(-{}^n\kappa)]_{sr} \quad (42)$$

where the exponential matrices are defined through Eq. (8).

When n successive unitary transformations have been carried out on the MCSCF reference state, we obtain

$$\begin{aligned} |{}^n0\rangle &= \exp[i{}^n\hat{\kappa}(n)] \exp[i{}^n\hat{S}(n)] \cdots \exp[i{}^2\hat{\kappa}(2)] \\ &\times \exp[i{}^2\hat{S}(2)] \exp[i{}^1\hat{\kappa}(1)] \exp[i{}^1\hat{S}(1)] |0\rangle \end{aligned} \quad (43)$$

where again the n ~ denotes that the operators are expressed in the basis obtained after n successive unitary transformations of the orbitals and the states. Equation (38) and successive application of the relation

$$\begin{aligned} \exp[i({}^2\hat{S}(2))] &= \exp[i\hat{\kappa}(1)] \exp[i\hat{S}(1)] \exp[i\hat{S}(2)] \\ &\times \exp[-i\hat{S}(1)] \exp[-i\hat{\kappa}(1)] \end{aligned} \quad (44)$$

allow us to express $|^n0\rangle$ in Eq. (43) in terms of the initial set of orbitals and states:¹⁴

$$\begin{aligned} |^n0\rangle &= \exp[i\hat{\kappa}(1)]\exp[i\hat{\kappa}(2)]\cdots\exp[i\hat{\kappa}(n)] \\ &\quad \times \exp[i\hat{S}(1)]\exp[i\hat{S}(2)]\cdots\exp[i\hat{S}(n)]|0\rangle \end{aligned} \quad (45)$$

The coefficients for the states $\{|^n0\rangle, |^nk\rangle\}$ therefore may be determined from the relation

$${}^nC = {}^0C \exp(-{}^1S) \exp(-{}^2S) \cdots \exp(-{}^nS) \quad (46)$$

The orbitals and states used in the $(n+1)$ st iteration are thus determined by multiplying n exponential matrices each containing a set of rotational parameters.

Equations (42) and (46) clearly display that a nonlinear transformation of the rotational parameters is carried out between each step of the iterative procedure. The basis of orbitals in which we express our energy function (and consequently our iterative function) is changed at each step of the iterative process in such a nonlinear iterative procedure.

We shall now discuss how the partial derivatives of the total energy may be relatively straightforwardly evaluated at the point $\underline{\lambda} = \mathbf{0}$. This fact is used for practical MCSCF implementation of the Newton-Raphson algorithm but also in many other practical implementations of MSCSF iterative schemes. At this initial point $\underline{\lambda} = \mathbf{0}$, the partial derivatives may be evaluated by expanding the exponential operators

$$\begin{aligned} E(\underline{\lambda}) &= \langle \tilde{0} | H | \tilde{0} \rangle = \langle 0 | \exp(-i\hat{S}) \exp(-i\hat{\kappa}) H \exp(i\hat{\kappa}) \exp(i\hat{S}) | 0 \rangle \\ &= E(\mathbf{0}) - i \langle 0 | [\hat{S} + \hat{\kappa}, H] | 0 \rangle - \frac{1}{2} \langle 0 | [\hat{S}, [\hat{S}, H]] | 0 \rangle - \frac{1}{2} \langle 0 | [\hat{\kappa}, [\hat{\kappa}, H]] | 0 \rangle \\ &\quad - \langle 0 | [\hat{S}, [\hat{\kappa}, H]] | 0 \rangle + \frac{i}{6} \langle 0 | [\hat{S}, [\hat{S}, [\hat{S}, H]]] | 0 \rangle \\ &\quad + \frac{i}{6} \langle 0 | [\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, H]]] | 0 \rangle + \frac{i}{2} \langle 0 | [\hat{S}, [\hat{S}, [\hat{\kappa}, H]]] | 0 \rangle \\ &\quad + \frac{i}{2} \langle 0 | [\hat{S}, [\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, H]]]] | 0 \rangle + \frac{1}{24} \langle 0 | [\hat{S}, [\hat{S}, [\hat{S}, [\hat{S}, H]]]] | 0 \rangle \\ &\quad + \frac{1}{24} \langle 0 | [\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, H]]]] | 0 \rangle + \frac{1}{6} \langle 0 | [\hat{S}, [\hat{S}, [\hat{S}, [\hat{\kappa}, H]]]] | 0 \rangle \\ &\quad + \frac{1}{4} \langle 0 | [\hat{S}, [\hat{S}, [\hat{\kappa}, [\hat{\kappa}, H]]]] | 0 \rangle + \frac{1}{6} \langle 0 | [\hat{S}, [\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, H]]]] | 0 \rangle \\ &\quad + \cdots \end{aligned} \quad (47)$$

All terms through fourth order in $\hat{\kappa}$ and \hat{S} are written out explicitly. The n

tupple symmetric commutator for the operators D_1, D_2, \dots, D_n may be defined^{16, 20} as

$$[D_1, D_2, \dots, D_n, H] = \frac{1}{n!} P(1, 2, \dots, n) [D_1, [D_2, \dots, [D_n, H] \dots]] \quad (48)$$

where $P(1, 2, \dots, n)$ is a permutation operator which contains the $n!$ permutations of the indices $1, 2, \dots, n$. Equation (47) may then be written as

$$\begin{aligned} E(\hat{\kappa}, \hat{S}) = & \langle 0 | H | 0 \rangle - i \langle 0 | [\hat{S} + \hat{\kappa}, h] | 0 \rangle - \frac{1}{2} \langle 0 | [\hat{S}, \hat{S}, H] | 0 \rangle \\ & - \frac{1}{2} \langle 0 | [\hat{\kappa}, \hat{\kappa}, H] | 0 \rangle - \langle 0 | [\hat{S}, [\hat{\kappa}, H]] | 0 \rangle + \frac{i}{6} \langle 0 | [\hat{S}, \hat{S}, \hat{S}, H] | 0 \rangle \\ & + \frac{i}{6} \langle 0 | [\hat{\kappa}, \hat{\kappa}, \hat{\kappa}, H] | 0 \rangle + \frac{i}{2} \langle 0 | [\hat{S}, \hat{S}, [\hat{\kappa}, H]] | 0 \rangle \\ & + \frac{i}{2} \langle 0 | [\hat{S}, [\hat{\kappa}, \hat{\kappa}, H]] | 0 \rangle + \frac{1}{24} \langle 0 | [\hat{S}, \hat{S}, \hat{S}, \hat{S}, H] | 0 \rangle \\ & + \frac{1}{24} \langle 0 | [\hat{\kappa}, \hat{\kappa}, \hat{\kappa}, \hat{\kappa}, H] | 0 \rangle + \frac{1}{6} \langle 0 | [\hat{S}, \hat{S}, \hat{S}, [\hat{\kappa}, H]] | 0 \rangle \\ & + \frac{1}{4} \langle 0 | [\hat{S}, \hat{S}, [\hat{\kappa}, \hat{\kappa}, H]] | 0 \rangle + \frac{1}{6} \langle 0 | [\hat{S}, [\hat{\kappa}, \hat{\kappa}, \hat{\kappa}, H]] | 0 \rangle \\ & + \dots \end{aligned} \quad (49)$$

We have used the fact that each of the operators $\hat{\kappa}$ and \hat{S} commutes with itself.

By using a supermatrix notation where the variational parameters $\{\kappa_{rs}\}$ and $\{S_{n0}\}$ and the excitation operators

$$\mathbf{Q}^+ = \{a_r^+ a_s\}, r > s; \mathbf{R}^+ = (\langle n \rangle \langle 0 |) \quad (50)$$

both form column vectors

$$\underline{\lambda} = \begin{pmatrix} \kappa \\ \mathbf{S} \end{pmatrix} \quad (51)$$

$$\mathbf{T}^+ = \begin{pmatrix} \mathbf{Q}^+ \\ \mathbf{R}^+ \end{pmatrix} \quad (52)$$

we can write the total energy in Eq. (49) as

$$\begin{aligned} E(\underline{\lambda}) = & E(\mathbf{0}) + F_i(\mathbf{0})\lambda_i + \frac{1}{2}G_{ij}(\mathbf{0})\lambda_i\lambda_j \\ & + \frac{1}{6}K_{ijk}(\mathbf{0})\lambda_i\lambda_j\lambda_k + \frac{1}{24}M_{ijkl}(\mathbf{0})\lambda_i\lambda_j\lambda_k\lambda_l + \dots \end{aligned} \quad (53)$$

where the matrices $F_i(\mathbf{0})$, $G_{ij}(\mathbf{0})$, $K_{ijk}(\mathbf{0})$, and $M_{ijkl}(\mathbf{0})$ are identified as

$$F_i(\mathbf{0}) = \langle 0 | [T_i^+ - T_i, H] | 0 \rangle \quad (54)$$

$$G_{ij}(\mathbf{0}) = \langle 0 | [T_i^+ - T_i, T_j^+ - T_j, H] | 0 \rangle \quad (55)$$

$$K_{ijk}(\mathbf{0}) = \langle 0 | [T_i^+ - T_i, T_j^+ - T_j, T_k^+ - T_k, H] | 0 \rangle \quad (56)$$

$$M_{ijkl}(\mathbf{0}) = \langle 0 | [T_i^+ - T_i, T_j^+ - T_j, T_k^+ - T_k, T_l^+ - T_l, H] | 0 \rangle \quad (57)$$

Again note that the Einstein summation convention is used. All terms in the n -tuple commutator that couple the configuration and orbital space are defined such that the Hamiltonian first operates on the orbital space excitation operators and then on the configuration space excitation operators; for example; all the six terms

$$\begin{aligned} [\mathbf{R}^+, \mathbf{R}, \mathbf{Q}, \mathbf{Q}^+, H] &= [\mathbf{Q}, \mathbf{Q}^+, \mathbf{R}^+, \mathbf{R}, H] = [\mathbf{R}, \mathbf{Q}, \mathbf{Q}^+, \mathbf{R}^+, H] \\ &= [\mathbf{R}^+, \mathbf{Q}, \mathbf{Q}^+, \mathbf{R}, H] = [\mathbf{R}, \mathbf{R}^+, \mathbf{Q}, \mathbf{Q}^+, H] \\ &= [\mathbf{Q}, \mathbf{Q}^+, \mathbf{R}, \mathbf{R}^+, H] \end{aligned}$$

are defined to equal

$$[\mathbf{R}^+, \mathbf{R}, [\mathbf{Q}, \mathbf{Q}^+, H]] \quad (58)$$

Equations (54)–(57) thus give explicit expressions for partial derivatives at $\underline{\lambda} = \mathbf{0}$. When partial derivatives are evaluated at the point ${}^0\underline{\lambda} = \mathbf{0}$ we omit reference to this point in the partial derivative matrices. Explicit expressions for evaluating the matrices \mathbf{F} and \mathbf{G} are given in Section III and Appendix C.

Stationary points on the energy hypersurface satisfy the condition $\delta E = 0$, that is,

$$F_i + G_{ij}\lambda_j + \frac{1}{2}K_{ijk}\lambda_j\lambda_k + \frac{1}{6}M_{ijkl}\lambda_j\lambda_k\lambda_l + \dots = 0 \quad (59)$$

The nonlinear Newton–Raphson iterative function therefore becomes

$$\underline{\lambda} = -\mathbf{G}^{-1}\mathbf{F} \quad (60)$$

We now discuss the convergence properties when a sequence of k nonlinear Newton–Raphson iterations is carried out. The n th nonlinear Newton–Raphson iteration formally may be written as

$${}^n\tilde{\lambda}_i = - \tilde{G}_{ij}^{-1} {}^{n-1}\tilde{F}_j \quad (61)$$

where here ${}^{n-1}$ denotes that the matrices are evaluated for successively transformed orbitals and states involving ${}^1\tilde{\lambda}, {}^2\tilde{\lambda}, \dots, {}^{n-1}\tilde{\lambda}$ obtained from Eq. (61). The orbital and state transformations are carried out as described in Eqs. (42) and (46). The error term of the n th iteration may be expressed as

$${}^n\tilde{e}_i = \frac{1}{2} \tilde{L}_{ijl} {}^{n-1}\tilde{e}_j {}^{n-1}\tilde{e}_l \quad (62)$$

where we have used Eqs. (31) and (33). The error vector of the $n-1$ iteration may similarly be written as

$${}^{n-1}\tilde{e}_i = \frac{1}{2} \tilde{L}_{ijl} {}^{n-2}\tilde{e}_j {}^{n-2}\tilde{e}_l \quad (63)$$

The ${}^{n-1}$ basis is related to the ${}^{n-2}$ through multiplications involving the exponential unitary matrices $\exp(-{}^{n-1}\mathbf{k})$ and $\exp(-{}^{n-1}\mathbf{S})$. The matrices of the ${}^{n-1}$ basis may therefore be expressed in the ${}^{n-2}$ basis as

$$\tilde{L}_{ijl} = \tilde{L}_{ijl} (1 + O({}^{n-1}\tilde{\lambda})) \quad (64)$$

$${}^n\tilde{e}_i = {}^{n-2}\tilde{e}_i (1 + O({}^{n-1}\tilde{\lambda})) \quad (65)$$

Since ${}^{n-1}\tilde{\lambda}$ contains error terms of the order $O(({}^{n-2}{}^{n-2}\tilde{e})^2)$, Eqs. (64) and (65) may be written as

$$\tilde{L}_{ijl} = \tilde{L}_{ijl} \left(1 + O\left(\left({}^{n-2}{}^{n-2}\tilde{e}\right)^2\right) \right) \quad (66)$$

$${}^n\tilde{e}_i = {}^{n-2}\tilde{e}_i \left(1 + O\left(\left({}^{n-2}{}^{n-2}\tilde{e}\right)^2\right) \right) \quad (67)$$

and, when we use Eqs. (63), (66), and (67), Eq. (62) becomes

$${}^n\tilde{e}_i = \frac{1}{8} \tilde{L}_{ijl} \tilde{L}_{jmn} {}^{n-2}\tilde{e}_m {}^{n-2}\tilde{e}_n {}^{n-2}\tilde{e}_q {}^{n-2}\tilde{e}_r \quad (68)$$

Successive application of Eqs. (63), (66), and (67) then shows that the error term of the n th iteration of a nonlinear Newton-Raphson iteration ${}^{n-1}\tilde{e}$ becomes identical to the error term ${}^n\mathbf{e}$ of Eq. (34) obtained after n linear Newton-Raphson iterations. We have thus proved that n steps of a Newton-Raphson iteration procedure that performs either a linear or a nonlinear transformation of the variables between each step of the iterative procedure have the same order of convergence and contain the same error terms. Since the nonlinear and linear Newton-Raphson approaches have the same convergence properties and since the nonlinear Newton-Raphson approach only requires evaluation of partial derivatives at $\underline{\lambda} = \mathbf{0}$, practical applications of the MCSCF Newton-Raphson iterative approach have exclusively used a nonlinear approach.^{1,2,22} Henceforth, for convenience, we will not use tilde with the nonlinear Newton-Raphson procedure.

Thus, the nomenclature we use is generally as follows. At iteration point $n - 1$ the energy is evaluated, \mathbf{F} and \mathbf{G} are constructed, and ${}^n\underline{\lambda}$ is determined. Hence at iteration point $n - 1$ the n th iteration is performed to obtain the rotational parameters ${}^n\underline{\lambda}$.

To understand the local convergence characteristics of an iterative calculation, it is necessary to get an estimate of the magnitude of the error vector at each step of the iterative procedure. The error vector describes the distance between the actual parameter values and the parameter values of the stationary point. Since the parameter values of the stationary point first are determined at convergence, an accurate measure of the error vector in principle cannot be obtained before convergence. However, the error vector at iteration point $n - 1$ is

$${}^{n-1}\mathbf{e} = {}^{n-1}\underline{\lambda} - \underline{\alpha} = {}^{n-1}\underline{\lambda} - {}^n\underline{\lambda} + {}^n\underline{\lambda} - \underline{\alpha} = {}^{n-1}\underline{\lambda} - {}^n\underline{\lambda} + {}^n\mathbf{e} \quad (69)$$

If the length of the error vector at iteration point $n - 1$ is much smaller than at iteration point n ,

$$\|{}^{n-1}\mathbf{e}\| \gg \|{}^n\mathbf{e}\| \quad (70)$$

then a rough estimate of the size of the error vector at iteration point $n - 1$ becomes the length of the set of rotational parameters that are determined in the n th iteration

$$\|{}^{n-1}\mathbf{e}\| \approx \|{}^{n-1}\underline{\lambda} - {}^n\underline{\lambda}\| \quad (71)$$

Equation (70) is relatively well satisfied when a second- or higher-order iterative scheme is applied in the local region, and Eq. (71) then becomes a

reasonable measure of the length of the error vector. However, when parameter values are far from convergence, we are not justified in neglecting ${}^n\mathbf{e}$ compared to ${}^{n-1}\underline{\lambda} - {}^n\underline{\lambda}$ in Eq. (69) and the set of the rotational parameters of the n th iteration cannot be used with confidence for characterizing the global convergence of a sequence of iterations. This is especially true when parameter values are far from convergence and a step sign and size control algorithm is applied (see Section V). When a nonlinear iterative procedure is carried out, the basis of orbitals is changed at each step of the iterative procedure. The set of ${}^{n-1}\underline{\lambda}$ parameter values consequently are zero and $\|{}^n\lambda\|$ then directly may be used as an estimate of the magnitude of the error vector at iteration point $n - 1$. Thus the initial error $\|{}^0\mathbf{e}\|$ is approximately given by $\|{}^1\underline{\lambda}\|$, and so on.

C. Implementation of the Nonlinear Newton–Raphson Approach

1. One-Step Approach

The nonlinear Newton–Raphson equation is defined in Eq. (60). A computationally more tractable form is obtained by using the right-hand side of Eqs. (51) and (52), respectively, rather than $\underline{\lambda}$ and \underline{T}^+ . Equation (60) then becomes

$$\begin{pmatrix} \kappa \\ \mathbf{s} \end{pmatrix} = -(\mathbf{A} - \mathbf{B})^{-1} \begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix} = -\mathbf{G}^{-1} \mathbf{F} \quad (72)$$

where

$$\mathbf{W} = \langle 0 | [\mathbf{Q}^+, H] | 0 \rangle \quad (73)$$

$$\mathbf{V} = \langle 0 | [\mathbf{R}^+, H] | 0 \rangle \quad (74)$$

and

$$(\mathbf{A} - \mathbf{B}) = \frac{1}{2} \mathbf{G}$$

$$\begin{pmatrix} \mathbf{W} \\ \mathbf{V} \end{pmatrix} = \frac{1}{2} \mathbf{F}$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^{OO} & \mathbf{A}^{OC} \\ \mathbf{A}^{CO} & \mathbf{A}^{CC} \end{pmatrix} = - \begin{pmatrix} \langle 0 | [\mathbf{Q}, \mathbf{Q}^+, H] | 0 \rangle & \langle 0 | [\mathbf{R}, [\mathbf{Q}^+, H]] | 0 \rangle \\ \langle 0 | [\mathbf{R}^+, [\mathbf{Q}, H]] | 0 \rangle & \langle 0 | [\mathbf{R}, \mathbf{R}^+, H] | 0 \rangle \end{pmatrix} \quad (75)$$

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}^{OO} & \mathbf{B}^{OC} \\ \mathbf{B}^{CO} & \mathbf{B}^{CC} \end{pmatrix} = - \begin{pmatrix} \langle 0 | [\mathbf{Q}^+, \mathbf{Q}^+, H] | 0 \rangle & \langle 0 | [\mathbf{R}^+, [\mathbf{Q}^+, H]] | 0 \rangle \\ \langle 0 | [\mathbf{R}^+, [\mathbf{Q}^+, H]] | 0 \rangle & \langle 0 | [\mathbf{R}^+, \mathbf{R}^+, H] | 0 \rangle \end{pmatrix} \quad (76)$$

Once $\underline{\kappa}$ and \mathbf{S} are obtained, a new orbital basis [see Eq. (8)] and state expansion coefficients [see Eq. (4)] are defined. The new orbitals and state expansion coefficients are used to construct $\mathbf{A} - \mathbf{B}$, \mathbf{W} , and \mathbf{V} . This requires a two-electron transformation (see Section IX.A). Subsequently a new $\underline{\kappa}$ and a new \mathbf{S} are again obtained from Eq. (72), and so on (as above). This sequence is continued until all elements in \mathbf{W} and \mathbf{V} (or $\underline{\kappa}$ and \mathbf{S}) are smaller than a certain tolerance. The expressions $\mathbf{W} = \mathbf{0}$ and $\mathbf{V} = \mathbf{0}$ are sometimes referred to together as the generalized Brillouin theorem (GBT). When Eq. (72) is used as an iterative function, the calculation is often referred to as the one-step nonlinear Newton-Raphson approach.^{11,12}

2. Two-Step Approach

Equation (72) may be used as an iterative function in a slightly different way if a configuration interaction (CI) calculation is carried out in the MCSCF configuration space before Eq. (72) is applied. When the configuration interaction calculation is carried out

$$\langle m | H | l \rangle = E_l \delta_{ml} \quad (77)$$

and

$$V_n = \langle 0 | [| n \rangle \langle 0 |, H] | 0 \rangle = 0 \quad (78)$$

$$A_{nm}^{CC} = -\langle 0 | [R_n, R_m^+, H] | 0 \rangle = \delta_{nm} (E_m - E_0) \quad (79)$$

$$B_{nm}^{CC} = 0$$

Equation (72) may therefore be partitioned,^{11,41} using Eq. (78), as

$$\underline{\kappa} = -\left(\mathbf{A}^{OO} - \mathbf{B}^{OO} - (\mathbf{A}^{OC} - \mathbf{B}^{OC})(\mathbf{A}^{CC})^{-1}(\mathbf{A}^{CO} - \mathbf{B}^{CO}) \right)^{-1} \mathbf{W} \quad (80)$$

A similar iterative sequence to the one-step procedure may be used to obtain all elements of \mathbf{W} (or $\underline{\kappa}$) smaller than a certain tolerance. However, an additional CI is required in each iteration subsequent to the transformation and prior to construction of $\mathbf{A} - \mathbf{B}$ and \mathbf{W} . When Eq. (80) is used as an iterative function, the calculation is often referred to as the two-step nonlinear Newton-Raphson approach.^{10,11,13}

If energy difference $E_m - E_0$ occurring in diagonal elements of \mathbf{A}^{CC} [see Eq. (79)] is relative small, the last term in the inverse matrix may be very large and the two-step procedure is then less tractable for use as an iterative function than the one-step procedure. In a later section we will explicitly compare the convergence characteristics of the one- and two-step nonlinear Newton-Raphson calculations. We will also present an example of where it

is dangerous to use the two-step procedure. The nonlinear Newton–Raphson calculations we report will generally, if not otherwise specified, be one-step calculations.

The terms $\mathbf{A}^{\text{CO}} - \mathbf{B}^{\text{CO}}$ which couple the configuration and orbital variation have been neglected in some calculations. When the coupling terms have been neglected, convergence problems have often been encountered.^{11,13,15,36} In some cases an iterative scheme which neglects the coupling cannot converge even if the initial guess of orbitals and states are infinitesimally close to the desired stationary point.^{13,15}

When converging to the $(n - 1)$ st excited state of a given symmetry, the Hessian matrix for that state should usually have $n - 1$ negative eigenvalues. Since the Hessian matrix appears directly in the one-step Newton–Raphson iterative function, this fact may be used to direct the calculation to the desired stationary point (see Section V). When the two-step Newton–Raphson procedure is applied only eigenvalues of the reduced Hessian,

$$\mathbf{A}^{\text{OO}} - \mathbf{B}^{\text{OO}}(\mathbf{A}^{\text{OC}} - \mathbf{B}^{\text{OC}})(\mathbf{A}^{\text{CC}})^{-1}(\mathbf{A}^{\text{CO}} - \mathbf{B}^{\text{CO}})$$

and of the configuration block of the Hessian, \mathbf{A}^{CC} , are directly available. In Appendix D we show that if the Hessian matrix is required to have n negative eigenvalues and the configurational block of the Hessian matrix has m negative eigenvalues, then the reduced Hessian matrix needs to have $n - m$ negative eigenvalues. The step size and sign control algorithm of Section V may therefore also be used directly on Eq. (80) to ensure convergence to the desired solution. However, it may be very difficult to ensure convergence to the desired state if an energy difference $E_m - E_0$ of Eq. (79) is very small.

D. Matrix Elements in a Nonlinear Newton–Raphson Approach

We now consider the specific case where \hat{H} is the electronic Hamiltonian of an isolated atom or molecule, that is,

$$\hat{H} = \sum_{ij} h_{ij} a_i^+ a_j + \frac{1}{2} \sum_{ijkl} (ik|jl) a_i^+ a_j^+ a_l a_k$$

where

$$h_{ij} = \langle \phi_i | \hat{h} | \phi_j \rangle \quad (\hat{h} \text{ is the sum of one-body operators})$$

and

$$(ik|jl) = \langle \phi_i(1) \phi_j(2) | \frac{1}{r_{12}} | \phi_k(1) \phi_l(2) \rangle$$

where the ϕ 's are spin orbitals.

Let us now consider explicitly the evaluation of some of the matrix elements appearing in the one-step Newton–Raphson approach.^{11,23,26–27}

$$A_{nm}^{CC} = -\langle 0 | [R_n, R_m^+, H] | 0 \rangle = \langle n | H | m \rangle - \delta_{nm} \langle 0 | H | 0 \rangle, \quad B_{nm}^{CC} = 0 \quad (81)$$

$$V_n = \langle 0 | [|n\rangle \langle 0|, H] | 0 \rangle = -\langle 0 | H | n \rangle \quad (82)$$

Thus, \mathbf{A}^{CC} and \mathbf{V} contain all matrix elements contained in the MCSCF CI calculation. When the iterative procedure has converged, all elements of \mathbf{V} are zero and the interaction between the MCSCF reference state $|0\rangle$ and the residual states are eliminated. The diagonal and off-diagonal matrix elements of the Hamiltonian in the residual space $\langle |k\rangle$ may, however, all be nonvanishing.

Except for \mathbf{A}^{CC} and \mathbf{V} , the form of the matrix elements in the one- and two-step Newton–Raphson approaches are the same. The matrix elements of \mathbf{W} , $\mathbf{A}^{OO} - \mathbf{B}^{OO}$ may be derived from Eqs. (83) and (84) by index substitution

$$\begin{aligned} \langle 0 | [a_{i\sigma}^+ a_{u\sigma}, H] | 0 \rangle &= h_{up} \langle 0 | a_{i\sigma}^+ a_{p\sigma} | 0 \rangle - h_{pt} \langle 0 | a_{p\sigma}^+ a_{u\sigma} | 0 \rangle \\ &\quad - (pr|qt)\rho_{qpru}^{(2)} + (ur|qs)\rho_{iqsr}^{(2)} \end{aligned} \quad (83)$$

$$\begin{aligned} \langle 0 | [a_{i\sigma}^+ a_{k\sigma}, [H, a_{i\sigma'}^+ a_{u\sigma'}]] | 0 \rangle &= h_{kt} \langle 0 | a_{i\sigma}^+ a_{u\sigma} | 0 \rangle + h_{ul} \langle 0 | a_{i\sigma}^+ a_{k\sigma} | 0 \rangle \\ &\quad - \delta_{kt} h_{up} \langle 0 | a_{i\sigma}^+ a_{p\sigma} | 0 \rangle - \delta_{lu} h_{pt} \langle 0 | a_{p\sigma}^+ a_{k\sigma} | 0 \rangle \\ &\quad - \delta_{lu} (pr|qt)\rho_{pqkr}^{(2)} \\ &\quad - \delta_{kt} (ur|qs)\rho_{iqsr}^{(2)} + (pl|qt)\rho_{pquk}^{(2)} \\ &\quad - (ur|ks)\rho_{ilsr}^{(2)} + (kr|pt)\rho_{lpur}^{(2)} + (kt|pr)\rho_{plur}^{(2)} \\ &\quad + (ul|qs)\rho_{iqsk}^{(2)} + (us|ql)\rho_{qtsk}^{(2)} \end{aligned} \quad (84)$$

where

$$\rho_{ijkl}^{(2)} = \langle 0 | a_{i\sigma}^+ a_{j\sigma'}^+ a_{k\sigma'} a_{l\sigma} | 0 \rangle \quad (85)$$

In these equations we have explicitly introduced spin (σ and σ' run over the electron spin indices α and β). Note again that the Einstein summation convention is used. The excitation operators in Eqs. (83) and (84) have been coupled to singlet spin symmetry since they appear in the operator $\hat{\kappa}$ which must preserve the symmetry of the reference state when forming $\exp(i\hat{\kappa})|0\rangle$. The matrix elements of \mathbf{W} , and $\mathbf{A}^{OO} - \mathbf{B}^{OO}$ can be expressed in terms of one- and two-electron integrals and the one- and two-electron density matrices.

The elements of $\mathbf{A}^{\text{OC}} - \mathbf{B}^{\text{OC}}$ may be reduced as follows:

$$\langle 0 | [| 0 \rangle \langle n |, [a_{i\sigma}^+ a_{u\sigma}, H]] | 0 \rangle = \langle n | [a_{i\sigma}^+ a_{u\sigma}, H] | 0 \rangle \quad (86)$$

Then an explicit formula for Eq. (86) may be obtained from Eq. (83) by replacing one- and two-electron density matrices with the corresponding transition density matrix elements.

All indices in the two-electron density matrix must refer to occupied or partly occupied orbitals to give a nonvanishing two-electron density matrix element. Since in Eq. (83) three indices are common for the two-electron integral multiplying the two-electron density matrix element, only one index in the two-electron integrals in Eq. (83) can refer to a completely unoccupied orbital. Also, in Eq. (84) there are at least two indices common for a certain two-electron integral and the multiplying density matrix element. Hence, for the two-electron integrals in Eq. (84) only at most two indices can refer to completely unoccupied orbitals.

We note that we have not explicitly written any additional symmetry indices on the formulas. This is, of course, straightforward to do; however, a rather cumbersome nomenclature results. In all calculations we have reported, spatial and spin symmetry has, of course, been explicitly incorporated.¹¹⁻²²

Further analysis of the gradient and Hessian matrix elements is given in Appendix C.

IV. CHARACTERIZATION OF AN MCSCF STATE

Before we proceed further we will analyze in more detail the requirements that may be imposed on an MCSCF state to assure that the MCSCF state is a proper representation of the exact n th state. So far the MCSCF state is only required to be variationally correct, that is, to satisfy the GBT (i.e., be at a stationary point on the energy hypersurface). This is, of course, an insufficient condition to ensure that the MCSCF state also is a proper representation of the n th exact state, and additional conditions should be fulfilled.

An MCSCF state optimization is performed both with respect to linear (configuration expansion coefficients) and nonlinear (orbital expansion coefficients) variational parameters. In CI calculations the optimization is only performed with respect to the linear configuration expansion parameter and the Hyleraas–Undheim–McDonald theorem⁴² (the n th root in energy of a certain symmetry of the CI secular problem is an upper bound to the n th exact state of a certain symmetry) is often used as a necessary and sufficient condition for identifying the states. The Hyleraas–Undheim–McDonald theorem is only valid when variations are restricted to linear

variation parameters. Hence the Hyleraas–Undheim–McDonald theorem can only be applied within the MCSCF configuration space. If the Hyleraas–Undheim–McDonald theorem is applied within the MCSCF configuration space we would require the CI within the MCSCF configuration space to have $n - 1$ roots lower in energy than the root for the desired n th state. Since the MCSCF optimization also is performed with respect to the nonlinear orbital expansion parameters, it may be more appropriate to use criteria for characterizing an MCSCF state that are applicable to both the linear and nonlinear part of the variational space.

The MCSCF state is an approximation to the n th root in energy of a certain symmetry of a full CI calculation. The Hessian matrix corresponding to the n th root of a certain symmetry of the full CI solution has $n - 1$ negative Hessian eigenvalues. It is reasonable for the MCSCF state to have $n - 1$ negative Hessian eigenvalues, thereby simulating the full CI solution.

When characterizing the MCSCF state it is also appropriate to examine the linear response of the MCSCF state to an external one-electron perturbation. The linear response [or multiconfigurational time-dependent Hartree–Fock (MCTDHF)] calculation determines a set of excitation energies.^{43–53} When describing the n th state in energy of a certain symmetry it is relevant to require that the response calculation for that symmetry gives $n - 1$ negative excitation energies. In addition, the multiconfigurational stability conditions should be fulfilled.^{16, 43–44, 52–54}

In summary, the MCSCF state which represents the n th state of a certain symmetry should ideally have the following characteristics:

1. The full (unpartitioned) MCSCF Hessian should have $n - 1$ negative eigenvalues.
2. The MCTDHF calculations using the converged MCSCF orbitals and state expansion coefficients should be stable and have $n - 1$ negative excitation energies to states of the same symmetry as the MCSCF state.
3. The CI with the MCSCF configuration state functions should have $n - 1$ roots lower in energy.

For the lowest state of a given symmetry all these conditions must be satisfied at convergence. For an excited state of a given symmetry all of the above criteria should also be fulfilled at convergence. However, with a limited MCSCF configuration state space and with finite basis sets some (or, in rare cases, all) of these will not be met. This is particularly true when there are two (or more) states of the same symmetry that are very close in energy. Since it has not been traditional to examine these criteria either when converging or at convergence, we suspect that many of the previously reported MCSCF calculations are erroneous.

A further consideration is, of course, to examine the magnitude of state expansion coefficients. This is particularly relevant when some or all of the foregoing three criteria are not fulfilled. However, assignments based primarily on these magnitudes should be made only with extreme caution.

Another important reason for discussing at this point the characteristics of an MCSCF state is that these characteristics can be used to ensure that an MCSCF calculation is proceeding to the region of the correct stationary point. For example, when far from convergence (the global convergence problem) the Newton–Raphson procedure may give a few large step length amplitudes. These large amplitudes may take a calculation to an undesired place on energy hypersurface. The Hessian may have too many ($> n - 1$) or too few ($< n - 1$) negative eigenvalues. These may also cause a calculation to proceed to an undesired stationary point. Hence some sort of step size and sign control algorithm (or higher order procedure) needs to be used when far from convergence. Step size and sign control is discussed in the next section. In all cases the algorithm should be *firmly* based on principles which will lead us to the correct stationary point. For example, we have found that the most reliable and efficient constraint procedures should be designed to force the full Hessian to have the proper number, $n - 1$, of negative eigenvalues in every iteration.

A. n th State of the CI Using the MCSCF Configuration State Functions (CSFs)

It is important that the n th state of a certain symmetry has $n - 1$ negative eigenvalues of the converged full MCSCF Hessian. This condition does not, however, imply that the state needs to be the n th root of a certain symmetry of the CI using the MCSCF configuration state functions. The CI matrix [actually $2\langle i|H|j\rangle - 2\delta_{ij}\langle 0|H|0\rangle$; see Eq. (81)] is only one subblock (\mathbf{G}^{CC}) of the full Hessian. Even though the full Hessian has $n - 1$ negative eigenvalues at convergence, the state we are converging to may not be the n th state of the CI. Such a solution where the state is not the n th root of the CI (root flipping) will be variationally correct (and may, in fact, even be a very good approximation) but, of course, may not be an upper bound to the n th state.⁴²

Root flipping occurs when the n th state in the MCSCF CI becomes the $(n - 1)$ st [or $(n - 2)$ nd or $(n - 3)$ rd...] state as convergence progresses. This occurs because the orbitals and state expansion coefficients are being optimized for the original n th state and not for the other states of the same symmetry. When root flipping occurs the Hyleraas–Undheim–McDonald theorem⁴² may no longer be used to characterize the MCSCF state and the MCSCF state in general cannot assuredly be taken as an upper bound to the n th exact eigenstate. We have demonstrated through the correct application of the ideas on redundant variables and linear dependency that for many

cases root flipping is an artifact of the choice of MCSCF configurations.¹⁷ This is because there is often a freedom of choice between state and orbital variables for MCSCF optimization. Hence, a converged calculation with a smaller MCSCF CI space may have root flipping, whereas a calculation with a larger MCSCF CI space may have no root flipping. However, both converged MCSCF states will have the same total MCSCF energy. Hence, even though root flipping has occurred, the energy may still be an upper bound since certainly the energy of the calculation with no root flipping is an upper bound. We conclude that, in general, criteria 1 and 2 (above) are probably more significant than criterion 3 and that the current rather strong emphasis^{29–30, 32, 35–36} on having MCSCF calculations with no root flipping is probably somewhat misplaced.

B. Eigenvalues of the Hessian

Eigenvalues of the converged Hessian matrix of Eq. (14) are positive for the ground state. Similarly, a condition on the first, second,...excited state is that the corresponding Hessian matrix has one, two,...negative eigenvalues. In the limit of the full CI, all orbital optimization variables are redundant and hence the Hessian will just be \mathbf{G}^{CC} . The matrix elements of $\frac{1}{2}\mathbf{G}^{CC}$ are

$$\langle n|H|m\rangle - \delta_{mn}\langle 0|H|0\rangle \quad (81)$$

where n and m are CI states. In this case the eigenvalues of \mathbf{G}^{CC} and the full Hessian are the same. Obviously, in the limit of the full CI with an adequate basis set the Hessian will also have the correct number of negative eigenvalues.

With a smaller choice of configurations, the eigenvalues of \mathbf{G}^{CC} and \mathbf{G} are no longer the same. In fact, at convergence \mathbf{G}^{CC} may not have the appropriate number ($n - 1$) negative eigenvalues for the n th state while \mathbf{G} may have $n - 1$ negative eigenvalues. Thus, when root flipping occurs, our MCSCF reference state energy may not be an upper bound to the energy of this state. (However, see the discussion following and in Section V and Appendix A concerning reparametrization of an MCSCF problem). With a reasonable choice of configurations and provided the full Hessian has the proper number of negative eigenvalues, the converged MCSCF state may be a good representation of an excited state even when root flipping has occurred. The fact that an energy is an upper bound to the exact energy is, of course, no guarantee that the corresponding state is a good representation of an exact solution to the Schrödinger equation. Techniques which attempt to force \mathbf{G}^{CC} to have $n - 1$ negative eigenvalues are variationally constrained and may lead to undesired results.

In the two-step procedure, the eigenvalues of the MCSCF CI problem and the eigenvalues of the reduced Hessian matrix are separately determined. The eigenvalue problem for the full Hessian

$$\begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \mathbf{U}_O \\ \mathbf{U}_C \end{pmatrix} = \epsilon \begin{pmatrix} \mathbf{U}_O \\ \mathbf{U}_C \end{pmatrix} \quad (87)$$

transforms, using partitioning theory⁴¹ to an equation of reduced dimension

$$[\mathbf{G}^{OO} + \mathbf{G}^{OC}(\epsilon \mathbf{1} - \mathbf{G}^{CC})^{-1} \mathbf{G}^{CO}] \mathbf{U}_O = \epsilon \mathbf{U}_O \quad (88)$$

The resolvent can then be expanded

$$\begin{aligned} & [\mathbf{G}^{OO} - \mathbf{G}^{OC}(\mathbf{G}^{CC} - \epsilon \mathbf{1})^{-1} \mathbf{G}^{CO}] \mathbf{U}_O \\ &= (\mathbf{G}^{OO} - \mathbf{G}^{OC}(\mathbf{G}^{CC})^{-1} \mathbf{G}^{CO} - \epsilon \mathbf{G}^{OC}(\mathbf{G}^{CC})^{-1} \mathbf{1}(\mathbf{G}^{CC})^{-1} \mathbf{G}^{OC} - \dots) \mathbf{U}_O \end{aligned} \quad (89)$$

Hence, the eigenvalues and eigenvectors of the reduced Hessian are not the eigenvalues and eigenvectors of the full Hessian. They will be similar to the eigenvalues and vectors of the full Hessian only if the third, fourth, etc. terms in Eq. (89) are small and may be ignored.

The eigenvalues of the reduced Hessian are related to the eigenvalues of the Hessian matrix [Eq. (87)] through the nonorthogonal transformation

$$\begin{aligned} & \begin{pmatrix} 1 & -\mathbf{G}^{OC} \mathbf{G}^{CC^{-1}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\mathbf{G}^{CC^{-1}} \mathbf{G}^{CO} & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{G}^{OO} - \mathbf{G}^{OC} \mathbf{G}^{CC^{-1}} \mathbf{G}^{CO} & 0 \\ 0 & \mathbf{G}^{CC} \end{pmatrix} \end{aligned} \quad (90)$$

This transformation may be used to relate the number of negative eigenvalues of the Hessian matrix to the number of negative eigenvalues of the reduced Hessian matrix. If we define an arbitrary vector \mathbf{Y} as

$$\mathbf{Y} = \begin{pmatrix} 1 & 0 \\ -\mathbf{G}^{CC^{-1}} \mathbf{G}^{CO} & 1 \end{pmatrix} \mathbf{X} \quad (91)$$

for any vector \mathbf{X} satisfying $|\mathbf{X}| > 0$, it is easily seen that if the Hessian matrix is positive definite

$$\mathbf{Y}^T \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \mathbf{Y} > 0 \quad (92)$$

then the eigenvalues of

$$\begin{pmatrix} \mathbf{G}^{OO} - \mathbf{G}^{OC}\mathbf{G}^{CC^{-1}}\mathbf{G}^{CO} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{CC} \end{pmatrix} \quad (93)$$

have to be positive. Similarly, if the eigenvalues of Eq. (93) are positive, the eigenvalues of the Hessian matrix are positive. If the Hessian matrix has one negative eigenvalue, Eq. (93) contains at least one negative eigenvalue, but from the preceding analysis it is difficult to determine the number of negative eigenvalues in Eq. (93).

A more thorough analysis based on a multidimensional partitioning⁴¹ of the full Hessian (see Appendix D) shows that the number of negative eigenvalues of Eq. (93) is the same as the number of negative eigenvalues of Eq. (87). Hence, if there is root flipping, negative eigenvalues will appear in the upper-left-hand block (reduced Hessian matrix) of Eq. (93).

A final point is that since the CI coefficient problem is linear and the orbital optimization problem is nonlinear it is expected that there may be several stationary points with the same number of negative eigenvalues of the full Hessian (e.g., the BO calculations reported in Ref. 17). In such cases, a more detailed examination of the CI coefficients and criteria 1–3, above, is usually required.

C. Excitation Energies in the Multiconfigurational Time-Dependent Hartree–Fock Approximation

We will now describe how excitation energies (positive and negative) may be determined in a MCTDHF [also known as the multiconfigurational random phase approximation (MCRPA)] calculation.^{43–50}

In the MCTDHF approximation the linear response of an MCSCF state to a frequency-dependent one-electron perturbation is examined. The MCTDHF approximation has been derived previously and excitation energies are determined as eigenvalues of the generalized eigenvalue problem

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{S} & \Delta \\ -\Delta & -\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix} \quad (94)$$

where the \mathbf{A} and \mathbf{B} matrices are defined in Eqs. (75) and (76) and

$$\mathbf{S} = \langle 0 | [\mathbf{T}, \mathbf{T}^+] | 0 \rangle \quad (95)$$

$$\Delta = \langle 0 | [\mathbf{T}, \mathbf{T}] | 0 \rangle \quad (96)$$

Here $\begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix}$ are the eigenvectors and ω is the excitation energy, $\omega_i = E_i - E_0$.

The solution of Eq. (94) may be determined through performing a series of transformations involving matrices of only half the dimension of Eq. (94).^{16, 43–44, 52} To achieve this reduction, we write Eq. (94)

$$\mathbf{A}\mathbf{Z} + \mathbf{B}\mathbf{Y} = \omega\mathbf{S}\mathbf{Z} + \omega\Delta\mathbf{Y} \quad (97)$$

$$\mathbf{B}\mathbf{Z} + \mathbf{A}\mathbf{Y} = -\omega\Delta\mathbf{Z} - \omega\mathbf{S}\mathbf{Y} \quad (98)$$

Successively adding and subtracting the above two equations gives

$$(\mathbf{A} + \mathbf{B})(\mathbf{Z} + \mathbf{Y}) = \omega(\mathbf{S} - \Delta)(\mathbf{Z} - \mathbf{Y}) \quad (99)$$

$$(\mathbf{A} - \mathbf{B})(\mathbf{Z} - \mathbf{Y}) = \omega(\mathbf{S} + \Delta)(\mathbf{Z} + \mathbf{Y}) \quad (100)$$

Equation (99) may then be rearranged

$$\mathbf{Z} + \mathbf{Y} = \omega(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{S} - \Delta)(\mathbf{Z} - \mathbf{Y}) \quad (101)$$

and inserted into Eq. (100) to give the nonhermitian eigenvalue problem of half the dimensions of Eq. (94):

$$(\mathbf{S} - \Delta)^{-1}(\mathbf{A} + \mathbf{B})(\mathbf{S} + \Delta)^{-1}(\mathbf{A} - \mathbf{B})(\mathbf{Z} - \mathbf{Y}) = \omega^2(\mathbf{Z} - \mathbf{Y}) \quad (102)$$

If $\mathbf{S} - \Delta$ or $\mathbf{S} + \Delta$ are singular or near singular, it may be advantageous to rearrange Eq. (102) to be

$$\frac{1}{\omega^2}(\mathbf{Z} - \mathbf{Y}) = (\mathbf{A} - \mathbf{B})^{-1}(\mathbf{S} + \Delta)(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{S} - \Delta)(\mathbf{Z} - \mathbf{Y}) \quad (103)$$

which gives $1/\omega^2$ as eigenvalues. If $\mathbf{A} - \mathbf{B}$ is positive definite, Eq. (102) may be arranged to a hermitian eigenvalue problem

$$\begin{aligned} &(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{S} - \Delta)^{-1}(\mathbf{A} + \mathbf{B})(\mathbf{S} + \Delta)^{-1}(\mathbf{A} - \mathbf{B})^{1/2} \\ &(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{Z} - \mathbf{Y}) = \omega^2(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{Z} - \mathbf{Y}) \end{aligned} \quad (104)$$

since Δ is antisymmetric.

Excitation energies $\{\omega_i\}$ and the corresponding eigenvectors $\begin{pmatrix} {}^i\mathbf{Z} \\ {}^i\mathbf{Y} \end{pmatrix}$ are obtained from Eqs. (94), (102), (103), or (104). Another set of solutions is $\{-\omega_i\}$

and $\begin{pmatrix} i\mathbf{Y} \\ i\mathbf{Z} \end{pmatrix}$. A comparison with a spectral representation of the polarization propagator shows that the excitation energy is ω if the eigenvector $\begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix}$ is normalized to 1:

$$(\mathbf{Z}^T \quad \mathbf{Y}^T) \begin{pmatrix} \mathbf{S} & \Delta \\ -\Delta & -\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix} = 1 \quad (105)$$

while the excitation energy is $-\omega$ if the eigenvector $\begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix}$ is normalized to -1 :

$$(\mathbf{Z}^T \quad \mathbf{Y}^T) \begin{pmatrix} \mathbf{S} & \Delta \\ -\Delta & -\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{Z} \\ \mathbf{Y} \end{pmatrix} = -1 \quad (106)$$

That is, if we want the MCSCF state $|0\rangle$ to represent the third lowest state of a given symmetry, we require the MCTDHF calculation to give *two* negative excitation energies, that is, two negative ω_i 's with corresponding positive norm eigenvectors.

D. Stability Condition for a Multiconfigurational Hartree-Fock State

When solving the nonhermitian eigenvalue problem in Eq. (102) negative (ω^2 negative) or complex roots may be encountered. If negative or complex roots occur as eigenvalues of Eq. (102) the MCTDHF approximation is said to have an instability.¹⁶ We analyze in the following the conditions under which such instabilities may occur. If $\mathbf{A} - \mathbf{B}$ is positive definite, instabilities are not encountered if the matrix $(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{S} - \Delta)^{-1}(\mathbf{A} + \mathbf{B})(\mathbf{S} + \Delta)^{-1}(\mathbf{A} - \mathbf{B})^{1/2}$ is positive definite [see Eq. (104)], that is,

$$\mathbf{X}(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{S} - \Delta)^{-1}(\mathbf{A} + \mathbf{B})(\mathbf{S} + \Delta)^{-1}(\mathbf{A} - \mathbf{B})^{1/2}\mathbf{X}^T \geq 0 \quad (107)$$

Defining a vector

$$\mathbf{U} = \mathbf{X}(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{S} - \Delta)^{-1} \quad (108)$$

we may write Eq. (107) as

$$\mathbf{U}(\mathbf{A} + \mathbf{B})\mathbf{U}^T \geq 0 \quad (109)$$

which tells us that $\mathbf{A} + \mathbf{B}$ has to be positive definite to ensure that Eq. (107) is fulfilled. Thus if $\mathbf{A} - \mathbf{B}$ is positive definite and $\mathbf{A} + \mathbf{B}$ is not, an MCTDHF

instability is encountered. When determining the solution to Eq. (94) we might alternatively have determined $\mathbf{Z} - \mathbf{Y}$ from Eq. (100) and then inserted $\mathbf{Z} - \mathbf{Y}$ into Eq. (99). A derivation similar to the one described shows that, if $\mathbf{A} + \mathbf{B}$ is positive definite, then an MCTDHF instability is encountered if $\mathbf{A} - \mathbf{B}$ is not positive definite. If both $\mathbf{A} \pm \mathbf{B}$ are positive definite, instabilities are not encountered. If both $\mathbf{A} \pm \mathbf{B}$ are not positive definite, an explicit solution to Eq. (102) has to be determined before it is clear whether an instability occurs. If both $\mathbf{A} \pm \mathbf{B}$ are not positive definite and instabilities are not encountered, negative excitation energies are obtained in the MCTDHF calculation. We recall that the curvature of the energy hypersurface at a point representing an MCSCF state refers to twice the $\mathbf{A} - \mathbf{B}$ matrix which occurs in the MCTDHF stability condition. Further, if we previously had examined the variations in the energy resulting from purely imaginary orbital variations, the second derivative of the total energy would involve the $\mathbf{A} + \mathbf{B}$ matrix. Hence, the condition that $\mathbf{A} \pm \mathbf{B}$ is positive definite is satisfied if the MCSCF state represents a local minimum both with respect to real and imaginary orbital variations. This point has previously been mentioned by Dalgaard.⁴⁴ When $\mathbf{A} \pm \mathbf{B}$ are not positive definite, the MCSCF state represents a saddle point on the energy hypersurface.

The aforementioned stability condition is not always trivially fulfilled for a state that satisfies the GBT. We describe below one example where the stability condition is not fulfilled for a state which satisfies the GBT. Include in the configuration space one main configuration and all singly excited configurations relative to the main configuration. The orbital optimization excitation operators connecting orbitals that both are occupied or both are unoccupied in the main configuration are redundant operators for that particular choice of configurations. The Hartree–Fock (HF) state for the main configuration represents a stationary point in the MCSCF calculation just described. The HF state, however, is surely not the state that represents the ground state in this MCSCF calculation, nor does this state represent an excited state. We have considered cases which cover the above-described choice of configurations and found that the HF ground state was unstable in the MCSCF calculation. We found that $\mathbf{A} - \mathbf{B}$ was not positive definite while $\mathbf{A} + \mathbf{B}$ was positive definite and did thus not have the problem of giving the HF ground state a physical interpretation in the MCSCF calculation. In the aforementioned calculations physical intuition might have told us that the stationary point which represented the HF ground state did not show the right characteristics to represent either an MCSCF ground or an excited state. However, in more complicated cases, it may be very useful to use the stability condition as a constraint the MCSCF state should satisfy to represent the exact state, that is, in addition to having the proper number of negative excitation energies instabilities are not encountered.

V. CALCULATIONS WITH THE NEWTON-RAPHSON APPROACH

We have previously discussed the Newton-Raphson approach for MCSCF optimization (see Section III). This approach demonstrates second order convergence. The order concept is only defined when close to convergence^{9, 20, 22}; when farther from convergence, the Newton-Raphson approach may give step length amplitudes which are very large.^{13, 15, 17} The full Hessian G may also have an incorrect structure, that is, too few or too many negative eigenvalues,^{13, 15, 17} leading to a Newton-Raphson step in an undesired direction. Large step length amplitudes or step length amplitudes in the wrong direction may introduce large fluctuations in the step length amplitudes of the subsequent iterations. A divergent sequence of iterations may be observed if the large step length amplitudes or step length amplitudes of the wrong sign are used uncritically.

When the Newton-Raphson approach does not straightforwardly converge, in numerical analysis it has been advocated to use restricted step methods for obtaining global convergence.⁹ In a restricted step method for the lowest state of a certain symmetry a minimum for the second-order approximation to the Taylor series is determined subject to the constraint that the norm of $\underline{\lambda}$ has to be less than h , where h is a constant defining the trust region of the second order Taylor series expansion

$$\min_{\underline{\lambda}} q(\underline{\lambda}) = E_0 + F_i \lambda_i + \frac{1}{2} G_{ij} \lambda_i \lambda_j, \quad \|\underline{\lambda}\| \leq h \quad (110)$$

The value of h is chosen as large as possible subject to the condition that a certain measure of agreement exists between $q(\underline{\lambda})$ and $E(\underline{\lambda})$ in Eq. (47). Fletcher⁹ describes an algorithm for quantifying the agreement between $q(\underline{\lambda})$ and $E(\underline{\lambda})$. In iteration k a set of ${}^k\underline{\lambda}$ parameters is determined and

$$\Delta q^{(k)} = q({}^k\underline{\lambda}) - q(\mathbf{0}) \quad (111)$$

$$\Delta E^{(k)} = E({}^k\underline{\lambda}) - E(\mathbf{0}) \quad (112)$$

are evaluated.

The ratio

$$r^{(k)} = \frac{\Delta E^{(k)}}{\Delta q^{(k)}} \quad (113)$$

then measures the accuracy to which $q({}^k\underline{\lambda})$ approximates $E({}^k\underline{\lambda})$ in the sense that the closer $r^{(k)}$ is to unity, the better is the agreement. The algorithm of

Fletcher changes h adaptively, attempting to maintain a certain degree of agreement between $q(\lambda^k)$ and $E(\lambda^k)$ as measured by $r^{(k)}$ while keeping h as large as possible. The details of the algorithm of Fletcher may be found in Ref. 9. Fletcher proves that a restricted step algorithm is globally converging for the lowest state of a certain symmetry.

An essential part of the Fletcher restricted step method is that large step sizes never are allowed. A region of the energy hypersurface is first examined and never left before assurance is obtained that the region contains no stationary points. The step length amplitudes are then in a sense determined such as to bring the calculation to the part of the restricted region that shows the most promise for determining a stationary point. The stationary point for $q(\lambda)$ when no restrictions are imposed on the step length amplitude is of course the set of Newton-Raphson parameters.

In our implementation of restricted step methods we have chosen as an alternative to solve Eq. (110) to modify the Newton-Raphson step length amplitudes, ensuring that the modified Newton-Raphson step never becomes large and always is in the right direction.^{13, 15, 17} A comparison of the global convergence properties of the Newton-Raphson step size and sign control algorithm¹⁵ and Fletcher's restricted step method will be reported very soon.⁵⁵

For now we will show how we have implemented the step size and sign control algorithm in MCSCF Newton-Raphson calculations. We then report a series of nonlinear step size and sign controlled Newton-Raphson calculations to illustrate the local and global convergence characteristics of an MCSCF Newton-Raphson calculation. Calculations indicate that the MCSCF Newton-Raphson approach reliably can be used to get an MCSCF calculation to converge. However, the calculations also indicate that alternative iterative procedures may be derived which efficiently may be used to obtain both global²⁰ and local²⁰⁻²² convergence of an MCSCF calculation. These approaches will be discussed in subsequent sections.

A. Step Size and Sign Control in the Newton-Raphson Approach

The MCSCF Newton-Raphson approach is used to determine state wavefunctions for both the lowest and for excited states of a given symmetry. The algorithm we apply is constructed such that deviations from the Newton-Raphson step occur when (1) undesired negative eigenvalues show up in the Hessian matrix or (2) the Newton-Raphson approach gives very large step length amplitudes. In developing this algorithm as well as previously developed algorithms¹⁵ we have been particularly concerned with designing techniques which properly account for the characteristics of an MCSCF state as discussed in Section IV. Excited states represent saddle points on the energy hypersurface. We have discussed how the MCSCF state

that represents the n th state of a certain symmetry may be required to have $n - 1$ negative Hessian eigenvalues. The constraint algorithm therefore needs to include a feature that ensures convergence to a state that has the proper number of negative Hessian eigenvalues. A further criterion is that, of course, no step length amplitude should be so large that a calculation is moved far away from the stationary point of interest. We note that in usual calculations¹¹⁻²³ most step length amplitudes are well behaved and only a few amplitudes need to be reduced and/or changed in sign. Hence for an MCSCF problem we do not recommend scaling^{31, 32} the length of $\underline{\lambda}$. Furthermore, procedures which may constrain a calculation even in the local region³² are definitely *not* advocated. Before we describe how step size and sign control modifications are introduced, we analyze a Newton–Raphson step in more detail.

The nonlinear Newton–Raphson iterative function in Eq. (60) may be transformed into a basis where the Hessian matrix is diagonal:^{13, 15}

$$\mathbf{U} \mathbf{G} \mathbf{U}^+ = \boldsymbol{\varepsilon}, \quad \varepsilon_{ij} = \delta_{ij} \varepsilon_j \quad (114)$$

The Newton–Raphson equation then becomes

$$\bar{\underline{\lambda}} = -\boldsymbol{\varepsilon}^{-1} \bar{\mathbf{F}} \quad (115)$$

where

$$\bar{\mathbf{F}} = \mathbf{U} \mathbf{F}; \bar{\underline{\lambda}} = \mathbf{U} \underline{\lambda} \quad (116)$$

In the bar basis each mode can in a sense be described independently. The second order energy change in a Newton–Raphson iteration may further be expressed as

$$\Delta E(2) = F_i \lambda_i + \frac{1}{2} G_{ij} \lambda_i \lambda_j \quad (117)$$

$$\begin{aligned} &= -F_i G_{ij}^{-1} F_j + \frac{1}{2} G_{ij} G_{im}^{-1} F_m G_{jk}^{-1} F_k \\ &= -\frac{1}{2} G_{ij}^{-1} F_i F_j \end{aligned} \quad (118)$$

In the bar basis the second order energy change therefore becomes

$$\Delta E(2) = -\frac{1}{2} \varepsilon_i^{-1} \bar{F}_i^2 = -\frac{1}{2} \bar{\lambda}_i^2 \varepsilon_i \quad (119)$$

and hence only contains terms which consist of sums of products of uncoupled contributions. The variables in the bar basis may in that sense be considered independent.

In an MCSCF optimization problem we consider optimization of both orbital and state variables. In the original basis these two types of variables are distinct whereas in the bar basis they become mixed. To obtain a measure of a variable's orbital/state character we introduce a function τ_i , defined as the norm of the part of the eigenvector U_i [see Eq. (114)] which is in the configuration space, that is,

$$\tau_i = \sqrt{\sum_j U_{ij}^2} \quad (120)$$

where j is summed over all state variables. If $\tau_i = 1$ we thus have a pure configurational i th mode and if $\tau_i = 0$ we purely orbital i th mode.

We are now ready to describe how the step size control algorithm may be implemented.

1. Incorrect Negative Hessian Eigenvalues

Undesired negative Hessian eigenvalues may originate because the original point on the energy hypersurface is close to a stationary point with a "wrong" number of negative Hessian eigenvalues or the undesired negative eigenvalues may be of more accidental nature (e.g., due to the fact that orbital optimization is a highly nonlinear problem there may be several stationary points with the same number of negative eigenvalues; this may also be caused by basis set limitations). If the Hessian matrix has a number of negative eigenvalues that deviate from the desired number, it is usually necessary to change from the Newton-Raphson step to assure rapid convergence to the proper state. The most simple situation is encountered when optimizing the lowest state of a given symmetry, i.e., a minimization problem. If we ensure that the energy [or even the second-order energy $E(2)$] decreases during each iteration, convergence is very close to being ensured since a monotonic decreasing sequence of numbers has to converge for a lower bound function. If we denote the positive (negative) Hessian eigenvalues by $\epsilon^p(\epsilon^n)$, the second-order energy change may be divided into a positive ΔE_p and a negative ΔE_n contribution:

$$\Delta E(2) = \Delta E_p + \Delta E_n = -\frac{1}{2}\epsilon_i^n \bar{\lambda}_i^2 - \frac{1}{2}\epsilon_i^p \bar{\lambda}_i^2 \quad (121)$$

For minimization, deviation from the Newton-Raphson step should be implemented such that the term ΔE_n becomes negative. This might be done in the most simple way by changing signs of the step length amplitudes and eigenvalues which correspond to negative eigenvalues, that is

$$\bar{\lambda}_i \rightarrow -\bar{\lambda}_i \quad \text{and} \quad \epsilon_i^n \rightarrow |\epsilon_i^n|, \quad \epsilon_i^n < 0 \quad (122)$$

The energy shift which corresponds to such a change becomes

$$\Delta E_2 = -\frac{1}{2}\epsilon_i''\bar{\lambda}_i^2 - \frac{1}{2}|\epsilon_i''|\bar{\lambda}_i^2 \quad (123)$$

which is a sum of two negative terms.

When optimizing an excited state of a given symmetry, it is more difficult to determine an efficient strategy for changing from the Newton–Raphson step if undesired negative Hessian eigenvalues occur. The problem consists of distinguishing between the desired and undesired negative Hessian eigenvalues. This differentiation is usually relatively simple when considering optimization problems in which no root flipping occurs since the desired negative eigenvalues then are associated with the configuration space (that is, the τ values for the corresponding mode are large (> 0.7)). The undesired negative eigenvalues then normally will be primarily of orbital nature ($\tau < 0.1$). In iterations where root flipping occurs, τ cannot usually be used to distinguish between desired and undesired negative eigenvalues since the desired negative eigenvalues may have predominantly orbital character. It is then necessary to use some more ad hoc rules to differentiate between the desired and undesired eigenvalues. One such rule is that incorrect negative eigenvalues often are small compared to the desired negative eigenvalues. If the negative eigenvalues are of the same order of magnitude, one may have to rely on a trial-and-error procedure to determine which negative eigenvalues are desired and which are not. However, this situation appears very seldom in practical calculations. After identification of all undesired negative eigenvalues, the corresponding Newton–Raphson step length amplitudes and the eigenvalues are negated to assure convergence to the desired stationary point.

2. Step Size Control

In the initial couple of Newton–Raphson iterations large step size elements $\bar{\lambda}_i$ are often encountered for a few modes.^{13,15,17} These large step size elements have to be avoided to get the Newton–Raphson sequence of iterations to reliably converge. In the following we describe how such elements may be constrained. We will control the step sizes in two parts. Initially we apply the general constraint

$$\text{if } |\bar{\lambda}_i| > K_1 \text{ set } |\bar{\lambda}_i| = K_1 \quad (124)$$

where the size of K_1 will be discussed later. Step sizes then will be constrained according to the orbital/configuration nature, that is,

$$\text{if } |\bar{\lambda}_i| > \frac{K_2}{\sqrt{\tau_i}} \text{ set } |\bar{\lambda}_i| = \frac{K_2}{\sqrt{\tau_i}} \quad (125)$$

where K_2 is smaller than K_1 . The size of K_2 will be discussed in more detail later. The second requirement [Eq. (125)] is motivated by the fact that the orbitals may change form completely during the iterative procedure (e.g., from being diffuse to being fairly tight) while for the most typical reasonable initial guesses the dominant configuration amplitudes seldom change by more than some 10–20% from the initial to the final iteration. This means that changes in the configuration amplitudes should usually be restricted more than changes of the orbital expansion coefficients.

The actual size of K_1 and K_2 also should depend on the number of step size amplitudes that have to be constrained. If many amplitudes have to be constrained $K_1(K_2)$ should be smaller than if just one step element has to be constrained in order to ensure that the total change in the orbital/configuration coefficients is no larger than a maximal value. This requirement is implemented into K_1 and K_2 as

$$K_1 = \frac{K_1^0}{\sqrt{n_1}}$$

$$K_2 = \frac{K_2^0}{\sqrt{n_2}} \quad (126)$$

where K_1^0 (and K_2^0) are constants and $n_1(n_2)$ are the number of step size elements which are larger than $K_1^0(K_2^0)$. The precise values of K_1^0 and K_2^0 are not essential and reasonable regions have been determined¹⁷ to be

$$K_1^0: 0.4-0.6$$

$$K_2^0: 0.1-0.15 \quad (127)$$

where the upper bounds are used in relatively easy optimization problems while the lower bounds are used in more difficult cases. The actual values we have used in the nonlinear Newton–Raphson calculations we describe are $K_1^0 = 0.426$ and $K_2^0 = 0.113$. This value of K_1^0 is chosen to allow that a complete change in orbital nature may take place in approximately three iterations. The K_2^0 value ensures that the dominant configuration amplitudes change a maximum of some 10–15% in three iterations.

Finally, we separately constrain very small modes ($|\epsilon_i| < 0.002$) so that $|\bar{\lambda}_i|$ is at most 0.20.

The step size and sign control algorithm we have described is a slight modification of the restricted step length methods that are used in numerical analysis.⁹ The restricted step methods are proven to assure convergence

to a local minimum if the sequence of iterations remains in a closed region and if the Hessian matrix is bounded.⁹ The step size and sign control algorithm we have described above is therefore expected to reliably converge to the lowest state of a given symmetry. Convergence to excited states requires more careful investigation.

There have been several other (usually brief) discussions of constraint procedures in the SCF^{56–58} and MCSCF literature.^{24,28,29,31,32,59} Usually these techniques do not monitor and control the individual step length amplitudes or the number of negative eigenvalues of the Hessian. Without correct step size control these procedures may move to an undesired place on the energy hypersurface that may be far from the desired stationary point. Without eigenvalue sign control, convergence may proceed to a state that is not the state of interest or even to an “unphysical” place on the energy hypersurface.

We further note that the procedure we have described above for step size and sign control,¹⁷ as well as the mode damping procedure,¹⁵ will in general only be invoked when far from convergence, that is, when higher order terms in Eq. (53) are important. Closer to convergence the step length amplitudes all approach zero and no constraint procedure is used.

Recently, Shepard et al.³² have proposed and used a procedure which reduces step length amplitudes whenever a Hessian eigenvalue is smaller in magnitude than 0.1 a.u. This reduction is performed *regardless* of the magnitude of the corresponding step length amplitudes. As expected, such a constraint procedure is rather effective in destroying second-order convergence. We have never used or advocated such a constraint procedure.^{13, 15, 17}

B. Small Hessian Eigenvalues

Small Hessian eigenvalues play a very central role for the understanding of the convergence characteristics of a Newton–Raphson MCSCF calculation. We can easily see that small eigenvalues may cause problems by reexamining Eq. (115):

$$\bar{\lambda}_i = \epsilon_i^{-1} \bar{F}_i \quad (\text{no sum}) \quad (115)$$

When far from convergence an \bar{F}_i may be large compared to the corresponding ϵ_i . Thus $\bar{\lambda}_i$ may be large. Near inflection points an ϵ_i may be small and even of the wrong sign (of course, an eigenvalue of \mathbf{G} may also be of the wrong sign because we are near the incorrect stationary point).

The eigenvalues of the configuration block $(\mathbf{A}^{CC} - \mathbf{B}^{CC})$ of the $(\mathbf{A} - \mathbf{B})$ matrix are obtained directly from Eq. (79) to be $E_m - E_0$. When a total energy difference $E_m - E_0$ is small, for example, near an avoided curve crossing, a small Hessian eigenvalue will occur, provided the configuration-orbital coupling elements are not very large. The small Hessian eigenvalue that cor-

responds to the total energy difference $E_m - E_0$ then of course has a predominantly configurational nature [τ of Eq. (120) is close to 1].

Small Hessian eigenvalues also may have their origin in the orbital space excitation operators due to the presence of low-lying "virtual orbitals" or due to an effect which we now discuss. Suppose that in an MCSCF calculation we have included a redundant variable that corresponds to an excitation between two completely occupied orbitals cl, cl' . The Hessian matrix elements which couple the $a_{cl}^+ a_{cl'} - a_{cl'}^+ a_{cl}$ excitation operator to the configuration space becomes zero for that case during a sequence of iterations since

$$a_{cl'}^+ a_{cl'} |n\rangle = a_{cl'}^+ a_{cl} |n\rangle = a_{cl}^+ a_{cl'} |0\rangle = a_{cl'}^+ a_{cl} |0\rangle = 0 \quad (128)$$

The orbital part of the Hessian matrix, however, becomes a sum of gradient matrix elements. Using Eq. (128):

$$\begin{aligned} & \langle 0 | [(a_i^+ a_j - a_j^+ a_i), (a_{cl}^+ a_{cl'} - a_{cl'}^+ a_{cl}), H] | 0 \rangle \\ &= \frac{1}{2} \langle 0 | [(a_i^+ a_j - a_j^+ a_i), [(a_{cl}^+ a_{cl'} - a_{cl'}^+ a_{cl}), H]] | 0 \rangle \\ &= -\frac{1}{2} \langle 0 | [H, [a_i^+ a_j - a_j^+ a_i, a_{cl}^+ a_{cl'} - a_{cl'}^+ a_{cl}]] | 0 \rangle \\ &= \frac{1}{2} \delta_{jcl} \langle 0 | [a_i^+ a_{cl'} - a_{cl'}^+ a_i, H] | 0 \rangle + \frac{1}{2} \delta_{icl'} \langle 0 | [a_j^+ a_{cl} - a_{cl}^+ a_j, H] | 0 \rangle \\ &\quad + \frac{1}{2} \delta_{icl} \langle 0 | [a_{cl'}^+ a_j - a_j^+ a_{cl'}, H] | 0 \rangle + \frac{1}{2} \delta_{jcl'} \langle 0 | [a_{cl}^+ a_i - a_i^+ a_{cl}, H] | 0 \rangle \end{aligned} \quad (129)$$

which first are zero when convergence has been reached. The Hessian matrix therefore has an eigenvalue that is approaching zero when convergence is approached and that is zero at convergence. An excitation between a completely occupied orbital and an orbital that is very close to being completely occupied gives a very similar result to the one discussed above and a small Hessian eigenvalue shows up also for such a case. These Hessian eigenvalues will be smaller the closer the partly occupied orbital is to being completely occupied. An analysis similar to the one in Eq. (129) can be carried out for excitations between empty and nearly empty orbitals. Small Hessian eigenvalues therefore are also expected to show up in such cases.

C. Step Size and Sign Controlled Nonlinear Newton–Raphson Calculations

To illustrate the convergence characteristic of the step size and sign controlled nonlinear Newton–Raphson approach we report some calculations on the $B^3\Sigma_u^-$ and $E^3\Sigma_u^-$ states of O_2 . The $B^3\Sigma_u^-$ and $E^3\Sigma_u^-$ states are the two lowest states of ${}^3\Sigma_u^-$ symmetry and have an avoided crossing due to a va-

TABLE I
The Basis Set for O₂ (34 STO)

Type	Exponent
1s	6.83768
	9.46635
2s	1.67543
	2.68801
2p	1.65864
	3.69445
2p _{±1}	0.30
	0.70
3d _{0, ±1}	2.0

lence-Rydberg mixing of configurations. Some Newton-Raphson calculations on the $B^3\Sigma_u^-$ and $E^3\Sigma_u^-$ states have previously been reported.¹⁷ The complexity of the calculations varies substantially over the region of the potential energy curve. In this section and in the following sections we will often perform calculations on these states to study the efficacy of various MCSCF techniques.

The Newton-Raphson MCSCF calculations we consider use 34 STO with two diffuse sets of p orbitals. The basis set is given in Table I. The diffuse functions are included to reliably describe the Rydberg configuration. All the calculations we report except when otherwise specified will be one-step calculations, and the initial guess of orbitals is a set of grand canonical Hartree-Fock orbitals⁴⁰ with occupation $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_u^4 1\pi_g^2$. In Section V.D a more thorough discussion of the requirements that may be imposed on the initial guess of orbitals and states is given. The calculations we consider include the four configurations in Table II. These configurations have been reported to be the dominant configurations of the CI calculations of Ref. 60.

TABLE II
Configurations for O₂ Calculations

Configuration ^a	Number of states
V_π	Core $3\sigma_g^2 1\pi_u^2 1\pi_g^3$ 1
V_σ	Core $3\sigma_g^1 1\pi_u^4 1\pi_g^2 3\sigma_u^1$ 2
$V_{\sigma'}$	Core $3\sigma_g^1 1\pi_u^2 1\pi_g^4 3\sigma_u^1$ 2
Ry	Core $3\sigma_g^2 1\pi_u^4 1\pi_g^1 2\pi_u^1$ 1

^aCore $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2$.

Initially we consider a $B^3\Sigma_u^-$ calculation at 2.13 a.u. and an $E^3\Sigma_u^-$ calculation at 2.10 a.u. These calculations are typical for the complexity of an average Newton-Raphson calculation on a lowest and first excited state of a given symmetry. In Tables III and IV the overall convergence characteristics of the $B^3\Sigma_u^-$ calculation at 2.13 a.u. and the $E^3\Sigma_u^-$ at 2.10 a.u. are reported. The calculations converge in six and seven iterations, respectively, to an accuracy of 10^{-10} a.u. in the total energy. The $B^3\Sigma_u^-$ state which is the lowest state of ${}^3\Sigma_u^-$ symmetry has one negative Hessian eigenvalue and three large step length amplitudes in the first iteration (iteration point 0). The step size and sign controlled algorithm thus is applied in the first iteration. In each of the second and third iterations (iteration points 1 and 2) one large step length amplitude is constrained. After the third iteration, the Newton-Raphson calculation is in the local region and the calculation converges rapidly and reliably, as expected for a second-order approach. During the entire calculation, including the initial three iterations where step size and sign control are applied, the total energy decreases rapidly and monotonically.

The $E^3\Sigma_u^-$ calculation in Table IV shows basically the same convergence characteristics as the $B^3\Sigma_u^-$ calculation in Table III. Since the $E^3\Sigma_u^-$ state is the next lowest state of ${}^3\Sigma_u^-$ symmetry, the converged Hessian should have one negative eigenvalue. No spurious negative Hessian eigenvalues were ob-

TABLE III
Convergence Characteristics of a Step Size and Sign Controlled
Newton-Raphson Calculation for the $B^3\Sigma_u^-$ State of O₂ at 2.13 a.u.

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda}\ \sim \ {}^n\underline{\epsilon}\ $
0	0.0946557369	5.02×10^{-1}	5.66×10^{-1} ^c
1	0.0246933274	2.18×10^{-1}	3.17×10^{-1} ^c
2	0.0055035052	3.62×10^{-2}	3.35×10^{-1} ^c
3	0.0009142549	3.42×10^{-2}	9.56×10^{-2}
4	0.0000433969	6.56×10^{-3}	2.65×10^{-2}
5	0.0000002359	4.30×10^{-5}	2.38×10^{-3}
6	0.0000000000	3.05×10^{-7}	1.52×10^{-5}
7	0.0000000000	$< 10^{-10}$	$< 10^{-8}$

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and ${}^{n+1}\underline{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|{}^{n+1}\underline{\lambda}\|$ is an approximation to $\|{}^n\underline{\epsilon}\| = \|{}^n\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy – 149.3086149361 a.u.

^cThe step size and sign control algorithm¹⁷ is applied in this iteration.

TABLE IV
Convergence Characteristics of a Step Size and Sign Controlled One-Step
Newton-Raphson Calculation for the $E^3\Sigma_u^-$ State of O₂ at 2.10 a.u.

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ \lambda^{n+1} - \lambda^n\ \sim \ e^n\ $
0	0.0760320763	9.40×10^{-1}	5.21×10^{-1} ^c
1	-0.0024816821	4.24×10^{-2}	2.57×10^{-1} ^c
2	-0.0008260320	3.52×10^{-2}	4.62×10^{-1} ^c
3	0.0000947284	1.91×10^{-3}	2.76×10^{-1}
4	0.0000311601	2.06×10^{-3}	2.66×10^{-1}
5	0.0000073514	4.34×10^{-3}	6.66×10^{-2}
6	0.0000000906	2.76×10^{-4}	9.17×10^{-3}
7	0.0000000000	4.16×10^{-6}	1.03×10^{-4}
8	0.0000000000	$< 10^{-10}$	—

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and λ^{n+1} is determined. Thus, at iteration point n , iteration $n+1$ is performed; $\|\lambda^{n+1}\|$ is an approximation to $\|e^n\| = \|\lambda^n - \alpha\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy - 149.2781477108 a.u.

^cThe step size and sign control algorithm¹⁷ is applied in this iteration.

served in this calculation since only the one desired negative Hessian eigenvalue shows up in all iterations. In the initial three iterations (iteration points 0-2) large step length amplitudes are constrained. After the third iteration the calculation is in the local region and converges rapidly and reliably. In the first iteration (iteration point 0) the total energy is above the converged total energy while the total energies of the second and third iteration (iteration points 1 and 2) are below the converged total energy. In the last iterations the total energies are again above the converged total energy. In calculations on the lowest state of a given symmetry the total energy of the reference state will always bound the converged total energy from above during the iterative procedure (the variational principle). In excited state calculations no such bounds can be established. Because we are converging to the first excited state of ${}^3\Sigma_u^-$ symmetry, fluctuations around the converged total energy may be expected during the iterative procedure. The $E^3\Sigma_u^-$ calculation shows a definite converging trend (the approximate error vector norm gets smaller except at iteration point 2) during the whole sequence of iterations. The norm of the GBT matrix slightly increases between the third and the fourth iteration points and from the fourth to the fifth iteration points, but overall a rapid decrease in the norm of the GBT vector is observed.

After having described broadly the convergence characteristics of typical Newton-Raphson calculations we continue discussing separately the global

and local convergence characteristics of Newton–Raphson MCSCF calculations.

1. Global Convergence

A more complete understanding of the global convergence problem is obtained by transforming the Newton–Raphson equation to the form where the Hessian matrix is diagonal [Eq. (114)]. This form clearly demonstrates the need for using a step size and sign control algorithm in the initial couple of iterations since both very large step lengths and step sizes of the wrong sign may be encountered in the initial Newton–Raphson iterations.

As an example, in Table V we report the development of the lowest Hessian eigenvalues, the unconstrained $\bar{\lambda}_i$, constrained $\bar{\lambda}_i^c$, and the configuration mode content (τ_i) of the modes corresponding to the lowest 1–4 and 7–9 Hessian eigenvalues of the sequence of Newton–Raphson iterations in Table III ($B^3\Sigma_u^-$ at 2.13 a.u.). It is clear from Table V that one undesired negative Hessian eigenvalue shows up in the initial iteration (iteration point 0). The sign of the corresponding amplitude is changed by the step sign control algorithm. The modes corresponding to the lowest Hessian eigenvalues 2–4 are also constrained at iteration point zero. At iteration point 1 the Hessian matrix has the correct structure since it is positive definite. Only one very large step length amplitude is reduced. Iteration point 2 gives one step length amplitude of -0.5587 that is reduced to -0.3314 . Iteration point 3 is in the local region and hence the calculation converges rapidly and reliably from this point on. It is clear from Table V that large step length amplitudes can be associated with small Hessian eigenvalues which may fluctuate a great deal from iteration to iteration. The larger Hessian eigenvalues show relatively smaller fluctuations during the sequence of iterations and the corresponding modes give relative small step sizes. The changes in the Hessian eigenvalues 7–9 are less than 50% from the initial to the last iteration. The corresponding step sizes are also very moderate. The larger Hessian eigenvalues show even smaller relative changes and result in very small step sizes during the sequence of iterations. When the calculation approaches the local region, the fluctuations in the smaller Hessian eigenvalues gradually become smaller.

In regions where an error term analysis is valid it is easy to see how large fluctuations in the eigenvalues can lead to large errors in $\underline{\lambda}$. From Eq. (31) we see that

$${}^k e_i = \frac{1}{2} G_{in}^{-1}(\underline{\alpha}) K_{njk}(\underline{\alpha})^{k-1} e_j {}^{k-1} e_k \quad (31)$$

$$= \frac{1}{2} G_{in}^{-1}(\underline{\alpha}) \delta G_{nk}({}^k \underline{\lambda})^{k-1} e_k \quad (130)$$

where δG is the first term in the Taylor series expansion of G . From Eq. (31) it is obvious that second-order convergence is expected for the Newton–

TABLE V

The Lowest 1–4 and 7–9 Eigenvalues ϵ of the Hessian Matrix, the Unconstrained $\bar{\lambda}$, Constrained $\bar{\lambda}^c$ Step Length Amplitudes and the Configuration Mode Content τ of the Newton–Raphson Sequence of Table III ($B^3\Sigma_u^-$ of O₂ at 2.13 a.u.)

Iteration point ^a	Hessian eigenvalue number	Hessian eigenvalue ϵ_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i	Hessian eigenvalue number	Hessian eigenvalue ϵ_i	$n+1\bar{\lambda}_i$	τ_i
0	1	−0.2410	−0.1069	0.1069	0.63	7	0.4592	−0.1558	0.09
	2	0.0598	−0.7794	−0.3012	0.06	8	0.9006	−0.0277	0.95
	3	0.1840	0.6799	0.3012	0.11	9	1.2528	0.0013	0.85
	4	0.2606	0.3573	0.3012	0.03				
1	1	0.0010	−37.1987	−0.20	0.31	7	0.7874	−0.0201	0.17
	2	0.1134	−0.1884	0.40		8	1.0858	−0.0063	0.97
	3	0.2878	−0.0492	0.02		9	1.8206	0.0364	0.31
	4	0.3226	0.1226	0.07					
2	1	0.0514	−0.5587	−0.3314	0.15	7	0.8554	0.0009	0.20
	2	0.0850	−0.0243	0.40		8	1.0252	0.0012	0.05
	3	0.1824	−0.0419	0.01		9	1.1906	−0.0033	0.94
	4	0.2454	−0.0101	0.01					
3	1	0.0596	−0.0780	0.13		7	0.8062	0.0004	0.01
	2	0.1412	0.0052	0.01		8	1.2094	0.0000	0.98
	3	0.1940	−0.0181	0.14		9	1.2880	0.0106	0.10
	4	0.2250	−0.0436	0.59					
4	1	0.0484	−0.0232	0.10		7	0.8856	0.0000	0.01
	2	0.1528	−0.0029	0.04		8	1.0950	−0.0032	0.10
	3	0.1940	0.0097	0.50		9	1.2104	0.0001	0.98
	4	0.2212	0.0045	0.10					
5	1	0.0446	−0.0022	0.10		7	0.9030	0.0000	0.01
	2	0.1554	−0.0002	0.05		8	1.0428	−0.0082	0.11
	3	0.1866	−0.0008	0.51		9	1.2088	0.0000	0.98
	4	0.2228	0.0002	0.04					
6	1	0.0442	0.0000	0.10		7	0.9046	0.0000	0.01
	2	0.1556	0.0000	0.05		8	1.0536	0.0000	0.11
	3	0.1860	0.0000	0.50		9	1.2128	0.0000	0.98
	4	0.2232	0.0000						

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and $n+1\bar{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|n+1\bar{\lambda}\|$ is an approximation to $\|\bar{\lambda}\| = \|\bar{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

Raphson approach. However, from Eq. (130) we see that if the relative changes in \mathbf{G} are large, only linear convergence is expected. This is often true in intermediate regions (i.e., just prior to the local region) of the energy hypersurface.

The four configurations we are including in the O_2 calculations differ by at least two orbital replacements and the diagonal elements of the one-electron density matrix therefore become identical to the natural occupation numbers. The natural occupation numbers for the converged $B^3\Sigma_u^-$ state are $1\sigma_g(2)$, $2\sigma_g(2)$, $3\sigma_g(1.9434)$, $1\sigma_u(2)$, $2\sigma_u(2)$, $3\sigma_u(0.0566)$, $1\pi_u(1.6244)$, $2\pi_u(0.1154)$, $1\pi_g(1.2602)$. As expected from the analysis in Section V.B the eigenvector corresponding to the lowest Hessian eigenvalue has approximately 90% of the amplitude for the excitations $1\sigma_g$, $2\sigma_g \rightarrow 3\sigma_g$ and $3\sigma_u \rightarrow 4\sigma_u$, $5\sigma_u$, $6\sigma_u$, $7\sigma_u$ and $2\pi_u \rightarrow 3\pi_u$, $4\pi_u$, $5\pi_u$ (excitations between completely occupied and nearly occupied orbitals and between nearly empty and empty orbitals). The remaining amplitude of the lowest Hessian eigenvalue is primarily in the configuration space and is predominantly on the first excited ${}^3\Sigma_u^-$ state since this state is the only low-lying state of ${}^3\Sigma_u^-$ symmetry (0.2054 a.u. above the $B^3\Sigma_u^-$).

From Table V it is further seen that the mode corresponding to the lowest Hessian eigenvalue changes character during the iterative procedure. In the initial iteration the configuration content of the mode is 0.63, whereas in the last iteration the configuration content is only 0.10. The reason for this change is that the total energy difference between the $B^3\Sigma_u^-$ state and the first excited state of ${}^3\Sigma_u^-$ symmetry (the nonoptimized $E^3\Sigma_u^-$ state) gradually increases during the iterative procedure because the orbitals get more and more optimized for describing the $B^3\Sigma_u^-$ state. When the $B^3\Sigma_u^-$ state becomes more optimized the energy lowers but the E state becomes less and less optimized and therefore acquires a higher total energy. The initial iteration has an energy difference between the $B^3\Sigma_u^-$ state and the unoptimized $E^3\Sigma_u^-$ state of 0.0196, while at iteration point 6 the energy difference has increased to 0.2054.

An example of a much more difficult global convergence problem is the $B^3\Sigma_u^-$ calculation at 2.05 a.u. In Table VI the global convergence characteristics of the $B^3\Sigma_u^-$ calculation at 2.05 a.u. are reported. At the initial seven iteration points the Hessian matrix is structurally incorrect. After iteration 8 (iteration point 7) the calculation is in the local region. The smallest Hessian eigenvalues show very large relative changes until the local region is reached. However, with regard to energy, the calculation is gradually and monotonically approaching the converged total energy from above. The complexity of this calculation is due to the very many small eigenvalues of the Hessian matrix. Comparing the $B^3\Sigma_u^-$ converged calculations at 2.13 and

TABLE VI
 Global Convergence Characteristics of the Newton-Raphson Calculation
 for the $B^3\Sigma_u^-$ State of O₂ at 2.05 a.u.

Iteration point ^a	<i>E</i>	Hessian eigenvalue			
		ϵ_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i
0	-149.200847927	-0.0282	0.7595	-0.1851	0.47
		-0.0050	-0.0249	0.0249	0.00
		0.0010	5.2503	0.2000	0.02
		0.0098	-0.3606	-0.2130	0.01
		0.0130	1.9475	0.2130	0.30
		0.0480	-0.0001		0.00
		0.0252	-0.1849		0.01
1	-149.2718590247	-0.0792	0.0783	-0.0783	0.11
		0.0070	-0.0956		0.00
		0.0121	-0.1189		0.00
		0.0164	0.0001		0.00
		0.0618	0.0404		0.01
		0.1568	-0.0127		0.28
		0.2216	-0.0050		0.00
2	-149.2780178861	-0.0012	1.0275	-0.20	0.00
		-0.0104	0.0549	-0.0549	0.00
		0.0002	0.4940	0.20	0.00
		0.0006	-0.1416		0.00
		0.0032	0.0264		0.00
		0.0150	0.0032		0.00
		0.1586	0.0034		0.28
3	-149.2785502295	-0.0182	-0.0630	0.0630	0.02
		0.0006	-7.9726	-0.20	0.00
		0.0012	0.2568	0.20	0.00
		0.0018	0.0886		0.00
		0.0044	0.0002		0.00
		0.0178	0.0024		0.00
		0.1570	0.0057		0.28
4	-149.2789667154	-0.0128	-0.2063	0.20	0.01
		0.0006	-0.1201		0.00
		0.0014	-0.0776		0.00
		0.0020	0.0475		0.00
		0.0038	-0.0303		0.00
		0.0140	0.0022		0.00
		0.1434	0.0028		0.31
5	-149.2797356270	-0.0110	-0.4600	0.20	0.00
		0.0018	0.0761		0.00
		0.0036	-0.1522		0.00
		0.0042	0.0346		0.00
		0.0126	-0.0875		0.00

TABLE VI (Continued)

Iteration point ^a	<i>E</i>	Hessian eigenvalue ϵ_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i
6	-149.2803087224	0.0382	0.0001		0.00
		0.1388	-0.0069		0.39
		-0.0036	0.9592	-0.20	0.05
		0.0062	0.0406		0.00
		0.0114	-0.0067		0.00
		0.0212	0.0896		0.02
		0.0636	-0.0233		0.05
		0.1248	0.0203		0.40
7	-149.2809077764	0.1448	0.0000		0.00
		0.0050	-0.4895	-0.4260	0.04
		0.0121	0.0387		0.01
		0.0198	-0.0085		0.00
		0.0275	-0.0158		0.04
		0.0449	0.0169		0.47
		0.0715	0.0259		0.05
		0.1317	0.0001		0.01
⋮					
12	-149.2813989690	0.0052	0.0000		0.00
		0.0087	0.0000		0.00
		0.0103	0.0000		0.00
		0.0307	0.0000		0.00
		0.0418	0.0000		0.64
		0.0734	0.0000		0.00
		0.1386	0.0000		0.31

^aAt iteration point *n*, the energy, **F**, and **G** are evaluated and $n+1\bar{\lambda}$ is determined. Thus at iteration point *n*, iteration *n*+1 is performed.

2.05 a.u. shows that the smallest Hessian eigenvalues of the 2.13 a.u. calculation are about 10 times as large as the smallest Hessian eigenvalues at 2.05 a.u. The smaller Hessian eigenvalues of the 2.05 a.u. calculation occur because the natural occupation of the $3\sigma_g$ orbital (1.9964) is much closer to 2 and the natural occupation of the $3\sigma_u$ orbital (0.0036) is much closer to being zero in the 2.05 a.u. calculation than in the 2.13 a.u. calculation. As a consequence, the 2.05 a.u. calculation is a much more difficult global convergence problem than the 2.13 a.u. calculation. The gradual and monotonical decrease in the total energy of the calculation in the global region indicates that the step size and sign control algorithm works rather well.

In some cases step size and sign control algorithms may, however, show some deficiencies; for example, in the $B^3\Sigma_u^-$ calculation at 2.10 a.u. the step size and sign control procedure does not give a monotonic decrease in the total energy during the sequence of iterations. This is shown in Table VII, where we report how the lowest two total energies of a CI calculation in the MCSCF configuration space develop during the sequence of step size and sign controlled Newton-Raphson iterations. The boldface number is the total energy for the state that we are optimizing. Step size and sign control is applied in the initial 7 iterations (through iteration point 6). One negative Hessian eigenvalue shows up at iteration points 0 and 3 and mode reversal is applied in these iterations. At iteration point 3 an amplitude of -2.6826 corresponding to the negative Hessian eigenvalue -0.0030 a.u. is reversed and set equal to 0.4260 . Although mode reversal certainly has to be applied in this case, it appears that a step length amplitude of less than 0.4260 would be more optimal for this case and probably would have resulted in a lower total energy in the next iteration. We would, however, point out that such a deficiency of the step size and sign control algorithm of the aforementioned character is *very* exceptional and for that reason we have not modified the step size and sign algorithm.

The global convergence characteristics of excited state calculations are very similar to the characteristics that have been observed in the lowest state calculations. One difference is that the Hessian matrix in excited state calculations should have desired negative Hessian eigenvalues. To illustrate the global convergence characteristics of an excited state calculation we report in Table VIII the development of the lowest six Hessian eigenvalues ϵ_i , the unconstrained $\bar{\lambda}_i$, constrained $\bar{\lambda}_i^c$ step length amplitude, and configuration mode content τ_i of the sequence of Newton-Raphson iterations of Table IV ($E^3\Sigma_u^-$ at 2.10 a.u.). The Hessian matrix has just one negative Hessian eigenvalue in all iterations and hence no undesired negative eigenvalues show up in the calculation. The step size control algorithm is applied in the initial 3 iterations (at iteration points 0–2). It is seen that the character of the mode corresponding to the lowest Hessian eigenvalue changes character completely from initially (iteration point 0) being configurationally dominated ($\tau = 0.66$) to finally (iteration points 3–7) being of orbital nature ($\tau = 0.00$). This change is most easily understood examining the development of the lowest eigenvalues of the $A^{CC} - B^{CC}$ block of the Hessian matrix and assuming that the coupling with the orbital block is small. The eigenvalues of $A^{CC} - B^{CC}$ are the energy differences $E_n - E_0$ [see Eq. (79)], where E_0 and E_n are roots of the MCSCF CI eigenvalue problem. In Table IX we report how the lowest two roots of an MCSCF CI calculation develop during the iteration sequence of Table VIII. In all iterations the state that has been assigned $E^3\Sigma_u^-$ is the lowest root of the CI problem. In the initial iteration $E_1 - E_0$ is

TABLE VII
 The Lowest CI Eigenvalues of the MCSCF Configuration List (Table II) and the Weights (CI Amplitude²) of the Corresponding Dominant Configurations for the Newton-Raphson Sequence of the $B^3\Sigma_u^-$ State at 2.10 a.u.

Iteration point ^a	Configuration weight		
	Lowest eigenvalues ^b	Ry^c	V_π^c
0	-149.20211563 -149.18489141	0.830 0.169	0.156 0.751
1	-149.25836212 -149.10710741	0.141 0.848	0.781 0.108
2	-149.27099790 -148.97413671	0.092 0.873	0.832 0.058
3	-149.27796839 -148.92061763	0.098 0.842	0.835 0.056
4	-149.27135167 -148.94947506	0.087 0.871	0.840 0.053
5	-149.27823226 -148.97009508	0.114 0.840	0.819 0.073
6	-149.28499762 -149.02026003	0.162 0.799	0.776 0.114
7	-149.29118188 -149.08785027	0.258 0.718	0.694 0.200
8	-149.29313749 -149.14333597	0.403 0.581	0.560 0.336
9	-149.29352110 -149.13662670	0.391 0.594	0.570 0.324
10	-149.29352285 -149.13577628	0.387 0.598	0.576 0.322
11	-149.29352285 -149.13579476	0.387 0.598	0.576 0.322

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and ${}^{n+1}\lambda$ is determined. Thus at iteration point n , iteration $n+1$ is performed.

^bThe boldface values correspond to the $B^3\Sigma_u^-$ state. The lightface values refer to the (unoptimized) E state.

^cConfigurations Ry and V_π are defined in Table II.

TABLE VIII
 The Lowest Six Hessian Eigenvalues ϵ , the Corresponding Unconstrained
 $\bar{\lambda}$, Constrained $\bar{\lambda}^c$ Step Length, and Configuration Mode
 Content τ of the Newton-Raphson Sequence of Table IV
 $(E^3\Sigma_u^-$ of O₂ at 2.10 a.u.)

Iteration point ^a	Hessian eigenvalue ϵ_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i
0	-0.1300	-0.0087		0.66
	0.0348	0.0173		0.01
	0.0656	0.0419		0.01
	0.0828	0.3529		0.26
	0.1356	-0.7770	-0.4260 ^b	0.07
	0.2196	0.0268		0.00
1	-0.0488	0.2452	0.2098 ^b	0.29
	0.0210	-0.0274		0.00
	0.0370	0.0045		0.00
	0.0980	0.0641		0.37
	0.1046	0.0075		0.00
	0.2612	0.0261		0.05
2	-0.0108	0.5294	0.4260 ^b	0.03
	0.0032	-0.1387		0.00
	0.0052	-0.0886		0.00
	0.0138	0.0103		0.00
	0.0636	0.0027		0.00
	0.1120	-0.0595		0.48
3	-0.0042	0.2182		0.00
	0.0018	0.0891		0.00
	0.0024	-0.1428		0.00
	0.0110	0.0055		0.00
	0.0548	0.0007		0.00
	0.1300	-0.0167		0.39
4	-0.0028	-0.0641		0.00
	0.0010	-0.1789		0.00
	0.0016	-0.1866		0.00
	0.0058	0.0014		0.00
	0.0272	0.0015		0.00
	0.1308	-0.0036		0.38
5	-0.0056	0.0201		0.00
	0.0016	0.0397		0.00
	0.0024	0.0494		0.00
	0.0080	0.0026		0.00
	0.0376	0.0008		0.00
	0.1294	0.0013		0.39

TABLE VIII (Continued)

Iteration point ^a	Hessian eigenvalue ϵ_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i
6	-0.0050	0.0003		0.00
	0.0014	0.0065		0.00
	0.0020	0.0064		0.00
	0.0068	0.0004		0.00
	0.0324	0.0001		0.00
	0.1298	0.0001		0.39
7	-0.0050	0.0000		0.00
	0.0014	0.0001		0.00
	0.0020	0.0001		0.00
	0.0068	0.0000		0.00
	0.0322	0.0000		0.00
	0.1298	0.0000		0.39

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and $n+1\bar{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed. $\|n+1\bar{\lambda}\|$ is an approximation to $\|\bar{\lambda}\| = \|n\bar{\lambda} - \alpha\|$ [see Eq. (71)].

^bThe step size and sign control algorithm¹⁷ is applied to this mode in this iteration.

only 0.0173 a.u., whereas at convergence $E_1 - E_0$ has increased to 0.2526. The lowest Hessian eigenvalues therefore at convergence have a much smaller configurational content.

The $E^3\Sigma_u^-$ state is the first excited state of $^3\Sigma_u^-$ symmetry. It might seem a little surprising that we assign the lowest root of the MCSCF CI calculation to this state. Orbitals that are optimal for the $E^3\Sigma_u^-$ state (the first excited state of $^3\Sigma_u^-$ symmetry) may, however, be very far from optimal for the $B^3\Sigma_u^-$ state (the lowest state of $^3\Sigma_u^-$ symmetry) and the total energy of the optimized $E^3\Sigma_u^-$ may therefore be lower than the total energy of the $B^3\Sigma_u^-$ state (when nonoptimized orbitals are used to construct the $B^3\Sigma_u^-$ state). In the region of a potential energy curve where an avoided curve crossing occurs it is often observed that the n th state in energy of a certain symmetry is not assigned to the n th root of the MCSCF CI problem. Such a situation is often referred to as root flipping. In the present case the state we assign $E^3\Sigma_u^-$ has one negative Hessian eigenvalue and further the total energy of the converged state of $E^3\Sigma_u^-$ state is 0.01537514 a.u. above the total energy of the converged $B^3\Sigma_u^-$ state (see Table VII).

The global convergence of the $E^3\Sigma_u^-$ calculation of 2.10 a.u. is typical for an excited state calculation. Much simpler calculations may occasionally be encountered; for example, the $E^3\Sigma_u^-$ calculation at 2.13 a.u. is not far from being in the local region with an initial guess of grand canonical

TABLE IX

The Lowest CI Eigenvalues of the MCSCF Configuration List (Table II) and the Weights of the Corresponding Dominant Configurations for the Newton-Raphson Sequence of Table IV ($E^3\Sigma_u^-$ of O₂ at 2.10 a.u.).

Iteration point ^a	Lowest eigenvalues ^b	Configuration Weight <i>Ry</i> ^c	Configuration Weight <i>V_π</i> ^c
0	-149.20211563	0.830	0.156
	-149.18489141	0.169	0.751
1	-149.28073894	0.840	0.153
	-149.12849771	0.153	0.737
2	-149.27923909	0.928	0.070
	-149.08661756	0.069	0.845
3	-149.27814012	0.944	0.054
	-149.03228053	0.054	0.936
4	-149.27811808	0.944	0.056
	-149.02553854	0.056	0.941
5	-149.27814396	0.941	0.057
	-149.02717088	0.058	0.939
6	-149.27814763	0.941	0.057
	-149.02661730	0.058	0.938
7	-149.27814771	0.941	0.057
	-149.02658825	0.058	0.938

^aAt iteration point *n*, the energy, **F**, and **G** are evaluated and λ^{n+1} is determined. Thus at iteration point *n*, iteration *n* + 1 is performed.

^bThe boldface values correspond to the $E^3\Sigma_u^-$ state. The lightface values refer to the (unoptimized) *B* state.

^cConfiguration *Ry* and *V_π* are defined in Table II.

Hartree-Fock orbitals. In Table X we report the convergence characteristics of the $E^3\Sigma_u^-$ calculation at 2.13 a.u. The step size control algorithm is only applied in the initial iteration, where a step length of 0.63 is restricted to 0.404. The simplicity of the 2.13 a.u. calculation is caused by the fact that the numerical value of the smallest converged Hessian eigenvalue at 2.13 a.u. is 0.06 a.u. and thus is 50 times larger than the numerically smallest converged eigenvalue at 2.10 a.u. The total energy of the $E^3\Sigma_u^-$ state is also seen from Table X to fluctuate around the converged total energy.

In the two excited state $E^3\Sigma_u^-$ calculations reported so far it has not been necessary to reverse modes because the Hessian matrix had the desired number of negative eigenvalues. However, in more complicated excited state

TABLE X
Convergence Characteristics of the Step Size and Sign Controlled
Newton-Raphson Sequence for the $E^3\Sigma_u^-$ State of O₂ at 2.13 a.u.

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \underline{\lambda}\ \sim \ \lambda^n \underline{\mathbf{G}}\ $
0	0.0590242739	9.72×10^{-1}	4.62×10^{-1} ^c
1	-0.0003831946	8.36×10^{-2}	2.03×10^{-1}
2	-0.0002286547	1.36×10^{-2}	7.26×10^{-2}
3	0.0000018925	2.64×10^{-3}	5.66×10^{-3}
4	0.0000000000	1.02×10^{-5}	7.68×10^{-5}
5	0.0000000000	< 10^{-8}	< 10^{-8}

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and λ^{n+1} is determined. Thus, at iteration point n , iteration $n+1$ is performed. $\|\lambda^{n+1}\|$ is an approximation to $\|\lambda^n \underline{\mathbf{G}}\| = \|\lambda^n \underline{\lambda} - \underline{\mathbf{G}}\|$ [see Eq. 71].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy. -149.2533729769 a.u.

^cThe step size and sign control algorithm¹⁷ is applied in this iteration.

calculations undesired negative Hessian eigenvalues often show up. The excited state calculations reported in Ref. 17 on the $^4\Pi$ states of BO provide an example of a case in which mode reversal was needed to obtain convergence to the desired $^4\Pi$ state. In the calculation on the first excited state of $^4\Pi$ symmetry at 3.2 a.u. it was required to reverse modes in the initial 6 iterations. In the calculation on the second excited $^4\Pi$ state at 3.2 a.u. it was necessary to reverse modes in the initial 16 iterations before the Hessian matrix acquired the desired number of negative eigenvalues. However, these complications are exceptional and usual cases do not have these difficulties.¹¹⁻²³

When undesired negative eigenvalues occur in the Hessian matrix for excited state calculations, one has to face the problem of distinguishing between desired and undesired negative eigenvalues. We described in Section V.A the rules we have used. When it is required to reverse modes as many as 16 times before the Hessian matrix gets the desired number of negative eigenvalues, however, it is obvious that the procedure applied is not optimal for this case and has to be improved. We point out that the BO calculations are very complex and not typical. Our experience has been that the mode reversal technique works very well in more usual cases. A real advantage of the mode controlling technique is, of course, that it is firmly based on the theoretical characteristics of the state of interest and on assuring that no step length amplitude is extraordinarily large.

The global convergence characteristics of step size and sign controlled Newton–Raphson calculations may be summarized as follows. In the initial phase of the global convergence problem both step size and sign control have to be applied to ensure convergence to the desired state. As the calculation approaches the local region, sign control usually is not necessary since the Hessian matrix acquires the desired number of negative eigenvalues. In the last phase of the global convergence problem only step size control is normally required to bring the calculation into the local region. The complexity of the global convergence problem is closely associated with the magnitude of small Hessian eigenvalues. Generally, as the eigenvalues of the Hessian matrix become smaller, the global convergence problem becomes more complex. Furthermore, small Hessian eigenvalues show up in connection with orbitals that are very close to being completely occupied or completely unoccupied as well as when two or more MCSCF CI roots are close in energy or when the calculation is near an inflection point on the energy hypersurface.

2. Local Convergence

The error term analysis of a sequence of Newton–Raphson iterations in Eq. (34) is only valid in the local region. The error term which results when one Newton–Raphson iteration is carried out is given in Eq. (31). A measure of the size of the error vector is the norm of the vector given in Eq. (71). Introducing the Frobenius norm⁶¹ $\|\cdot\|$ for a matrix \mathbf{A} of dimension m and n

$$\|\mathbf{A}\| = \left(\sum_i^m \sum_j^n A_{ij}^2 \right)^{1/2} \quad (131)$$

allows us to rewrite the error term at the k th Newton–Raphson iteration point

$$\|{}^k \mathbf{e}\| \leq \frac{1}{2} \|\mathbf{G}^{-1}(\alpha) \mathbf{K}(\alpha)\| \|{}^{k-1} \mathbf{e}\| \|{}^{k-1} \mathbf{e}\| \quad (132)$$

or

$$\frac{\|{}^k \mathbf{e}\|}{\|{}^{k-1} \mathbf{e}\|^2} \leq \frac{1}{2} \|\mathbf{G}^{-1}(\alpha) \mathbf{K}(\alpha)\| \quad (133)$$

The ratio $\|{}^k \mathbf{e}\|/\|{}^{k-1} \mathbf{e}\|^2$ thus is smaller than or equal to the constant $\frac{1}{2} \|\mathbf{G}^{-1}(\alpha) \mathbf{K}(\alpha)\|$.

Let us now examine how the ratio $\|{}^k \mathbf{e}\|/\|{}^{k-1} \mathbf{e}\|^2$ develops in the local region in some of the calculations that have been described previously in de-

tail. Initially consider the $E^3\Sigma_u^-$ calculation at 2.13 a.u. reported in Table X. At iteration points $k = 2, 3$, and 4 , we obtain 1.76, 1.37, and 1.47, respectively, for the ratio $\|\epsilon^k\|/\|\epsilon^{k-1}\|^2$. The ratio is fairly constant during the iterative procedure. In Table XI we have reported the step length amplitudes of the modes corresponding to the lowest five Hessian eigenvalues. The step length norm corresponding to these five modes constitutes more than 99% of the full step length norm. Significant contributions originate from no more than two modes. The ratios $\|\epsilon^k\|/\|\epsilon^{k-1}\|^2$ are relatively constant in the local region.

The inverse Hessian eigenvalues enter directly into the error term [see Eq. (31)], and from Eq. (130) it is obvious that large fluctuations occur in the ratios $\|\epsilon^k\|/\|\epsilon^{k-1}\|^2$ when large fluctuations are observed in the small Hessian eigenvalues. In the $E^3\Sigma_u^-$ calculation at 2.10 a.u. of Table VIII the ratios become 3.49, 0.94, 2.07, and 1.22 for iteration points $k = 4, 5, 6$, and 7. Table VIII shows that also in this case only a few (about two or three) modes given significant contributions to the step length norm. The relative large fluctuations in the ratios of the 2.10 a.u. calculation compared to the 2.13 a.u. calculation are caused by the large fluctuations that occur in the smallest Hessian eigenvalues in the 2.10 a.u. calculation. The lowest Hessian eigenvalues at 2.10 are thus $-0.0042, -0.0028, -0.0056, -0.0050$, and -0.0050 for points $k = 3, 4, 5, 6$, and 7 (all in the local region), while the smallest positive eigenvalues are respectively 0.0018, 0.0010, 0.0016, 0.0014, and 0.0014. At 2.13 a.u. for points $k = 1, 2, 3$, and 4 (all in the local region) the lowest Hessian eigenvalues are $-0.1238, -0.1282, -0.1312$, and -0.1316 , while the smallest positive eigenvalues are respectively 0.053, 0.048, 0.060, and 0.060. Hence the fluctuations in the smallest eigenvalue are relatively small for the case at 2.13 a.u. compared to the case at 2.10 a.u.

We examine a condition under which the equality sign in Eq. (133) is valid. Let us consider the basis where $G(\alpha)$ is diagonal [the bar basis of Eqs. (114)–(116)] and let us assume that the m th component of the error vector $\bar{\epsilon}$ dominates in both the $(k-1)$ st and k th iteration point. The error vector of the k th iteration point may then be written as

$${}^k\bar{\epsilon}_m = \frac{1}{2} \epsilon_m^{-1} \bar{K}_{mmm} {}^{k-1}\bar{\epsilon}_m {}^{k-1}\bar{\epsilon}_m \quad (134)$$

where in Eq. (134) and the following equation no summation is implied by repeated indices. Since ${}^k\bar{\epsilon}_m / {}^{k-1}\bar{\epsilon}_m^2 = \frac{1}{2} \epsilon_m^{-1} \bar{K}_{mmm}$, the equality sign of Eq. (133) is expected to be approximately valid if one of the components of the error vector $\bar{\epsilon}$ dominates. Large amplitudes of ${}^{k+1}\bar{\lambda}$ may frequently be due to the presence of small eigenvalues of $G({}^k\lambda)$. If there are several small eigenvalues ϵ_i which give several large $\bar{\lambda}_i$, then equality (134) is not generally

TABLE XI

The Lowest Five Hessian Eigenvalues ε_i , the Corresponding Unconstrained $\bar{\lambda}_i$, Constrained $\bar{\lambda}_i^c$ Step Length Amplitudes, and Configuration Mode Content τ_i of the Newton-Raphson Sequence of Table X ($E^3\Sigma_u^-$ of O₂ at 2.13 a.u.)

Iteration point ^a	Hessian eigenvalue ε_i	$n+1\bar{\lambda}_i$	$n+1\bar{\lambda}_i^c$	τ_i
0	-0.2052	-0.0689		0.76
	0.0316	0.6456 ^b	0.3012	0.03
	0.0396	0.5580 ^b	0.3012	0.03
	0.0660	0.0112		0.00
	0.1968	-0.0189		0.18
1	-0.1238	-0.1936		0.88
	0.0532	-0.0438		0.00
	0.0656	0.1287		0.01
	0.0864	0.0000		0.00
	0.2068	0.0557		0.08
2	-0.1282	-0.0656		0.87
	0.0482	-0.0087		0.00
	0.0634	-0.0273		0.01
	0.0834	0.0048		0.00
	0.1850	0.0102		0.09
3	-0.1312	0.0009		0.87
	0.0604	0.0015		0.00
	0.0808	0.0052		0.01
	0.1042	-0.0009		0.00
	0.1714	0.0010		0.09
4	-0.1316	0.00003		0.87
	0.0600	0.00000		0.00
	0.0800	0.00004		0.01
	0.1034	0.00000		0.00
	0.1718	0.00000		0.09

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are evaluated and $n+1\bar{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|n+1\bar{\lambda}\|$ is an approximation to $\|\eta\bar{\alpha}\| = \|n\bar{\lambda} - \bar{\alpha}\|$ [see Eq. (71)].

^bThe step size and sign control algorithm¹⁷ is applied for this mode at this iteration.

expected to be valid for successive iterations (although this certainly may happen in some cases).

The development of the ratio $\|{}^k\mathbf{e}\|/\|{}^{k-1}\mathbf{e}\|^2$ for the $B^3\Sigma_u^-$ calculation at 2.13 a.u. (see Table V) has a fairly constant value for consecutive iterations. For iteration points $k = 4, 5$, and 6 (all in the local region) we obtain the ratios 2.90, 3.39, and 2.68, respectively. The relative small deviation in these ratios is in agreement with the fact that there is only one fairly small Hessian eigenvalue (~ 0.044) in the local region of this calculation which produces the dominant λ component.

The local convergence of these Newton-Raphson calculations is straightforwardly understood from the error term analysis of Section III. The ratios $\|{}^k\mathbf{e}\|/\|{}^{k-1}\mathbf{e}\|^2$ in the local region are approximately equal to a constant of the order 1–5. The size of the actual constant depends on the magnitude of the smallest Hessian eigenvalues.

3. Comparison of One-Step and Two-Step Newton-Raphson Calculations

To illustrate the convergence characteristics of the two-step Newton-Raphson calculations we report in Table XII the convergence characteristics of a two-step Newton-Raphson calculation on the $E^3\Sigma_u^-$ state at 2.10 a.u. The orbital and state rotation parameters are not determined simultaneously in the two-step approach, and a reliable measure of the total step length (and therefore also of the length of the error vector) of a given iteration is not easily available. We have for that reason only reported the difference between the total energy of a given step of the iterative procedure and the converged total energy to measure the error of a given iteration. In Table IV the convergence characteristics of the one-step Newton-Raphson approach are given. The convergence characteristics of the one- and two-step calculations are very similar, and both calculations converge to an accuracy of 10^{-10} a.u. in the total energy in seven iterations. The similar convergence characteristics of one- and two-step Newton-Raphson calculations have been observed in many other calculations.^{11,13,15,17}

However, in some cases convergence difficulties have been encountered with the two-step approach where the one-step approach has converged rapidly and reliably. To illustrate this point we will examine a sequence of converged reparametrized one- and two-step calculations of the $E^3\Sigma_u^-$ state at 2.13 a.u. We have previously discussed how the four-configuration calculation of Table II can be reparameterized to a five-, six-, and seven-configuration calculation by successively adding one of the configurations core $3\sigma_g^2 1\pi_u^4 1\pi_g^n n\pi_u$ (with $n = 3, 4$, and 5) to the four-configuration case.¹⁷ In Table XIII we report for the converged four-, five-, six-, and seven-configuration calculations the lowest two eigenvalues of the MCSCF CI matrix

TABLE XII
Convergence Characteristics of a Two-Step Newton-Raphson
Calculation for the $E^3\Sigma_u^-$ State at 2.10 a.u.

Iteration point ^a	$E - E^{\text{CONV}}$ ^b
0	0.0760320763 ^c
1	-0.0039816356 ^c
2	-0.0003771275 ^c
3	0.0000738856
4	0.0000049646
5	-0.0000020503
6	0.0000000068
7	0.0000000000

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} are determined and ${}^{n+1}\lambda$ is evaluated. Thus at iteration point n , iteration $n+1$ is performed.

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy - 149.2781477108 a.u.

^cStep size control is applied in this iteration.

(column 2), the Hessian matrix (column 3), the configuration block of the Hessian matrix (column 4), and the reduced Hessian matrix (column 5). The MCSCF CI calculations of column 2 are performed using the respective MCSCF configuration lists, and the boldface MCSCF CI energy is therefore the converged MCSCF total energy. The converged MCSCF energies are the same for each of these reparametrized MCSCF calculations. From column 2, we see that root flipping occurs for the four- and five-configuration cases but does not occur for the six- and seven-configuration cases.

Since we are converging to the $E^3\Sigma_u^-$ state, the Hessian matrix of all the reparametrized calculations has one negative eigenvalue (see column 3). If the root flipping does not occur at one point of the iterative procedure, the reduced Hessian will have purely positive eigenvalues. Columns 4 and 5 demonstrate that the one negative eigenvalue of the Hessian matrix gets distributed to either the \mathbf{G}^{CC} block (when there is no root flipping) or to the reduced Hessian (when root flipping occurs), in agreement with the derivation of Appendix D. When the root flipping occurs, the negative eigenvalue of the reduced Hessian is larger numerically than the negative eigenvalue of full Hessian, as would be expected from the analysis of Appendix D (see Fig. 1).

Table XIII demonstrates one of the difficulties which may be encountered with two-step Newton-Raphson MCSCF calculations when there are states

of the same symmetry that are fairly close in energy. The five-configuration case has a negative reduced Hessian eigenvalue, -756.1 a.u. This occurs because the two lowest $^3\Sigma_u^-$ CI states (see column 2) are very close in energy. The partitioned form of the Hessian eigenvalue equation [Eq. (D.3)] then has a vertical asymptote that will be very close to the line $E = 0$ [see the plot of the multivalued function $\epsilon(E)$ of Eq. (D.4) in Fig. 1]. Very small changes in the location of this asymptote will result in large changes of the negative eigenvalue. When a two-step calculation is carried out under such circumstances, very large fluctuations may occur in the eigenvalues of the reduced Hessian matrix, and it is not at all straightforward to devise a constraint procedure due to the behavior of these eigenvalues. In the two-step calculations reported in Table XIII, we also note that very large changes occur in the lowest eigenvalue of the reduced Hessian when the configuration list is changed. Contrary to this, the lowest eigenvalue of the Hessian matrix that is used in the one-step procedure (column 3) is very stable when the configuration list is increased. The eigenvalues of the Hessian matrix will not even fluctuate very much in the one-step procedure when the partitioning which leads to the two-step procedure is not defined (one energy difference $E_m - E_0$

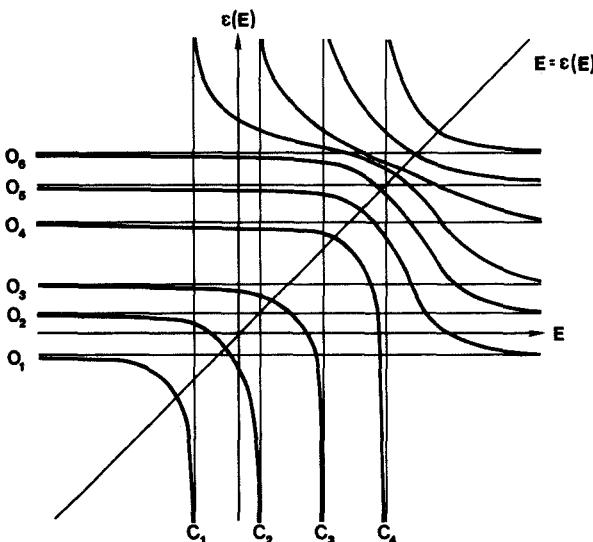


Fig. 1. A plot of the multivalued function $\epsilon(E)$ as a function of E for a case where root flipping occurs. The branches of this function represent the eigenvalues of the matrix $\mathbf{G}^{OO} - \mathbf{G}^{OC}(\mathbf{G}^{CC} - E\mathbf{I})^{-1}\mathbf{G}^{CO}$. The eigenvalues of the Hessian matrix are represented by the interaction of these branches with the line $E = \epsilon(E)$. The intersection of the branches with the line $E = 0$ are the eigenvalues of reduced Hessian matrix. The vertical and horizontal asymptotes are the eigenvalues of the matrices \mathbf{G}^{CC} and \mathbf{G}^{OO} , respectively.

TABLE XIII
Converged MCSCF Results for the $E^3\Sigma_u^-$ State of O₂ at 2.13 a.u.

		C1 ^a eigenvalues (a.u.)	One-step Hessian eigenvalues (a.u.)	G ^{CC} eigenvalues (a.u.)	Two-step reduced Hessian eigenvalues (a.u.)
4 conf. ^b	Lowest	-149.25337297 ^f	-0.13169982	0.02821812	-9.75475424
	Next lowest	-149.22515485	0.06002166	0.58710661	0.06014124
5 conf. ^c	Lowest	-149.25337297 ^f	-0.13496382	0.00059799	-756.16741520
	Next lowest	-149.25277498	0.06002612	0.12977741	0.06015418
6 conf. ^d	Lowest	-149.26482575	-0.13541722	-0.01145278	0.06015630
	Next lowest	-149.25337297 ^f	0.06002658	0.12611592	0.08036454
7 conf. ^e	Lowest	-149.27527358		-0.02190061	
	Next lowest	-149.25337297 ^f		0.11310116	

^aConfiguration interaction calculation performed using the configurations defining the MCSCF problem.

^bThe first four configurations in Table II.

^cConfigurations of footnote *b* with the additional configuration (core) $3\sigma_g^2 1\pi_u^4 1\pi_g^1 3\pi_u^1$.

^dConfigurations of footnote *c* with the additional configuration (core) $3\sigma_g^2 1\pi_u^4 1\pi_g^1 4\pi_u^1$.

^eConfigurations of footnote *d* with the additional configuration (core) $3\sigma_g^2 1\pi_u^4 1\pi_g^1 5\pi_u^1$.

^fConverged MCSCF total energy.

is equal to zero). The one-step procedure must therefore be advocated when MCSCF calculations are carried out on difficult cases, for instance, avoided crossings.

D. Initial Guess of Orbitals and States

In order to successfully carry out a Newton-Raphson sequence of iterations it is important to have an initial guess of orbitals easily available that is of such a quality that global convergence easily can be obtained with the step size and sign controlled Newton-Raphson procedure or other stable procedure. To assure convergence to the desired state, it is important to be able to obtain an initial guess of the reference state that resembles the desired state enough for the iterative procedure to converge to this state. In this section we address the problem of how to determine an initial guess of orbitals and states.

To assure convergence of a Newton-Raphson calculation it is our experience that it suffices to use as an initial guess of orbitals a set of grand canonical Hartree-Fock orbitals⁴⁰ with occupations corresponding ap-

proximately to the ones of the dominant configuration. Occupations which correspond to another configuration may in many cases also suffice;¹⁷ for example; in the O₂, $^3\Sigma_u^-$ calculations we have reported, the V_{π} and the Ry configurations of Table II are the dominant configurations and we have used occupation numbers corresponding to the ground state $X^3\Sigma_g^-$ in the grand canonical Hartree–Fock calculation. A set of canonical restricted Hartree–Fock orbitals with $N - 1$ potential virtuals^{38,39} may suffice equally well as initial guess of orbitals for a Newton–Raphson MCSCF calculation.¹⁸ We have often chosen to use a set of grand canonical Hartree–Fock orbitals as the initial guess because this set of orbitals is the easiest and most straightforward set of orbitals to determine and because the grand canonical Hartree–Fock equations are independent of the spacial and spin symmetry of the considered state, differing only in the specification of the set of occupation numbers.⁴⁰

With an initial guess of orbitals at hand, the initial guess of the state may be obtained by selecting an appropriate root of a CI calculation in the MCSCF configuration space. The Undheim–McDonald theorem⁴² is most often used for selecting this root. The Undheim–McDonald theorem tells that the n th lowest root of a certain symmetry of a CI problem is an upper bound for the n th state of that symmetry. The n th root of the MCSCF CI problem is therefore usually assigned to represent the n th state. Such an assignment suffices in most cases; an example is the $E^3\Sigma_u^-$ at 2.13 a.u. reported in Table X.

However, an assignment of the initial guess of the reference state based on the Undheim–McDonald theorem may lead to an erroneous result in some cases. An example is the $B^3\Sigma_u^-$ calculation at 2.10 a.u. reported in Table VII. Even though the $B^3\Sigma_u^-$ state is the lowest state of $^3\Sigma_u^-$ symmetry, the correct assignment corresponding to this initial guess of orbitals is the second lowest root of the MCSCF CI problem. The energy difference between the first and second root of the MCSCF CI problem is very small, and after the first iteration the state that has been assigned to the $B^3\Sigma_u^-$ state has dropped in energy to become the lowest root of the MCSCF CI problem, where it stays until convergence. The state we have assigned to be the $B^3\Sigma_u^-$ state has a dominant configuration amplitude on the V_{π} configuration. The amplitude of the dominant configuration changes very little and may be used when we choose the state from the MCSCF CI for a two-step calculation.

The reason why the second root of an MCSCF CI problem may be assigned as the lowest state is, of course, that orbitals which optimally describe one configuration may be very far from optimal for describing another configuration. Our initial guess of orbitals describes very poorly the dominant V_{π} configuration; it more optimally describes the Ry configuration.

When we are optimizing the $B^3\Sigma_u^-$ state the orbitals become more and more optimal for the V_n configuration, and after the first iteration the lowest root of the MCSCF CI problem has dominant weight for the V_n configuration. If we insist on using the lowest root of the initial MCSF CI problem as a guess for the $B^3\Sigma_u^-$ state and on converging to a state which had a positive definite Hessian matrix, convergence cannot be obtained. However, if we use the lowest root as an initial guess for the $E^3\Sigma_u^-$ state (which is required to have one negative Hessian eigenvalue), convergence is easily obtained as described in Table VIII. Root flipping occurs for the entire sequence of iterations in the $E^3\Sigma_u^-$ calculation at 2.10 a.u. (see Table IX). In the $B^3\Sigma_u^-$ calculation at 2.10 a.u. (Table VII) root flipping only occurs in the initial iteration. The assignment of a given root of an MCSCF CI problem to a given state may be, of course, very difficult when the Undheim-McDonald theorem cannot be used to carry out the assignment. When this is the case, a trial-and-error procedure may even have to be used before the correct assignment can be performed. As we have previously noted, since in most previous MCSCF calculations² no monitoring of the characteristics of the states was done, we expect that many MCSCF calculations reported in the literature are erroneous.

Occasionally when it is difficult to get a calculation to converge it may be advantageous to try to reparametrize the calculation. The philosophy of a reparametrization is to move some of the orbital parameters to the configuration space by including some additional configurations and to ensure at the same time that the total energy of the converged state does not depend on whether the reparametrization is performed or not. For example, in Ref. 17 we have shown how the O₂ MCSCF calculation which includes the four configurations of Table II may be reparameterized by including the additional configuration core $3\sigma_g^2 1\pi_u^4 1\pi_g^1 3\pi_u^1$. (See Ref. 17 and Appendix A, where a more thorough discussion can be found about this reparametrization.)

In the reported $E^3\Sigma_u^-$ calculation at 2.05 a.u. it was necessary to perform a reparametrization to get the calculation to converge. With an initial guess of grand canonical Hartree-Fock (HF) orbitals, the only "reasonable" initial CI state in the four-configuration calculation was far from having the configuration amplitudes of the converged state. The CI state was primarily valence-like (having large amplitude for the V_n configurations), with a low Rydberg weight of only 1%. The optimized state is highly correlated with a Rydberg weight of 42%. The use of grand canonical HF orbitals as an initial guess led in the four-configuration calculation to convergence to a "simple solution" without any Rydberg character at all. The convergence to this simple solution was not a result of initial spurious negative eigenvalues, since these were absent until very near the stationary point where the V_n configurations had a very large amplitude and the Ry configuration had zero am-

plitude for the obtained converged state. For a state that has zero amplitude on the *Ry* configuration, the operators $a_{n\pi_u}^+ a_{2\pi_u}$ ($n = 3, 4, \dots$) become redundant. The GBT amplitude corresponding to a redundant variable is zero at convergence. In this case, the Hessian matrix also had a zero eigenvalue at convergence due to the redundant variable. Because the GBT amplitudes went faster to zero than the Hessian eigenvalue [see Eq. (129)], no warning was given that convergence was being obtained to an undesired state until right at convergence where the Hessian matrix possessed the zero eigenvalue and amplitude of zero for the *Ry* configuration. So-called simple solutions such as this one may occasionally appear in MCSCF calculations if the initial guess of configurations and orbitals is far from the desired result. This kind of solution, in which some configurations completely drop out of the calculation and some operators become redundant, is rare but not all that unusual in MCSCF calculations.¹⁷ This case is the first one we have experienced where mode reversal was not sufficient to avoid convergence to the simple solution.

It was clear that our problems were caused by an inadequate description of the $2\pi_u$ orbital, so we decided to reparametrize the calculation to a five-configuration calculation. With the five-configuration parameterization, the initial guess for the *E* state came closer to the final configurational weights, basically because the description of the π_u space now is qualitatively more correctly treated in our initial guess. The restructuring of our parameter space led to convergence without problems. The total energy for this converged calculation with the five configurations included in the CI space is, of course, the same as the total energy obtained in a converged four-configuration calculation; however, the converged π_u orbitals in the two calculations are different. When the set of orbitals obtained in the converged five-configuration calculation was used as the initial guess of a four-configuration calculation, convergence was easily obtained to the proper stationary point with one negative eigenvalue of the Hessian and, of course, with the same total energy as in the five-configuration calculation. We obtained a weight of 55% *V_u* and 42% of *Ry* in the converged four-configuration calculation.

E. Conditions for Improved Iterative Methods

The previous derivation of the Newton–Raphson method as well as the convergence characteristics of actual calculations indicates that when it is possible to obtain a reasonable initial guess of the MCSCF state the constrained Newton–Raphson method is usually very stable and rapidly converging in situations where the Hessian matrix has no very small eigenvalues.^{11–23, 27–37} In cases where very small Hessian eigenvalues occur, the global convergence characteristics of the method are not as satisfactory. With constraints, some 3–20 iterations may be required to get the calculation into

the local region.¹⁷ In the local region the unconstrained Newton–Raphson method converges reliably and rapidly.

The conditions that an optimization procedure should satisfy to show improved convergence properties compared to a constrained Newton–Raphson model depend on whether the method is designed for obtaining improved local or global convergence. In the local region computational efficiency to reach the stationary point is the property which decides whether a method is feasible, whereas in global convergence the stability and reliability of the method for moving to the local region of the proper stationary point when the initial guess is “far away” is the decisive factor.

Large step lengths cannot be accepted in a global method. A global method has to be designed in such a way that step sizes are so moderate that a given region is not left until the method has detected that there are no stationary points in that region. Because the global convergence problems with the step size and sign controlled Newton–Raphson algorithm may be caused by the appearance of very small Hessian eigenvalues, alternative improved global methods are required that are less sensitive to the appearance of very small Hessian eigenvalues. Such a method is obtained, for example, if the iterative function is based on a cubic expansion of the energy function. We discuss the global convergence characteristics of an iterative cubic calculation in Section VIII.

In the local region we shall show that improved computational efficiency compared to a Newton–Raphson approach may be obtained if the Hessian matrix is kept fixed during a sequence of several iterations. When a fixed Hessian series of iterations is carried out, a new energy gradient is constructed in each iteration. The difference between the new energy gradient and the one of the previous iteration contains information about the Hessian at the new point, and it may be desirable also to use this information to update the Hessian matrix during the iterative sequence.^{9,21} We shall demonstrate that such Hessian update schemes constitute an efficient way for obtaining local convergence. Furthermore, Hessian updates will be shown to sometimes have attractive global convergence properties when implemented with a step size (and sign) control algorithm. In the next two sections we shall describe in more detail the convergence characteristics of improved local approaches.

VI. GENERALIZED AND FIXED HESSIAN APPROACHES

A. The Generalized Newton–Raphson Perturbative Approach

In local regions perturbation expansions can be used to solve Eq. (59). A convenient starting point for such an expansion is obtained by rearranging

Eq. (59) with $\underline{\lambda}$ equal to the stationary point $\underline{\alpha}$:

$$\underline{\alpha}_k = -G_{kl}^{-1}F_l - \frac{1}{2}G_{kl}^{-1}K_{lmn}\underline{\alpha}_m\underline{\alpha}_n - \frac{1}{6}G_{kl}^{-1}M_{lmnp}\underline{\alpha}_m\underline{\alpha}_n\underline{\alpha}_p + \dots \quad (135)$$

In a given iteration (k) we may then construct an approximation ${}^k\underline{\lambda}^{(p)}$ to $\underline{\alpha}$ so that ${}^k\underline{\lambda}^{(p)}$ fulfills Eq. (135) to order p . The set of iterative functions which then is obtained is often referred to as generalized Newton–Raphson perturbative procedures.^{16,18} We will now study the derivation and the convergence characteristics of these procedures. It is noted that Eq. (59) tacitly assumes an expansion around the origin (i.e., ${}^0\underline{\lambda} = \mathbf{0}$), so a nonlinear transformation is required between two iterations.

The simplest approximation to $\underline{\alpha}$ is obtained when $\underline{\alpha}$ is set equal to zero on the right-hand side of Eq. (135). We then obtain¹⁶

$${}^k\underline{\lambda}_k^{(1)} = -G_{kl}^{-1}F_l \quad (136)$$

which is the nonlinear Newton–Raphson iterative function of Eq. (60). A second step may be carried out on Eq. (135), where $\underline{\alpha}$ on the right-hand side of Eq. (135) is from Eq. (136). Keeping all terms that are quadratic in ${}^k\underline{\lambda}^{(1)}$ gives the function

$${}^k\underline{\lambda}_k^{(2)} = {}^k\underline{\lambda}_k^{(1)} - \frac{1}{2}G_{kl}^{-1}K_{lmn}{}^k\underline{\lambda}_m^{(1)}{}^k\underline{\lambda}_n^{(1)} = {}^k\underline{\lambda}_k^{(1)} - \frac{1}{2}G_{kl}^{-1}K_{lmn}G_{mi}^{-1}F_iG_{nj}^{-1}F_j \quad (137)$$

which is the Chebyshev formula.^{16,62}

A general series of formulas which are consistent through still higher powers in ${}^k\underline{\lambda}^{(1)}$ may be derived through iterating on Eq. (135), keeping terms through the desired power in ${}^k\underline{\lambda}^{(1)}$. The iterative function ${}^k\underline{\lambda}^{(p)}$ is thus obtained by replacing $\underline{\alpha}$ on the right-hand side of Eq. (135) by ${}^k\underline{\lambda}^{(p-1)}$, keeping all terms through order $({}^k\underline{\lambda}^{(1)})^p$ in the power series expansion. The parameter set ${}^k\underline{\lambda}^{(\infty)}$ corresponds to the stationary point $\underline{\alpha}$ if the sequence converges.

Since ${}^k\underline{\lambda}^{(p)}$ contains all terms through order p in ${}^k\underline{\lambda}^{(1)}$, ${}^k\underline{\lambda}^{(p)} - \underline{\alpha}$ is correct up to $O[({}^k\underline{\lambda}^{(1)})^{p+1}]$. From Eq. (29) we see that ${}^k\underline{\lambda}^{(p)} - \underline{\alpha}$ is thus correct to $O({}^k\epsilon^{p+1})$. In Section VIII we derive the error term of the Chebyshev formula [Eq. (137)] and report the results of some Chebyshev calculations together with the results of some other cubically convergent schemes.

B. The Linear Fixed Hessian Approach

We will now prove that formulas which show quadratic, cubic, quartic, ... convergence characteristics may alternatively be derived by carrying out a sequence of fixed Hessian iterations where the $\underline{\lambda}$ parameters are determined from the equation²²

$${}^p\underline{\lambda}_i = -G_{ij}^{-1}(F_j + F_j({}^1\underline{\lambda}) + F_j({}^2\underline{\lambda}) + \dots + F_j({}^{p-1}\underline{\lambda})) \quad (138)$$

It is assumed that the $\underline{\lambda}$ parameters at the initial iteration point are zero, and

no explicit reference to this point is made in matrices evaluated at this point. The expansion in Eq. (138) replaces the calculation of the higher order derivatives at one point in Eq. (135) with the calculation of gradients at several points (e.g., at ${}^0\lambda = 0, {}^1\lambda, {}^2\lambda, \dots, {}^{p-1}\lambda$; see Appendix B). The matrix \mathbf{F} is thus evaluated at points recursively defined by Eq. (138).

We will prove that when Eq. (138) is applied n times, error in ${}^n\lambda$ will be of order $n+1$ in ${}^0\mathbf{e}$. Induction is used to prove this point. Since the initial iteration of Eq. (138) is a Newton-Raphson iteration, it is clear that the foregoing statement is true when $p=1$. For a given n it is now assumed that ${}^{n-1}\lambda$ contains errors of order $O({}^0\mathbf{e})^n$ and it is now proved that ${}^n\lambda$ contains errors of order $O({}^0\mathbf{e})^{n+1}$. To do this, rewrite Eq. (138) as

$${}^n\lambda_i = {}^{n-1}\lambda_i - G_{ij}^{-1} F_j({}^{n-1}\lambda) \quad (139)$$

Expanding G_{ij} and $F_j({}^{n-1}\lambda)$ around the exact solution $\underline{\alpha}$ using Eq. (27) and (28) gives

$$\begin{aligned} {}^n\lambda_i &= {}^{n-1}\lambda_i - \left(G_{rs}(\underline{\alpha}) + K_{rs,t}(\underline{\alpha}) [{}^0\lambda_t - \alpha_t] + \dots \right)_{ij}^{-1} \\ &\quad \cdot \left(G_{jk}(\underline{\alpha}) [{}^{n-1}\lambda_k - \alpha_k] + \frac{1}{2} K_{jkm}(\underline{\alpha}) [{}^{n-1}\lambda_k - \alpha_k] [{}^{n-1}\lambda_m - \alpha_m] + \dots \right) \end{aligned} \quad (140)$$

Expanding the inverse matrix of Eq. (140) then gives

$$\begin{aligned} {}^n\lambda_i &= \alpha_i + G_{ij}^{-1}(\underline{\alpha}) K_{jpq} [{}^0\lambda_p - \alpha_p] [{}^{n-1}\lambda_q - \alpha_q] + \dots \\ &\quad - \frac{1}{2} G_{ij}^{-1}(\underline{\alpha}) K_{jkm}(\underline{\alpha}) [{}^{n-1}\lambda_k - \alpha_k] [{}^{n-1}\lambda_m - \alpha_m] + \dots \end{aligned} \quad (141)$$

Hence ${}^n\lambda_i$ has an error vector

$${}^n e_i = L_{ipq} {}^0 e_p {}^{n-1} e_q \quad \text{for } n > 1, \quad {}^1 e_i = \frac{1}{2} L_{ipq} {}^0 e_p {}^0 e_q \quad (142)$$

where L_{ipq} is defined in Eq. (33). Successive use of Eq. (142) gives the error vector ${}^n\mathbf{e}$ in terms of ${}^0\mathbf{e}$

$${}^n e_i = L_{ipq} {}^0 e_p {}^{n-1} e_q = \frac{1}{2} L_{il_1 k_1} {}^0 e_{l_1} L_{k_1 l_2 k_2} {}^0 e_{l_2} L_{k_2 l_3 k_3} {}^0 e_{l_3} \cdots L_{k_{n-1} l_n k_n} {}^0 e_{l_n} {}^0 e_{k_n} \quad (143)$$

We have thus proved that such a sequence of n fixed Hessian iterations contains errors of order $n+1$ in ${}^0\mathbf{e}$ and has the error term given in Eq. (143).

A fixed Hessian sequence may alternatively be carried out as

$${}^p \underline{\lambda}^L = - G_{ij}^{-1} \left(F_j + {}^1 \tilde{F}_j + {}^2 \tilde{F}_j + \dots + {}^{p-1} \tilde{F}_j \right) \quad (144)$$

where ${}^{p-1}\bar{\mathbf{F}}$ is the *GBT* vector of Eq. (54) evaluated in the orbital basis

$${}^{p-1}a_r^+ = \sum_s a_s^+ [\exp(-{}^{p-1}\kappa)]_{sr} \quad (145)$$

and state basis defined through the coefficient matrix ${}^{p-1}\mathbf{C}$

$${}^{p-1}\mathbf{C} = {}^0\mathbf{C} \exp(-{}^{p-1}\mathbf{S}) \quad (146)$$

From the previous derivation it follows straightforwardly that the sequence of fixed Hessian iterations in Eq. (144) will have the same order of convergence and the same error terms as the one of Eq. (138).²² The advantage of using Eq. (144) is that partial derivative matrices only have to be evaluated at the point $\lambda = \mathbf{0}$. This vector is evaluated easily (see Section III.D, Appendix B, and Appendix C).

By applying Eq. (138) or Eq. (144) cubically, quartically, ..., convergent MCSCF schemes have been derived without ever explicitly constructing the third, fourth, ... derivatives of the total energy.

A sequence of iterations that uses Eq. (138) or Eq. (144) results in a linear transformation of the rotational parameters λ between each step of the iterative procedure, as opposed to the previously used fixed Hessian approaches^{17, 29, 35} that carry out a nonlinear transformation of the rotational parameters. When a sequence consists of n steps of the nonlinear fixed Hessian approach, it can be shown to have a total order of convergence of $n + 1$. The proof is more involved than the proof for a linear fixed Hessian approach (see Ref. 22, where a detailed derivation is carried out).

C. Optimal Use of Fixed Hessian Approaches

We now discuss how fixed Hessian approaches most efficiently may be used in an MCSCF calculation. In the previous section it was shown that two fixed Hessian iterations, namely, a Newton–Raphson iteration followed by an iteration in which the Hessian matrix is kept fixed, have a total order of convergence of 3 with the error vector [Eq. (143)]:

$${}^2e_i = \frac{1}{2} L_{ijk} {}^0e_k L_{jmn} {}^0e_m {}^0e_n \quad (147)$$

The error vector of the (initial) Newton–Raphson iteration [see Eq. (31)] is

$$e_j^{\text{NR}} = \frac{1}{2} L_{jmn} {}^0e_m {}^0e_n \quad (148)$$

so

$$\|e^{\text{NR}}\| \leq \frac{1}{2} \|L^0 e\| \|{}^0e\| \quad (149)$$

Using Eq. (147), Eq. (148) may be written as

$${}^2e_i = L_{ijk} {}^0e_k e_j^{\text{NR}} \quad (150)$$

that is

$$\|^2\mathbf{e}\| \leq \|L^0\mathbf{e}\| \|e^{NR}\| \quad (151)$$

When the Newton–Raphson iteration diverges, it is reasonable to assume that

$$\|e^{NR}\| > \|^0\mathbf{e}\| \quad (152)$$

Equations (149), (151), and (152) show that the upper bound for the length of the error vector $\|^2\mathbf{e}\|$ is larger than the upper bound for $\|e^{NR}\|$. It is therefore reasonable to apply fixed Hessian approaches only in the local region where a straightforward application of the Newton–Raphson approach (i.e., no constraints) converges.

In the local region we compare three different iterative schemes: (1) a Newton–Raphson approach, (2) a fixed Hessian approach, and (3) a combination of fixed Hessian and Newton–Raphson approaches. We initially compare the efficiency of a Newton–Raphson approach and a fixed Hessian approach. A Newton–Raphson iteration which is succeeded with $N^F - 1$ fixed Hessian iterations gives a total order of convergence of $N^F + 1$. (Note that the first step in a fixed Hessian sequence is a Newton–Raphson step.) When a Newton–Raphson iteration is succeeded by $N - 1$ Newton–Raphson iterations, the total order of convergence is 2^N . The total number of fixed Hessian steps (N^F) which give the same order of convergence as N Newton–Raphson steps is thus

$$N^F = 2^N - 1 \quad (153)$$

Table XIV exemplifies Eq. (153) for $N = 1, 2, \dots, 10$. The comparison in Table XIV is solely based on order of convergence. A more accurate comparison is obtained by analyzing the corresponding error terms. The error vector obtained when carrying out N^F steps of the fixed Hessian series [see Eq. (143)] is

$${}^{2^N-1}e_i^{FH} = \frac{1}{2} L_{il_1k_1} L_{k_1l_2k_2} \dots L_{k_{(2^N-2)}l_{(2^N-1)}k_{(2^N-1)}} {}^0e_{l_1} {}^0e_{l_2} \dots {}^0e_{l_{(2^N-1)}} {}^0e_{k_{(2^N-1)}} \quad (154)$$

while the error vector of the corresponding sequence of N Newton–Raphson iterations [see Eq. (34)] becomes

$${}^N e_i^{NR} = 2^{-(2^N-1)} L_{ik_1 \dots k_{(2^N)}} {}^0e_{k_1} \dots {}^0e_{k_{(2^N)}} \quad (155)$$

where $L^{(N)}$ is defined through the recurrence relation of Eq. (35). The error vectors in Eqs. (154) and (155) have the same order and both contain $(2^N - 1)$

TABLE XIV
Convergence Characteristics of the Newton-Raphson versus
Fixed Hessian Approaches to MCSCF

Number of Newton-Raphson iterations ^a	Equivalent number of fixed Hessian iterations ^a
1 $\left(\frac{1}{2}\right)$	1 $\left(\frac{1}{2}\right)$ (same as Newton-Raphson)
2 $\left(\frac{1}{8}\right)$	3 $\left(\frac{1}{2}\right)$
3 $\left(\frac{1}{128}\right)$	7 $\left(\frac{1}{2}\right)$
4 $\left(\frac{1}{32768}\right)$	15 $\left(\frac{1}{2}\right)$
5 $\left(\frac{1}{2.147 \times 10^9}\right)$	31 $\left(\frac{1}{2}\right)$
⋮	⋮
10 $\left(\frac{1}{2^{1023}}\right)$	1023 $\left(\frac{1}{2}\right)$

^aThe numbers in parentheses refer to the constant multiplying the error term. See Eqs. (34) and (143).

L supermatrices connected in different ways. If we assume that the difference in the way the **L** supermatrices are coupled together does not affect the actual values of the error terms, then the difference between the error terms depends only on the constants multiplying the error terms. The error vectors satisfy then the relation

$$\|{}^N\mathbf{e}^{\text{NR}}\| = 2^{-(2^N - 2)} \|{}^{2^N - 1}\mathbf{e}^{\text{FH}}\| \quad (156)$$

For $N = 2$ the error term of the Newton-Raphson iteration is thus 1/4 of the error term in the corresponding fixed Hessian series; for $N = 3$ the corresponding ratio is 1/64; and so on. In Table XIV we have written out explicitly the error term constant for $N = 1, \dots, 10$. When N increases, the error term of the Newton-Raphson sequence becomes smaller and smaller compared to the corresponding error terms of the fixed Hessian sequence. The error term analysis thus indicates that the number of fixed Hessian iterations of Table XIV is a lower bound for the actual number of fixed Hessian steps which have to be carried out to simulate the N Newton-Raphson steps.

The optimal use of a fixed Hessian approach is, of course, not to use a fixed Hessian procedure to exactly obtain local convergence. An uncritical

use of a fixed Hessian procedure may require a very large number of fixed Hessian steps, for example if seven Newton-Raphson iterations are required to obtain convergence, a fixed Hessian approach would require at least 127 steps in which the Hessian matrix was kept fixed.

Fixed Hessian approaches are most efficiently used when combined with Newton-Raphson iterations. Two Newton-Raphson iterations may for example efficiently be replaced by three fixed Hessian iterations (see Table XIV; recall that a fixed Hessian procedure only requires evaluation of the GBT matrix at each step of the iterative procedure). In Ref. 22 it is proved that the most efficient use of fixed Hessian approaches is obtained if about one third of the total computer time (including transformation time) is spent carrying out iterations in which the Hessian matrix is kept fixed with only GBT matrix elements constructed.

D. Numerical Results

In the previous sections we have analytically examined the local convergence characteristics of Newton-Raphson and fixed Hessian iterative schemes. In this section we will report the results of some sample calculations on the $E^3\Sigma_u^-$ state of O_2 to illustrate the derivations of the previous sections. Our analytical derivations indicate that the fixed Hessian calculation may have a smaller radius of convergence than the Newton-Raphson calculation. To illustrate this point, in Table XV we report a linear fixed

TABLE XV
Convergence Characteristics of the Newton-Raphson and the Linear Fixed Hessian Approach
for the $E^3\Sigma_u^-$ State of O_2 at 2.13 a.u.

Iteration point ^a	Newton-Raphson ^c				Fixed Hessian ^c			
	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\alpha}\ $	$E - E^{\text{CONV}}$	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\alpha}\ $		
0	0.0590242739	9.72×10^{-1}	8.72×10^{-1}	0.0590242739	9.72×10^{-1}	8.72×10^{-1}		
1	0.0126959929	2.86×10^{-1}	4.02×10^{-1}	0.0126959929	2.86×10^{-1}	1.89×10^0		
2	0.0000375222	6.00×10^{-2}	1.64×10^{-1}	0.0793910245	5.58×10^{-1}	1.71×10^0		
3	0.0000256747	1.06×10^{-2}	1.74×10^{-2}	0.0683720648	6.18×10^{-1}	3.83×10^0		
4	0.0000000001	1.02×10^{-4}	3.19×10^{-4}	0.2610539743	1.59×10^0	1.65×10^0		
5	0.0000000000	3.90×10^{-8}	$\sim 9 \times 10^{-8}$	0.1923140885	1.80×10^0	1.23×10^0		

^aAt iteration point n , the energy, \mathbf{F} , and \mathbf{G} (Newton-Raphson only) are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\alpha}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ denotes the difference between the total energy of the present step of the iterative procedure and the total energy of the converged calculation. The total energy of the converged calculation is 149.2533729769 a.u.

^cNo constraints have been applied in this calculation.

Hessian [Eq. (144)] calculation and an unconstrained Newton-Raphson calculation on the $E^3\Sigma_u^-$ state of O_2 at an internuclear distance of 2.13 a.u. with our usual initial guess of grand canonical Hartree-Fock orbitals with occupations $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_u^4 1\pi_g^2$.

The Newton-Raphson (NR) calculation in Table XV converges while the linear fixed Hessian calculation diverges. In both calculations the first iteration gives a step length norm of 0.872. A constrained NR calculation which used the same initial guess of orbitals and states has previously been reported in Table X. In the constrained NR calculations two large step length amplitudes were each reduced from 0.646 and 0.558 to 0.301 in the first iteration. The convergence rate for the NR calculations in Tables XV and X are almost identical, indicating that for this case the reduction of the step length amplitudes does not significantly affect convergence. Step sizes such as those of the initial step of the NR calculation of Table XV may actually be considered to be on the borderline for allowed step sizes in the local region. After the first iteration both the NR calculations of Tables XV and X are definitely in the local region and no constraints are applied.

The orbitals and states obtained after constraining the step length of the first iteration in the NR calculation in Table X were used as "initial" guess of orbitals and states for the fixed Hessian series of iterations given in Table XVI. The initial iteration (the Newton-Raphson iteration) of the fixed Hessian series resulted in a step length norm of 0.203. This calculation is thus within the local region of an NR calculation. The fixed Hessian series of calculations therefore, as expected, converges. The error term of a fixed Hessian sequence of iterations [Eq. (142)] shows that the error vector decreases linearly in a fixed Hessian sequence of iterations. The norm of the approximate error vector in Table XVI is observed to decrease with approximately a factor of 0.3–0.4 in each iteration. One reason that the norm of the approximate error vector of Table XVI is not decreasing with exactly the same factor is of course that the error term analysis only is precisely correct infinitesimally close to the stationary point. Higher order terms will cause some deviations from linearity when the initial guess is not very close to being converged. The convergence behavior of Table XVI is rather typical for a fixed Hessian series when the series is initiated as soon as the local region is obtained.

In some extreme cases, the error term of a fixed Hessian series of calculations may make the fixed Hessian sequence of iterations converge extremely slowly. As an example in Table XVII we report a fixed Hessian calculation on the $E^3\Sigma_u^-$ state of O_2 at 2.10 a.u. using the orbitals and states at NR iteration point 4 of Table IV as an initial guess. Even though the fixed Hessian calculation certainly is initiated in the local region (the norm of the Newton-Raphson step is 0.266) the error terms dominate the fixed Hessian

TABLE XVI

Convergence Characteristics ($E^3\Sigma_u^-$ State of O₂ at 2.13 a.u.) of the Linear Fixed Hessian Calculation After Newton-Raphson Iteration 1
(i.e., Starting at Iteration Point 1) of Table X

Iteration point ^a	$E - E^{\text{CONV}}$	$\ \mathbf{F}\ $	$\ \lambda^{n+1} - \lambda^n\ \sim \ \mathbf{e}\ $
0	-0.0003831946	8.36×10^{-2}	2.03×10^{-1}
1	-0.0002286547	1.36×10^{-2}	7.25×10^{-2}
2	0.0000317736	1.27×10^{-2}	3.04×10^{-2}
3	-0.0000032557	3.04×10^{-3}	1.16×10^{-2}
4	0.0000003723	2.24×10^{-3}	2.63×10^{-3}
5	-0.0000000322	1.94×10^{-4}	1.05×10^{-3}
6	0.0000000035	2.08×10^{-4}	3.18×10^{-4}
7	-0.0000000004	3.04×10^{-5}	1.25×10^{-4}
8	0.0000000000	2.40×10^{-5}	3.32×10^{-5}
9	0.0000000000	3.64×10^{-6}	1.33×10^{-5}
10	0.0000000000	2.54×10^{-6}	3.68×10^{-6}
11	0.0000000000	4.02×10^{-7}	1.49×10^{-6}
12	0.0000000000	2.80×10^{-7}	4.00×10^{-7}
13	0.0000000000	4.78×10^{-8}	1.62×10^{-7}
14	0.0000000000	3.60×10^{-8}	6.30×10^{-8}
15	0.0000000000	$< 10^{-8}$	$\sim 2.5 \times 10^{-8}$

^aAt iteration point n , the energy and \mathbf{F} are evaluated and λ^{n+1} is determined. Thus at iteration point n , iteration $n + 1$ is performed; $\|\lambda^{n+1} - \lambda^n\|$ is an approximation to $\|\mathbf{e}\| = \|\lambda - \alpha\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy. The converged total energy is -149.2533729769 a.u.

sequence of iterations to such an extent that the fixed Hessian sequence of iterations is hardly converging. If the individual step length amplitudes are analyzed in this calculation, it appears that the calculation is dominated completely by a fluctuation of a single κ element which has an initial value of -0.190 and after that becomes 0.140, -0.136, 0.118, Such a fluctuation is rather unusual but may occur and may make the rate of convergence of a fixed Hessian calculation extremely slow.

The error vector of the n th iteration of a fixed Hessian sequence is proportional to the error vector of the $(n - 1)$ st iteration with a proportionality matrix which contains the error vector of the initial iteration [see Eq. (142)]. If the fixed Hessian series therefore is initiated in a region where \mathbf{e} is small, the error term would be smaller than if the fixed Hessian series is initiated in a region where \mathbf{e} is large. In Table XVIII and XIX we report two fixed Hessian series of calculations on the $E^3\Sigma_u^+$ state of O₂ at 2.13 a.u., which are initiated after the second and third iteration of the Newton-Raphson calculation in Table X. Recall that Table XVI reported the fixed Hessian

TABLE XVII
Convergence Characteristics of a Linear Fixed Hessian Calculation
After Newton-Raphson Iteration 4 (i.e., Starting at Iteration Point 4)
of Table IV ($E^3\Sigma_u^-$ State of O₂ at 2.10 a.u.)

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\epsilon}\ $
0	0.0000311601	2.06×10^{-3}	2.66×10^{-1}
1	0.0000073515	4.34×10^{-3}	1.73×10^{-1}
2	0.0000101563	3.30×10^{-3}	1.63×10^{-1}
3	0.0000057317	3.06×10^{-3}	1.44×10^{-1}
4	0.0000071905	2.82×10^{-3}	1.37×10^{-1}
5	0.0000046223	2.64×10^{-3}	1.27×10^{-1}
6	0.0000056164	2.54×10^{-3}	1.22×10^{-1}
7	0.0000038491	2.40×10^{-3}	1.15×10^{-1}
8	0.0000045908	2.30×10^{-3}	1.11×10^{-1}
9	0.0000032803	2.20×10^{-3}	1.06×10^{-1}
10	0.0000038580	2.12×10^{-3}	1.02×10^{-1}

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n + 1$ is performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\epsilon}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy. The converged total energy is -149.2781477108 a.u.

TABLE XVIII
Convergence Characteristics ($E^3\Sigma_u^-$ State of O₂ at 2.13 a.u.) Using the
Linear Fixed Hessian Approach Initiated After Iteration 2 (i.e., Starting
at Iteration Point 2) of the Newton-Raphson Calculation in Table X

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\epsilon}\ $
0	-0.0002286547	1.36×10^{-2}	7.26×10^{-2}
1	0.0000018925	2.64×10^{-3}	7.22×10^{-3}
2	0.0000000813	2.24×10^{-4}	2.09×10^{-3}
3	0.0000000051	6.72×10^{-5}	4.76×10^{-4}
4	0.0000000003	1.55×10^{-5}	1.15×10^{-4}
:	:	:	:
:	:	:	:

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n + 1$ performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\epsilon}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy. The converged total energy is -149.2533729769 a.u.

series initiated after the first Newton-Raphson iteration of Table X. The ratios between norms $\|{}^{n+1}\underline{\lambda}\|/\|{}^n\underline{\lambda}\| \sim \|{}^n\mathbf{e}\|/\|{}^{n-1}\mathbf{e}\|$ in Table XVI are 0.3–0.4, whereas the ratios of Table XVIII are 0.2–0.3 in qualitative agreement with the fact that the error vector ${}^0\mathbf{e}$ of the fixed Hessian sequence in Table XVIII is smaller than the error vector of the sequence of Table XVI. The $\|{}^{n+1}\underline{\lambda}\|/\|{}^n\underline{\lambda}\|$ ratio of Table XIX is 0.004, which is substantially smaller than the ratios of Tables XVI and XVIII, in agreement with the fact that the initial error vector ${}^0\mathbf{e}$ of the fixed Hessian sequence of iterations in Table XIX is substantially smaller than the initial error vector of the fixed Hessian sequences of Tables XVI and XVIII.

In Table XIV we have reported the number of fixed Hessian steps that are required to simulate a certain number of Newton-Raphson steps to obtain an equal total order of convergence. When error terms are as important as in the fixed Hessian calculation of Table XVII, a comparison of the total order of convergence of fixed Hessian and Newton-Raphson sequences has very little meaning. However, in less pathological cases, Table XIV gives a relatively accurate measure of the number of fixed Hessian steps that are required to simulate a certain number of NR steps. To exemplify this we compare the NR sequence of iterations in Table X with the fixed Hessian series of Tables XVI, XVIII, and XIX. The fixed Hessian series of Tables XVI, XVIII, and XIX use the orbitals and states after the first, second, and third iterations of the NR series in Table X as the initial guess. Comparing Table X and XVI we see that the initial two NR iterations roughly can be replaced by three fixed Hessian iterations (one NR iteration followed by two iterations in which the Hessian matrix is kept fixed). The length of the error vec-

TABLE XIX
Convergence Characteristics of the Linear Fixed Hessian Initiated
After Iteration 3 (i.e., Starting at Iteration Point 3) of the
Newton-Raphson Calculation of Table X

Iteration point ^a	$E - E^{\text{CONV}}$ ^b	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\ \sim \ {}^n\mathbf{e}\ $
0	0.0000018925	2.64×10^{-3}	5.66×10^{-3}
1	0.0000000000	1.02×10^{-5}	5.42×10^{-5}
2	0.0000000000	1.60×10^{-7}	3.29×10^{-7}
3	0.0000000000	$< 10^{-8}$	$< 10^{-8}$

^aAt iteration point n , the energy and \mathbf{F} are evaluated and ${}^{n+1}\underline{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|{}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\|$ is an approximation to $\|{}^n\mathbf{e}\| = \|{}^n\underline{\lambda} - \underline{g}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy. The converged total energy is -149.2533729769 a.u.

tor after two NR iterations is 5.66×10^{-3} , whereas the error of three fixed Hessian steps is 1.16×10^{-2} . Three initial NR steps correspond according to Table XIV to seven fixed Hessian steps. Table X shows that three NR steps have an error vector with a norm about 7.68×10^{-5} while from Table XVI seven fixed Hessian steps have an error vector norm of approximately 1.25×10^{-4} in reasonable agreement with the predictions of Table XIV. According to Table XIV four NR iterations correspond to 15 fixed Hessian steps. Table XVI shows that the error vector after four NR iterations is less than 10^{-8} while the norm of the error vector after 15 fixed Hessian steps is about 2.5×10^{-8} . The predictions of Table XIV are thus roughly confirmed by the calculations of Tables X and XVI. A comparison of Table X and Tables XVIII and XIX shows the same trends.

In Table XIV we further reported the error term constant, that is, the constant which is a multiplicative factor on the error term. The error term constant decreases rapidly in the Newton-Raphson sequence while the constant always is $\frac{1}{2}$ in the fixed Hessian sequence. The error term analysis thus indicates that the predicted number of fixed Hessian steps is a lower limit to the actual number of fixed Hessian steps. The calculations in Tables X and XVI confirm this prediction. Because of the relation between the error term constants, the difference between the predicted and the actual number of fixed Hessian steps should increase as the number of Newton-Raphson steps increases. This behavior is not clearly observed in the calculations reported in Tables X and XVI. This may be due somewhat to our use of Eq. (71) to estimate the error.

All the fixed Hessian calculations reported so far have been carried out using Eq. (144) where a linear transformation of variables is performed in between each step of the fixed Hessian series of calculations. In Section VI.B we discussed that the linear [Eqs. (144) or (138)] and the nonlinear fixed Hessian approach have the same convergence characteristics. In Table XX we report a fixed Hessian calculation which differs from the calculation of Table XVI in that the nonlinear fixed Hessian procedure is used.²² The result of Tables XX and XVI are in very close agreement. The agreement is so pronounced that we expect the error terms of the two calculations to be essentially the same. The error term of the nonlinear fixed Hessian approach has not yet been derived.

Thus we conclude that in the local region fixed Hessian procedures may be advantageously employed to improve the overall computational efficiency of MCSCF procedures. With these procedures only new **F** matrix elements are calculated at each iteration point. Thus these procedures may yield considerable savings over usual Newton-Raphson techniques since a smaller two-electron integral transformation is required [if, e.g., Eq. (144) is used] and a small number of matrix elements (**F** instead of both **F** and **G**) are required in each iteration. These savings, of course, may be offset by the slower con-

TABLE XX
 Convergence Characteristics ($E^3\Sigma_u^-$ State of O₂ at 2.13 a.u.) of the
 Nonlinear Fixed Hessian Approach After Newton-Raphson Iteration 1
 (Starting at Iteration Point 1) of Table X

Iteration point ^a	$E - E^{\text{CONV}}$	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\epsilon}\ $
0	-0.0003831946	8.36×10^{-2}	2.03×10^{-1}
1	-0.0002286547	1.36×10^{-2}	7.25×10^{-2}
2	0.0000318517	1.31×10^{-2}	3.05×10^{-2}
3	-0.0000032488	2.98×10^{-3}	1.16×10^{-2}
4	0.0000003252	2.10×10^{-3}	2.62×10^{-3}
5	-0.0000000284	2.80×10^{-4}	1.03×10^{-3}
6	0.0000000028	1.91×10^{-4}	3.03×10^{-4}
7	-0.0000000003	2.64×10^{-5}	1.20×10^{-4}
8	0.0000000000	2.12×10^{-5}	3.07×10^{-5}

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n + 1$ is performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\epsilon}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy. The converged total energy is -149.2533729769 a.u.

vergence of the fixed Hessian procedures [see Eqs. (34) and (142)]. A further consideration, as shown above, is that the error vector when the fixed Hessian sequence begins should be fairly small [see Eq. (142)]. The most efficient use of fixed Hessian approaches is obtained if about $\frac{1}{3}$ of the total computer time is spent carrying out iterations in which the Hessian matrix is kept fixed with only GBT matrix elements constructed.

In the next section we examine procedures which update the initial Hessian matrix in each iteration using information contained in the gradient and step length vectors.^{9, 21} These techniques may significantly speed up convergence in the local region compared to fixed Hessian procedures. Furthermore, combined with a step size and sign control algorithm, Hessian update procedures may be very useful even when far from convergence.

VII. UPDATE METHODS

Let us assume the gradient

$$\mathbf{F}(\underline{\lambda}) = \frac{\partial E(\underline{\lambda})}{\partial \underline{\lambda}} \quad (157)$$

is known at two consecutive points ${}^k\underline{\lambda}$ and ${}^{k+1}\underline{\lambda}$ of a sequence of iterations

$${}^{k+1}\underline{\lambda} = {}^k\underline{\lambda} - \mathbf{H}_k^{-1} \mathbf{F}({}^k\underline{\lambda}) \quad (158)$$

where \mathbf{H}_k is an approximation to the Hessian matrix \mathbf{G} at $^k\underline{\lambda}$. If $\mathbf{F}(\underline{\lambda})$ is expanded around $^{k+1}\underline{\lambda}$, one obtains

$$\mathbf{F}(^k\underline{\lambda}) = \mathbf{F}(^{k+1}\underline{\lambda}) + \mathbf{G}(^{k+1}\underline{\lambda})(^k\underline{\lambda} - ^{k+1}\underline{\lambda}) + O(^k\underline{\lambda} - ^{k+1}\underline{\lambda})^2 \quad (159)$$

It is thus clear that $\mathbf{F}(^k\underline{\lambda}) - \mathbf{F}(^{k+1}\underline{\lambda})$ gives a finite difference approximation to the Hessian at $^{k+1}\underline{\lambda}$ multiplied with $^k\underline{\lambda} - ^{k+1}\underline{\lambda}$. In the most important update methods⁶³ this information about the exact Hessian is built into the Hessian approximation \mathbf{H}_{k+1} of the $(k+2)$ nd iteration of Eq. (158). A requirement to the Hessian approximation \mathbf{H}_{k+1} thus becomes that it satisfies the quasi-Newton condition⁹

$$\mathbf{H}_{k+1}[^k\underline{\lambda} - ^{k+1}\underline{\lambda}] = \mathbf{F}(^k\underline{\lambda}) - \mathbf{F}(^{k+1}\underline{\lambda}) \quad (160)$$

which implies

$$\mathbf{H}_{k+1}[^k\underline{\lambda} - ^{k+1}\underline{\lambda}] = \mathbf{G}(^{k+1}\underline{\lambda})[^k\underline{\lambda} - ^{k+1}\underline{\lambda}] + O(^k\underline{\lambda} - ^{k+1}\underline{\lambda})^2 \quad (161)$$

The projection of the approximate and the exact Hessian into the direction $^k\underline{\lambda} - ^{k+1}\underline{\lambda}$ thus differs in second order in $^k\underline{\lambda} - ^{k+1}\underline{\lambda}$.

All Hessian update methods discussed here use as the basic assumption that the quasi-Newton condition [Eq. (160)] has to be satisfied. The quasi-Newton condition contains, however, only m equations (m is the dimension of the gradient). Since the Hessian matrix has $m(m+1)/2$ unknown elements, the quasi-Newton condition therefore has to be supplemented with additional information (e.g., structural or numerical information about the exact Hessian) to determine a Hessian update. Through simulating the exact Hessian matrix numerically and structurally, Hessian update methods are partly able to obtain the superior convergence characteristics of the Newton–Raphson model without constructing the Hessian matrix explicitly.

To illustrate how the quasi-Newton condition may be supplemented with further conditions to uniquely define a Hessian update, we explicitly derive one of the most well-known update formulas, the Broyden rank-1 update.⁶⁴ The conditions which supplement the quasi-Newton condition to define some of the other well-known Hessian update methods will also be discussed, although the Broyden rank-1 formula will be the only one explicitly derived. We will thus discuss the conditions applied to determine the Powell symmetrization of the Broyden update, the Davidon–Fletcher–Powell update, and the Broyden–Fletcher–Goldfarb–Shanno update.

A. Broyden Asymmetric Rank-1 Update and Its Symmetrization

In the Broyden rank-1 update,⁶⁴ the quasi-Newton condition in Eq. (160) is supplemented with the conditions

$$\mathbf{H}_{k+1}\mathbf{P} = \mathbf{H}_k\mathbf{P} \quad (162)$$

for any vector \mathbf{P} which is orthogonal to ${}^k\underline{\lambda} - {}^{k+1}\underline{\lambda}$:

$$\mathbf{P}^T [{}^k\underline{\lambda} - {}^{k+1}\underline{\lambda}] = 0 \quad (163)$$

The motivation for supplementing the quasi-Newton condition with the conditions of Eqs. (162) and (163) is that information about the exact Jacobian at ${}^{k+1}\underline{\lambda}$ only is available in the direction ${}^k\underline{\lambda} - {}^{k+1}\underline{\lambda}$. (A Jacobian is the first derivative of a vector function. In this case the Jacobian and the Hessian are the same.) Introducing the shorthand notation

$$\mathbf{K}_k = {}^{k+1}\underline{\lambda} - {}^k\underline{\lambda} \quad (164)$$

$$\mathbf{L}_k = \mathbf{F}({}^{k+1}\underline{\lambda}) - \mathbf{F}({}^k\underline{\lambda}) \quad (165)$$

$$\mathbf{D}_{k+1} = \mathbf{H}_{k+1} - \mathbf{H}_k \quad (166)$$

we may write Eqs. (162) and (163) as

$$\mathbf{D}_{k+1} \mathbf{P} = \mathbf{0} \quad (167)$$

for

$$\mathbf{P}^T \mathbf{K}_k = 0 \quad (168)$$

Thus, \mathbf{D}_{k+1} may be written as

$$\mathbf{D}_{k+1} = \frac{\mathbf{D}'_{k+1} \mathbf{K}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \quad (169)$$

Here, \mathbf{D}'_{k+1} is so far undetermined. When we are introducing the notation of Eqs. (164)–(166), the quasi-Newton condition in Eq. (160) gives

$$\mathbf{H}_k \mathbf{K}_k + \mathbf{D}_{k+1} \mathbf{K}_k = \mathbf{L}_k \quad (170)$$

Then, by inserting Eq. (169) into Eq. (170),

$$\mathbf{D}'_{k+1} \frac{\mathbf{K}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \mathbf{K}_k = \mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k \quad (171)$$

we can identify uniquely $\mathbf{D}'_{k+1} \mathbf{K}_k$, and the updated approximate Hessian \mathbf{H}_{k+1} may therefore be written as

$$\mathbf{H}_{k+1}^B = \mathbf{H}_k + [\mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k] \frac{\mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \quad (172)$$

The implementation of the Broyden update [Eq. (172)] on a computer thus only requires simple matrix-vector multiplication.

The update correction to \mathbf{H}_k [last term in Eq. (172)] is not a symmetric matrix. When carrying out updates on an unsymmetric Jacobian, it is acceptable to get an unsymmetric update correction [Eq. (172)], whereas in updates on a Hessian matrix one may require the update correction to be symmetric. We consider now how a symmetric analog of Eq. (172) may be determined following the derivation of Powell.⁶⁵

Let \mathbf{H}_k be a symmetric approximation to the Hessian and define $\mathbf{H}_{k+1}^{(1/2)}$ as

$$\mathbf{H}_{k+1}^{(1/2)} = \mathbf{H}_k + \frac{\mathbf{T}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \quad (173)$$

where

$$\mathbf{T}_k = \mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k \quad (174)$$

Since $\mathbf{H}_{k+1}^{(1/2)}$ is not symmetric, Powell defined

$$\mathbf{H}_{k+1}^{(1)} = \frac{1}{2} [\mathbf{H}_{k+1}^{(1/2)} + \mathbf{H}_{k+1}^{(1/2)T}] \quad (175)$$

$\mathbf{H}_{k+1}^{(1)}$ does not satisfy the quasi-Newton condition, that is,

$$\mathbf{H}_{k+1}^{(1)} \mathbf{K}_k \neq \mathbf{L}_k \quad (176)$$

so it may also be updated to satisfy the quasi-Newton condition, that is,

$$\mathbf{H}_{k+1}^{(1+1/2)} = \mathbf{H}_{k+1}^{(1)} + \frac{\mathbf{T}_k^{(1)} \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \quad (177)$$

where

$$\mathbf{T}_k^{(1)} = \mathbf{L}_k - \mathbf{H}_{k+1}^{(1)} \mathbf{K}_k \quad (178)$$

Equation (177) is, however, not symmetric. A symmetrized version $\mathbf{H}_{k+1}^{(2)}$ may be determined from $\mathbf{H}_{k+1}^{(1+1/2)}$ by using an analog of Eq. (175) and so on. Powell⁶⁵ showed that the result of carrying out an infinite sequence of iterations corresponding to Eqs. (175) and (177) gives the limiting value

$$\mathbf{H}_{k+1}^P = \lim_{i \rightarrow \infty} \mathbf{H}_{k+1}^{(i)} = \mathbf{H}_k + \frac{1}{\mathbf{K}_k^T \mathbf{K}_k} \left[\mathbf{T}_k \mathbf{K}_k^T + \mathbf{K}_k \mathbf{T}_k^T - \frac{\mathbf{K}_k [\mathbf{T}_k^T \mathbf{K}_k] \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \right] \quad (179)$$

which satisfies the quasi-Newton condition, gives a symmetric update correction, and in a sense corresponds to a symmetrized version of Eq. (172). Whereas the Broyden asymmetric update formula [Eq. (172)] is of rank 1, its symmetrized analog [Eq. (179)] is of rank 2.⁹

B. The Broyden Family (DFP and BFGS Updates)

In the numerical analysis literature the Hessian update methods have primarily been used to determine minima of functions of many variables. According to the general trends of numerical analysis, the Hessian update methods that most efficiently are able to determine function minima are the Davidon–Fletcher–Powell (DFP) and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) updates (see, however, Ref. 66). The DFP and BFGS Hessian updates are members of the rank-2 Broyden update family. To understand the strengths (and weaknesses) of the DFP and BFGS update methods, we will discuss the conditions that have to be imposed on the quasi-Newton condition to define updates that belong to the Broyden family. The conditions are as follows:

1. The updated approximate Hessian matrix has to satisfy the quasi-Newton condition of Eq. (160).
2. The updated approximate Hessian matrix has to be symmetric.
3. The updated approximate Hessian matrix has to transform under linear transformations of the variables such that the step length vector of a Newton-like method [Eq. (158)] is invariant with respect to a linear transformation of the variables.
4. When the approximate Hessian matrix is positive definite at one step of the iterative procedure, then the updated Hessian matrix must also be positive definite. Neither of the previous updates [Eqs. (172) and (179)] have this property.
5. The Hessian update has to be of at most rank 2.

The Broyden family update has one parameter, which is not determined. Special choices of this parameter gives the DFP and BFGS update formulas. To get some insight into the structure of the Broyden family updates we will now discuss in more detail the contents of conditions 1–5.

1. The importance of the quasi-Newton condition for defining Hessian update methods has already been stressed. This requirement is essential for speeding up convergence rates from the linear convergence encountered in fixed Hessian procedures. For if

$$\lim_{k \rightarrow \infty} \frac{\|(\mathbf{G}^{(k+1)}\underline{\lambda}) - \mathbf{H}_{k+1})(^{k+1}\underline{\lambda} - ^k\underline{\lambda})\|}{\|^{k+1}\underline{\lambda} - ^k\underline{\lambda}\|} = 0, \quad (180)$$

the sequence $\{\underline{\lambda}^k\}$ converges superlinearly,⁶⁷ that is, $\|\underline{\lambda}^{k+1} - \underline{\lambda}^k\|/\|\underline{\lambda}^k - \underline{\lambda}^k\|$ goes toward zero.

2. The condition that the updated matrix has to be symmetric is, of course, only applicable to symmetric (Hessian) updates. When imposing the symmetry constraint on the Hessian update, we decrease the number of variables to be determined from m^2 to $m(m+1)/2$. This means, then, that more emphasis is given to the quasi-Newton condition.

3. To understand the reason for condition 3, it is instructive to examine the possible changes of the parameters $\underline{\lambda}^{k+1}$ that occur in a Newton–Raphson iteration

$$\underline{\lambda}^{k+1} = \underline{\lambda}^k - \mathbf{G}^{-1}(\underline{\lambda}^k) \mathbf{F}(\underline{\lambda}^k) \quad (181)$$

when a linear transformation is carried out among the variables

$$\underline{x} = \mathbf{A}\underline{\lambda} + \mathbf{a} \quad (182)$$

The function $E(\underline{\lambda})$ may be regarded as depending on either \underline{x} or $\underline{\lambda}$:

$$E(\underline{\lambda}) = E(\mathbf{A}^{-1}(\underline{x} - \mathbf{a})) \quad (183)$$

The partial derivatives with respect to $\underline{\lambda}$ may therefore straightforwardly be related to partial derivatives with respect to \underline{x} . Consider, for example, how the gradient with respect to the variables $\underline{\lambda}$ may be related to the gradient with respect to the variables \underline{x} :

$$F_i^{(\lambda)} = \frac{\partial E(\underline{\lambda})}{\partial \lambda_i} = (\mathbf{A}^T \mathbf{F}^{(x)})_i \quad (184)$$

Similarly, we may relate the second partial derivatives as follows:

$$\mathbf{G}^{(\lambda)} = \mathbf{A}^T \mathbf{G}^{(x)} \mathbf{A} \quad (185)$$

Thus, in the $(k+1)$ st iteration, a Newton–Raphson iteration [Eq. (181)], where \underline{x} is considered as variable, may be written [using Eqs. (182), (184), and (185)]:

$$\begin{aligned} \underline{x}^{k+1} &= \underline{x}^k - (\mathbf{G}^{(x)}(\underline{x}^k))^{-1} \mathbf{F}^{(x)}(\underline{x}^k) \\ &= \mathbf{A}^{k+1} \underline{\lambda} + \mathbf{a} \end{aligned} \quad (186)$$

From Eqs. (182) and (186) it is clear that the result of a Newton–Raphson iteration does not depend on whether the transformation is applied before

or after the iteration. If we require the update methods to possess the same invariance property with respect to a linear transformation of variables as the Newton-Raphson model, a necessary and sufficient condition on the updated matrix becomes

$$\mathbf{H}_{k+1}^{(\lambda)} = \mathbf{A}^T \mathbf{H}_{k+1}^{(x)} \mathbf{A} \quad (187)$$

If we assume that the Hessian approximation of the k th iteration \mathbf{H}_k possesses the invariance property, then the Hessian update matrix has to satisfy

$$\mathbf{H}_{k+1}^{(\lambda)} - \mathbf{H}_k^{(\lambda)} = \mathbf{A}^T (\mathbf{H}_{k+1}^{(x)} - \mathbf{H}_k^{(x)}) \mathbf{A} \quad (188)$$

to keep the invariance property.

4. This condition (i.e., that if the approximate Hessian matrix at one step of the iterative procedure is positive definite then the updated approximate Hessian has to be positive definite) is a useful property of an update when the update is used to determine function minima, as it ensures that the Hessian approximation will have the correct structure during the whole sequence of iterations when the structure is correct at one single step of the iterative procedure. In connection with the use of update methods for obtaining convergence in MCSCF calculations, this positive definite requirement is less important when the Hessian update method in MCSCF is used to determine excited states that have a nonpositive definite Hessian matrix.

5. The limitation that $\mathbf{H}_{k+1} - \mathbf{H}_k$ be of rank 2 complements the previously described constraints to define uniquely the Broyden family update. A rank-2 update correction may generally be written as

$$\mathbf{H}_{k+1} - \mathbf{H}_k = a\mathbf{UV}^T + b\mathbf{VU}^T + c\mathbf{VV}^T + d\mathbf{UU}^T \quad (189)$$

where a , b , c , and d are constants and \mathbf{U} and \mathbf{V} are vectors. Choosing the vectors \mathbf{U} and \mathbf{V} as

$$\mathbf{U} = \mathbf{H}_k \mathbf{K}_k; \quad \mathbf{V} = \mathbf{L} = \mathbf{F}({}^{k+1}\underline{\lambda}) - \mathbf{F}({}^k\underline{\lambda}) \quad (190)$$

the Hessian update correction in Eq. (189) satisfies the invariance condition (i.e., condition 3). Conditions 1 and 2 may then be used to define the Broyden family update as

$$\mathbf{H}_{k+1} = \mathbf{H}_{k+1}^0 + a'\mathbf{W}_k \mathbf{W}_k^T \quad (191)$$

where

$$\mathbf{W}_k = \frac{\mathbf{H}_k \mathbf{K}_k}{\mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k} - \frac{\mathbf{L}_k}{\mathbf{K}_k^T \mathbf{L}_k} \quad (192)$$

and

$$\mathbf{H}_{k+1}^0 = \mathbf{H}_k + \frac{\mathbf{L}_k \mathbf{L}_k^T}{\mathbf{L}_k^T \mathbf{K}_k} - \frac{\mathbf{H}_k \mathbf{K}_k \mathbf{K}_k^T \mathbf{H}_k}{\mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k} \quad (193)$$

where a' is an undefined parameter. [See Ref. 9, where a detailed derivation of Eqs. (191)–(193) is carried out.]

The conservation of positive definiteness of the approximate Hessian may be used to limit the interval of allowed a' parameters. In Ref. 9 it is shown how a positive definiteness requirement limits the interval of the parameter a' to be $[a'_{\min}, \infty]$, where a'_{\min} is the largest negative number for which $\mathbf{H}_{k+1}^0 + a'_{\min} \mathbf{W}_k \mathbf{W}_k^T$ is singular. It should be mentioned that to limit the interval of a' it is assumed that $\mathbf{K}_k^T \mathbf{L}_k > 0$. Which value of a' to use within the allowed interval must to a high degree rely on numerical experience. While members of the Broyden family behave identically when exact line searches are used,⁶⁸ differences are encountered with direct application of Eq. (158) (no line search). Extensive numerical tests are compiled in Ref. 69. The two most commonly used choices are the Broyden–Fletcher–Goldfarb–Shanno (BFGS) update⁷⁰

$$a' = 0$$

$$\mathbf{H}_{k+1}^{\text{BFGS}} = \mathbf{H}_{k+1}^0 = \mathbf{H}_k + \frac{\mathbf{L}_k \mathbf{L}_k^T}{\mathbf{L}_k^T \mathbf{K}_k} - \frac{\mathbf{H}_k \mathbf{K}_k \mathbf{K}_k^T \mathbf{H}_k}{\mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k} \quad (194)$$

and the Davidon–Fletcher–Powell (DFP) update⁷¹

$$a' = a'_{\text{DFP}} = \mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k$$

$$\mathbf{H}_{k+1}^{\text{DFP}} = \mathbf{H}_k + \left[1 + \frac{\mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k}{\mathbf{K}_k^T \mathbf{L}_k} \right] \frac{\mathbf{L}_k \mathbf{L}_k^T}{\mathbf{L}_k^T \mathbf{K}_k} - \frac{\mathbf{H}_k \mathbf{K}_k \mathbf{L}_k^T + \mathbf{L}_k \mathbf{K}_k \mathbf{H}_k}{\mathbf{L}_k^T \mathbf{K}_k} \quad (195)$$

A rank-1 procedure can also be generated from Eq. (192):⁷²

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{(\mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k)(\mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k)^T}{(\mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k)^T \mathbf{K}_k} \quad (196)$$

This procedure has some disadvantages compared to, say, the DFP and BFGS updates: it does not retain positive definiteness, and it can be rather unstable. One theoretical advantage of the update (196) is that a quadratic function is minimized in at most $m + 1$ iterations under reasonable condi-

tions. Furthermore, the change is given by one vector and one scalar, which is simplifying for very large dimensions. The rank-1 update [Eq. (196)] has been used in MCSCF optimization by Eade and Robb.⁷³ Their implementation is, however, not optimal since they do not use the quasi-Newton condition correctly owing to neglect of changes of CI coefficients. The actual performance of the update [Eq. (196)] reported by Eade and Robb is, however, very encouraging.

C. A Variational Approach to Update Procedures

The update methods mentioned hitherto can be generated by minimizing a norm of the change $\mathbf{H}_{k+1} - \mathbf{H}_k$.⁷⁴ This point of view elucidates the relation between the different updates and thus provides help in choosing the best update for a given situation. We will therefore discuss the updates in this perspective. To do this, we define a general weighted norm

$$\|\mathbf{W}\|_{\mathbf{A}} = \|\mathbf{AWA}\| \quad (197)$$

where $\|\cdot\|$ is the Frobenius norm. Further, let \mathbf{U}_{k+1} (\mathbf{U}_{k+1}^S) be all matrices (symmetric matrices) that satisfy the quasi-Newton condition of iteration $k + 2$.

The Broyden rank-1 update \mathbf{H}_{k+1}^B minimizes $\|\mathbf{H}'_{k+1} - \mathbf{H}_k\|_{\mathbf{I}}$, where \mathbf{H}'_{k+1} belongs to \mathbf{U}_{k+1} . This is easily seen since for any \mathbf{H}'_{k+1}

$$\begin{aligned} \|\mathbf{H}_{k+1}^B - \mathbf{H}_k\|_{\mathbf{I}} &= \left\| (\mathbf{L}_k - \mathbf{H}_k \mathbf{K}_k) \frac{\mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \right\|_{\mathbf{I}} \\ &= \left\| (\mathbf{H}'_{k+1} - \mathbf{H}_k) \frac{\mathbf{K}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{K}_k} \right\|_{\mathbf{I}} \\ &\leq \|\mathbf{H}'_{k+1} - \mathbf{H}_k\|_{\mathbf{I}} \end{aligned} \quad (198)$$

where \mathbf{I} represents the unit matrix (i.e., $\mathbf{A} = \mathbf{I}$). When it is essential to minimize the Frobenius norm of the change $\mathbf{H}_{k+1} - \mathbf{H}_k$, the Broyden rank-1 method seems an obvious choice. If $\|\mathbf{H}'_{k+1} - \mathbf{H}_k\|_{\mathbf{I}}$ is minimized with \mathbf{H}'_{k+1} restricted to \mathbf{U}_{k+1}^S , the Powell symmetrization of Broyden's rank-1 method is obtained.⁷⁴ This is shown by observing that $\|(\mathbf{H}'_{k+1} - \mathbf{H}_k)\mathbf{V}\|_{\mathbf{I}}$ is smaller than $\|(\mathbf{H}'_{k+1} - \mathbf{H}_k)\mathbf{V}\|_{\mathbf{I}}$ for all choices of \mathbf{V} (\mathbf{V} is any arbitrary vector).

When a pattern in $\mathbf{H}_k - \mathbf{G}^{(k+1)}\underline{\lambda}$ exists, it can be advantageous to incorporate this by choosing a matrix \mathbf{A} that reflects this pattern and then minimize $\|\mathbf{H}'_{k+1} - \mathbf{H}_k\|_{\mathbf{A}}$ to obtain \mathbf{H}_{k+1} . If $\mathbf{H}_k - \mathbf{G}^{(k+1)}\underline{\lambda}$ is expected to be proportional to $G^{(k+1)}\underline{\lambda}_{ij}$, it seems sound to weight $(\mathbf{H}_{k+1} - \mathbf{H}_k)$ so that large elements are allowed to change more than small elements. In cases where $\mathbf{G}^{(k+1)}\underline{\lambda}$ is positive definite a proper choice of \mathbf{A} is then $(\mathbf{G}^{(k+1)}\underline{\lambda})^{-1/2}$. The minimizer of $\|\mathbf{H}'_{k+1} - \mathbf{H}_k\|_{(\mathbf{G}^{(k+1)}\underline{\lambda})^{-1/2}}$ with \mathbf{H}'_{k+1} restricted to \mathbf{U}_{k+1}^S can be shown to be $\mathbf{H}_{k+1}^{\text{DFP}}$. In the same way the minimizer with \mathbf{H}'_{k+1} restricted to \mathbf{U}_{k+1}^S of $\|\mathbf{H}'_{k+1} - \mathbf{H}_k^{-1}\|_{(\mathbf{G}^{(k+1)}\underline{\lambda})^{1/2}}$ is the BFGS update. The BFGS and DFP updates thus seem appropriate when the relative errors in \mathbf{H}_{k+1} (compared to $\mathbf{G}^{(k+1)}\underline{\lambda}$) are of the same magnitude and when $\mathbf{G}^{(k+1)}\underline{\lambda}$ is positive definite.

D. Inverse Updates

In the previous sections updates were carried out directly on the Hessian (Jacobian). The step length parameters ${}^{k+1}\underline{\lambda}$ were then determined from ${}^k\underline{\lambda}$, $\mathbf{F}({}^k\underline{\lambda})$ and \mathbf{H}_k^{-1} as

$${}^{k+1}\underline{\lambda} = {}^k\underline{\lambda} - \mathbf{H}_k^{-1} \mathbf{F}({}^k\underline{\lambda}) \quad (199)$$

which corresponds to solving a set of linear inhomogenous equations

$$\mathbf{H}_k({}^{k+1}\underline{\lambda} - {}^k\underline{\lambda}) = -\mathbf{F}({}^k\underline{\lambda}) \quad (200)$$

The solution of Eq. (199) for very large dimensions is a very time-consuming part of each iteration. It therefore becomes worthwhile to examine whether the update corrections can be carried out directly on the inverse matrix:

$$\mathbf{H}_{k+1}^{-1} = \mathbf{H}_k^{-1} + (\mathbf{H}')^{-1} \quad (201)$$

That is, can $(\mathbf{H}')^{-1}$ be determined directly. If \mathbf{H}_k^{-1} is available ${}^{k+1}\underline{\lambda}$ can be constructed directly by a multiplication of a matrix with a vector [see Eq. (199)] which is much faster than solving a set of linear inhomogeneous equations. The inverse update formulas may easily be determined from the equation

$$\mathbf{H}_{k+1}(\mathbf{H}_k^{-1} + (\mathbf{H}')^{-1}) = \mathbf{1} \quad (202)$$

and corresponding to the update formulas of Eqs. (172), (179), (194), and

(195) we obtain

$$(\mathbf{H}_{k+1}^B)^{-1} = \mathbf{H}_k^{-1} - (\mathbf{H}_k^{-1}\mathbf{L}_k - \mathbf{K}_k) \frac{\mathbf{K}_k^T \mathbf{H}_k^{-1}}{\mathbf{K}_k^T \mathbf{H}_k^{-1} \mathbf{L}_k} \quad (203)$$

$$\begin{aligned} (\mathbf{H}_{k+1}^P) &= \mathbf{H}_k^{-1} + \frac{1}{\mathbf{K}_k^T \mathbf{K}_k} \left\{ (\mathbf{L}_k - \mathbf{H}_k^{-1} \mathbf{K}_k) \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{L}_k^T - \mathbf{K}_k^T \mathbf{H}_k^{-1}) \right. \\ &\quad \left. - \frac{1}{\mathbf{K}_k^T \mathbf{K}_k} [\mathbf{K}_k^T \mathbf{L}_k - \mathbf{K}_k^T \mathbf{H}_k^{-1} \mathbf{K}_k] \mathbf{K}_k \mathbf{K}_k^T \right\} \end{aligned} \quad (204)$$

$$\begin{aligned} (\mathbf{H}_{k+1}^{BFGS})^{-1} &= \mathbf{H}_k^{-1} + \left(1 + \frac{\mathbf{L}_k^T \mathbf{H}_k^{-1} \mathbf{L}_k}{\mathbf{K}_k^T \mathbf{L}_k} \right) \frac{\mathbf{K}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{L}_k} - \frac{\mathbf{K}_k \mathbf{L}_k^T \mathbf{H}_k^{-1} + \mathbf{H}_k^{-1} \mathbf{L}_k^T \mathbf{K}_k}{\mathbf{K}_k^T \mathbf{L}_k} \end{aligned} \quad (205)$$

$$(\mathbf{H}_{k+1}^{DFP})^{-1} = \mathbf{H}_k^{-1} - \frac{\mathbf{H}_k^{-1} \mathbf{L}_k \mathbf{L}_k^T \mathbf{H}_k^{-1}}{\mathbf{L}_k^T \mathbf{H}_k^{-1} \mathbf{L}_k} + \frac{\mathbf{K}_k \mathbf{K}_k^T}{\mathbf{K}_k^T \mathbf{L}_k} \quad (206)$$

It is observed that $(\mathbf{H}_{k+1}^{BFGS})^{-1}$ is obtained from \mathbf{H}_{k+1}^{DFP} when the substitutions

$$\mathbf{H}_k \rightarrow \mathbf{H}_k^{-1}; \quad \mathbf{L}_k \rightarrow \mathbf{K}_k; \quad \mathbf{K}_k \rightarrow \mathbf{L}_k \quad (207)$$

are carried out. Similarly, \mathbf{H}_{k+1}^{BFGS} is related to $(\mathbf{H}_{k+1}^{DFP})^{-1}$ through the same substitution. The DFP and BFGS update corrections are said to be dual.⁹

E. Convergence Characteristics of Hessian Update Methods

The practical implementation of a Hessian update method consists of carrying out a sequence of iterations similar to the sequence of the linear fixed Hessian approach [Eq. (138)] but where the Hessian matrix gets updated in each iteration. If the Hessian matrix were actually constructed and used in each iteration we would perform just sequence of linear Newton–Raphson iterations [Eq. (20)]. When a sequence of N linear fixed Hessian iterations is carried out the total order of convergence in ${}^0\mathbf{e}$ is $N + 1$, whereas when N linear Newton–Raphson iterations are carried out we obtain a total order of convergence of 2^N . Since the Hessian update corrections are determined to simulate the exact Hessian matrix, a sequence of N update iterations is expected to have a total order of convergence in between $N + 1$ and 2^N depending on the efficiency of the update method used to simulate the exact Hessian matrix. The number of fixed Hessian iterations performed in the local region to obtain convergence is usually small (~ 10). It is therefore not realistic to expect that the Hessian updates will be very accurate and will well

simulate the exact Hessian. Since the quasi-Newton condition gives information about the exact Hessian in directions corresponding to actual step lengths, it may still be expected that Hessian updates can considerably improve the structure of the part of the Hessian matrix that corresponds to variables that change significantly (i.e., that are far from converged and thereby improve considerably the convergence properties of a fixed Hessian procedure). In the next subsection we demonstrate through numerical examples that the update methods constitute a very efficient way for obtaining local convergence of an MCSCF calculation. The Hessian update methods can also be shown to converge superlinearly.⁶⁷

The convergence problems that appear in a sequence of step size and sign controlled Newton–Raphson iterations are associated with large fluctuations in a few modes and are caused by small eigenvalues of the Hessian matrix (see Section V.B). Because the Hessian update methods update the Hessian matrix corresponding to the directions (modes) that give most convergence problems, we expect the Hessian update methods to simulate to a certain degree the global convergence characteristics of the Newton–Raphson approach. In the next subsection we return to this question and demonstrate through numerical examples how the update methods may have very promising global convergence properties.

The performance of the various update methods cannot be expected to be identical. Neither can it be expected that a single method is superior in all cases since optimization in MCSCF is carried out both to minima and saddle points. The previous subsections have provided some information about the usefulness of the different updates.

In ground state calculations the DFP and BFGS updates are appealing since they usually guarantee retention of positive definite approximations to the Hessian. However, the weighting matrix $(G^{(k+1)}\lambda)^{-1/2}$ used in obtaining the DFP variationally does not reflect an optimal weighting, since the relative errors in the Hessian approximations are not of the same magnitude. As described in Section V the smallest eigenvalues of the Hessian may change significantly between two points of iteration whereas larger eigenvalues are more stable. So if an exact Hessian from one point is used as the starting approximation to the Hessian at another point, large elements are approximated relatively better than small elements. This is in contrast with the assumption of using a weighting matrix $G^{(k+1)}\lambda^{-1/2}$. The updates based on minimization with the neutral norm $\|\cdot\|_I$ can thus not be entirely disregarded, although they may not retain positive definite matrices.

For optimization to saddle points the DFP and BFGS updates have no theoretical advantages, and the fact that H_{k+1}^P and H_{k+1}^B minimizes the neutral norm $\|\cdot\|_I$ can be of importance. Further, the denominators in H_{k+1}^P and H_{k+1}^B [Eqs. (172) and (179)] are the square of the step length, which remains

strictly positive. The denominators in $\mathbf{H}_{k+1}^{\text{BFGS}}$ and \mathbf{H}^{DFP} [Eqs. (194) and (195)] contain $\mathbf{L}_k^T \mathbf{K}_k$ and $\mathbf{K}_k^T \mathbf{H}_k \mathbf{K}_k$. They can become near-singular. Therefore, it is expected that the Broyden rank-1 and its symmetrized form can be used with advantage when converging to a saddle point.

F. Numerical Results

To illustrate how update methods work in actual calculations we will now describe calculations carried out on the $B^3\Sigma_u^-$ state at 2.13 a.u. and on the $E^3\Sigma_u^-$ state at 2.13 and 2.10 a.u. To examine the methods for ground state calculations we initially consider the $B^3\Sigma_u^-$ calculation at 2.13 a.u.

In Table XXI we report the result of a Broyden rank-1 calculation. The initial Hessian for both the update and the fixed Hessian approaches is from after iteration 2 (at iteration point 2) of the Newton–Raphson calculation in Table III. Updates begin in the subsequent iteration. The update sequence converges, while a series of linear fixed Hessian iterations with the same initial guess of orbitals diverges. This indicates the importance of updating or recalculating the Hessian matrix during the iterative procedure. Thus, the region of convergence is extended by the update of the Hessian matrix.

To investigate the region of convergence in more detail we describe some update calculations that are initiated at previous steps of the Newton–Raphson series of iterations in Table III. Step size and sign control were carried out in these steps of the Newton–Raphson calculation. Step size and sign control were implemented in the update calculations in a way similar to that of the Newton–Raphson approach (see Section V). Step size and sign control is usually carried out on symmetric matrices. The Broyden rank-1 update gives an unsymmetric Hessian approximation, so the update calculations we initially report for obtaining global convergence will be using the Powell symmetrization of the Broyden rank-1 update (the BFGS and DFP update methods will be discussed later).

A Powell update calculation is reported in Table XXII, which uses an initial Hessian calculated after the first iteration (at iteration point 1) of Table III. The Powell sequence of iterations in Table XXII converges. In the initializing Newton–Raphson iteration a step length amplitude of 37.2 was constrained, and in the subsequent Powell iteration a step length amplitude was constrained. The Powell sequence of iterations fluctuates in the initial six iterations, and the total energy at iteration point 5 is farther away from the converged total energy than the total energy at iteration point 2. During the iterations where fluctuations in the total energy occur, the approximation to the Hessian matrix gradually improves. After iteration point 5 the step sizes are accurate enough to ensure rapid convergence. The gradual improvement of the Hessian approximation is reflected in a stabilization of the smallest eigenvalues of the Hessian approximation. To illustrate this point,

TABLE XXI
Fixed Hessian and Broyden Rank-1 Update Which Use an Initial \mathbf{G} from Iteration Point 2 of Table III ($B^3\Sigma_u^-$ of O₂ at 2.13 a.u.)

Iteration point ^b	Broyden rank-1 update ^a				Fixed Hessian ^a			
	$E - E^{\text{CONV}}^c$	$\ \mathbf{F}\ $	$\ \lambda^{n+1} - \lambda^n\ \sim \ \epsilon\ $		$E - E^{\text{CONV}}^c$	$\ \mathbf{F}\ $	$\ \lambda^{n+1} - \lambda^n\ \sim \ \epsilon\ $	
0	0.0055035052	3.62×10^{-2}	3.35×10^{-1} ^d		0.005503502	3.62×10^{-2}	3.35×10^{-1} ^d	
1	0.0009142549	3.42×10^{-2}	1.73×10^{-1}		0.0009142549	3.42×10^{-2}	2.72×10^{-1}	
2	0.0008941631	2.58×10^{-2}	7.96×10^{-2}		0.0031811111	3.54×10^{-2}	4.34×10^{-1}	
3	0.0000767315	1.03×10^{-2}	2.96×10^{-2}		0.0055891388	7.96×10^{-2}	1.03	
4	0.0000036257	1.57×10^{-3}	6.04×10^{-3}		0.0388891594	2.64×10^{-1}	2.33	
5	0.0000000457	2.20×10^{-4}	8.71×10^{-4}		0.1468562383	1.56×10^0	4.42×10^{-1}	
6	0.0000000064	5.88×10^{-5}	4.75×10^{-4}					
7	0.0000000000	4.41×10^{-6}	3.26×10^{-5}					

^aThe first iteration listed is the Newton-Raphson iteration. Updates begin at iteration point 1. The fixed Hessian approach uses \mathbf{G} from iteration point 0.

^bAt iteration point n , the energy and \mathbf{F} are evaluated and λ^{n+1} is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|\lambda^{n+1} - \lambda^n\|$ is an approximation to $\|\epsilon\| = \|\lambda - \alpha\|$ [see Eq. (71)].

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy – 149.3086149361 a.u.

^dThe step size and sign algorithm¹⁷ is applied in this iteration.

TABLE XXII
Powell Update Which Uses an Initial G from Iteration Point 1 of
Table III ($B^3\Sigma_u^-$ of O₂ at 2.13 a.u.).

Iteration point ^a	$E - E^{\text{CONV}}$	$\ \mathbf{F}\ $	$\ \lambda^{n+1} - \lambda^n\ \sim \ \mathbf{e}\ $
0 ^b	0.0246933274	2.18×10^{-1}	3.17×10^{-1} ^d
1	0.0055035052	3.62×10^{-2}	2.44×10^{-1}
2	0.0006525723	2.40×10^{-2}	1.80×10^{-1}
3	0.0003031638	2.98×10^{-2}	4.10×10^{-2}
4	0.0000577782	7.78×10^{-3}	2.27×10^{-1}
5	0.0013706643	2.60×10^{-2}	2.04×10^{-1}
6	0.0000027061	1.84×10^{-3}	7.64×10^{-3}
7	0.0000004323	1.11×10^{-3}	1.96×10^{-3}
8	0.0000000258	1.35×10^{-4}	6.14×10^{-4}
9	0.0000000865	6.56×10^{-5}	3.52×10^{-4}
10	0.0000000008	2.42×10^{-5}	5.85×10^{-5}
11	0.0000000002	9.06×10^{-6}	3.96×10^{-5}
12	0.0000000000	4.00×10^{-6}	4.31×10^{-5}

^aAt iteration point n , the energy and \mathbf{F} are evaluated and λ^{n+1} is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|\lambda^{n+1} - \lambda^n\|$ is an approximation to $\|\mathbf{e}\| = \|\lambda - \alpha\|$ [see Eq. (71)].

^bInitializing constrained Newton-Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy = 149.3086149361 a.u.

^dThe step size and sign algorithm¹⁷ is applied in this iteration.

in Table XXIII we report the development of the lowest four eigenvalues of the approximation Hessian matrix during the sequence of iterations of Table XXII. The four lowest eigenvalues of the exact Hessian matrix at the converged point are 0.0442, 0.1556, 0.1860, and 0.2232 (see Table V).

It may be seen in Table XXIII that the lowest eigenvalues fluctuate a little in the initial iterations. As a general trend, however, the smallest eigenvalues of the approximate Hessian matrix are rapidly improved during the iterative procedure, and close agreement is observed between the four lowest eigenvalues of the converged approximate Hessian matrix and the exact Hessian matrix. The gradual decrease in $E - E^{\text{CONV}}$, which is interrupted at iteration point 5, can be ascribed to the fact that the approximate Hessian matrix at iteration point 4 has the smallest eigenvalue (0.0082). In all previous and subsequent iterations the smallest eigenvalue of the approximate Hessian is 5–10 times as large as the smallest eigenvalue at iteration point 4. This small eigenvalue at iteration point 4 disappears after the Hessian update at iteration point 5 is carried out.

TABLE XXIII
 Lowest Four Eigenvalues of the Hessian Update Matrix in the
 Powell Update Sequence of Iterations in Table XXII
 $(B^3\Sigma_u^-$ of O₂ at 2.13 a.u.)

Iteration point ^a	Lowest eigenvalues			
0	0.0010	0.1134	0.2878	0.3226
1	0.0240	0.0862	0.2860	0.3262
2	0.0506	0.1056	0.2860	0.3256
3	0.0448	0.1012	0.2848	0.3272
4	0.0082	0.1204	0.2776	0.3302
5	0.0522	0.1364	0.2754	0.3302
6	0.0548	0.1336	0.2756	0.3302
7	0.0382	0.1780	0.2751	0.3302
8	0.0578	0.1760	0.2710	0.3306
9	0.0434	0.1756	0.2560	0.3314
10	0.0544	0.1780	0.2500	0.3308
11	0.0884	0.1541	0.1798	0.3362
12	0.0390	0.1622	0.1798	0.3340
13	0.0434	0.1622	0.1812	0.3334

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $n+1\lambda$ is determined. Thus at iteration point n , iteration $n+1$ is performed.

The convergence characteristics of a Hessian update (as well as a Newton-Raphson) MCSCF calculation are very strongly affected by the modes corresponding to the small eigenvalues of the Hessian matrix. The small eigenvalues of the Hessian matrix get substantially improved when updates are carried out. The larger eigenvalues are not affected much by the update; for example, the lowest seven to nine eigenvalues at iteration point 0 are 0.7874, 1.0858, 1.8206, respectively, whereas at iteration point 13 they are 0.8022, 1.1780, and 1.8284. The corresponding converged Hessian eigenvalues (nonlinear transformation) are 0.9046, 1.0536, and 1.2128 (see Table V).

Thus the Powell update is able to converge from a region where it is necessary to restrict step sizes in both the initializing Newton-Raphson iteration and in the consecutive sequence of Powell update iterations. When the Powell sequence of iterations are started out at one previous iteration of the Newton-Raphson procedure (the initial \mathbf{G} is from iteration point 0 of Table III, and updates start on the subsequent iteration), the Powell sequence of iterations is found not to converge. The initializing Hessian matrix is structurally incorrect at this step of the Newton-Raphson sequence, and in the current implementation the Powell sequence of iterations is not able to improve the structure of the Hessian matrix enough to make the Powell se-

quence of iterations converge. (See Ref. 21 for a more thorough discussion of this matter.)

The BFGS and DFP methods have also been applied, using as an initial \mathbf{G} the Newton-Raphson calculation at iteration point 1 of Table III. The BFGS and DFP calculations are reported in Ref. 21 and have similar convergence characteristics as those obtained for the Powell calculation of Table XXII. The DFP and BFGS calculations converge in 10 and 12 iterations, respectively, whereas the Powell sequence of iterations converges in 12 iterations. Calculations which further investigate the global convergence characteristics of BFGS and DFP updates on the lowest state of a given symmetry are currently being carried out.⁷⁵

After having discussed the region of convergence of update methods in ground state calculations, we now carry out a corresponding investigation for excited states. To do so, we describe Powell and Broyden rank-1 update calculations on the $E^3\Sigma_u^-$ state of O_2 at the internuclear distances of 2.13 and 2.10 a.u. The Newton-Raphson calculation at 2.13 a.u. is a very simple MCSCF calculation and requires only five Newton-Raphson iterations to converge (see Table X). The Newton-Raphson calculation at 2.10 a.u. is more complicated and requires eight iterations to converge (see Table IV). The difference in the complexity of the two calculations can be ascribed to the fact that the converged Hessian matrix at 2.13 a.u. has an eigenvalue at convergence of 6.0×10^{-2} (see Table XI) whereas the smallest Hessian eigenvalue (in magnitude) at convergence is 1.4×10^{-3} at 2.10 a.u. (see Table VIII).

In Table XXIV we report the Broyden rank-1 and the fixed Hessian calculation at 2.13 a.u. The fixed Hessian sequence of iterations diverges, while the Broyden rank-1 calculation converges. The Broyden rank-1 calculation converges in 11 iterations, while the Newton-Raphson iteration converges in 5 iterations. The convergence characteristics of the Broyden calculation of Table XXIV show great similarity with the Powell calculation of Table XXII. In the initial 6 Broyden iterations in Table XXIV the calculation does not show a definite convergence trend. At iteration point 3, $E - E^{\text{CONV}}$ is 10 times as large as the corresponding value of iteration point 1. The energy difference $E - E^{\text{CONV}}$ at iteration point 6 is also of the same magnitude as the value at iteration point 1. After iteration point 6, the Broyden calculation converges rapidly and monotonically. During the initial sequence of iterations the approximative Hessian matrix gains more and more similarity with the exact Hessian matrix. After iteration point 6, steps which bring the calculation to converge rapidly can be taken. Since the fixed Hessian calculation diverges, the region of convergence is thus extended by updating the Hessian matrix.

To further examine the extension in the region of convergence we now describe some calculations on the $E^3\Sigma_u^-$ state of O_2 at the internuclear dis-

tance of 2.10 a.u. The Newton-Raphson calculation is given in Table IV. Step sizes were controlled in the initial three Newton-Raphson iterations (see Table VIII). The Broyden rank-1 update calculations initiated using \mathbf{G} at Newton-Raphson iteration points 4 and 3 of Table IV are reported in Tables XXV and XXVI. The Broyden rank-1 calculation that has been started out closest to convergence (Table XXV) converges rapidly and monotonically. When the Broyden calculations have been started farther from convergence (Table XXVI), $E - E^{\text{CONV}}$ very slowly improves during the initial iterations and then converges rapidly and monotonically. A Powell update sequence has also been initiated using \mathbf{G} from Newton-Raphson iteration point 2 of Table IV. Here, the Powell update calculation is not converging, and it appears that the region of convergence of the Powell updates are smaller for excited states than for ground state calculations—in particular when the converged Hessian matrix has very small eigenvalues.

In order to compare the region of convergence of the Broyden rank-1 update methods with the region of convergence of the DFP and BFGS update methods when applied on excited states, the calculations on the $E^3\Sigma_u^-$ state at 2.10 a.u. were repeated using BFGS and DFP updates. The Broyden rank-1 update calculation converges when the update is initialized using \mathbf{G}

TABLE XXIV

Convergence Characteristics of the Broyden Rank-1 Update and Linear Fixed Hessian Calculations of the $E^3\Sigma_u^-$ State of O_2 at 2.13 a.u. Using an Initial \mathbf{G} from Iteration Point 0 of Table X

Iteration point ^a	Broyden rank-1			Fixed Hessian		
	$E - E^{\text{CONV}}_c$	$\ \mathbf{F}\ $	$\ {}^{n+1}\lambda - {}^n\lambda\ \sim \ {}^n\mathbf{e}\ $	$E - E^{\text{CONV}}_c$	$\ \mathbf{F}\ $	$\ {}^{n+1}\lambda - {}^n\lambda\ \sim \ {}^n\mathbf{e}\ $
0 ^b	0.0590242739	9.72×10^{-1}	4.62×10^{-1} ^b	0.0590242739	9.72×10^{-2}	4.62×10^{-1} ^d
1	-0.0003831946	8.36×10^{-2}	1.37×10^{-1}	-0.0003831946	8.36×10^{-2}	1.45×10^{-1}
2	-0.0002434388	4.66×10^{-2}	2.92×10^{-1}	0.0000353020	5.16×10^{-2}	3.50×10^{-1}
3	0.0038464235	1.33×10^{-1}	2.18×10^{-1}	0.0068589871	1.90×10^{-1}	1.34
4	0.0002185597	5.06×10^{-2}	1.41×10^{-1}	0.0522554397	4.08×10^{-1}	1.34
5	0.0001156611	2.04×10^{-2}	1.97×10^{-1}	0.0400059992	5.10×10^{-1}	3.64
6	0.0001917947	3.04×10^{-2}	1.29×10^{-1}	0.2462884409	1.57×10^0	5.32
7	0.0000104537	4.62×10^{-2}	1.02×10^{-2}			
8	0.0000025132	2.96×10^{-2}	8.69×10^{-3}			
9	0.0000014093	1.95×10^{-3}	4.26×10^{-3}			
10	0.0000000010	8.38×10^{-5}	2.41×10^{-4}			
11	0.0000000000	4.64×10^{-5}	9.19×10^{-5}			

^aAt iteration point n , the energy and \mathbf{F} are evaluated and ${}^{n+1}\lambda$ is determined. Thus at iteration point n , iteration $n + 1$ is performed. $\|{}^{n+1}\lambda - {}^n\lambda\|$ is an approximation to $\|{}^n\mathbf{e}\| = \|{}^n\lambda - \mathbf{a}\|$ [see Eq. (71)].

^bInitializing constrained Newton-Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy = 149.2533729769 a.u.

^dThe step size and sign control algorithm¹⁷ is applied in this iteration.

TABLE XXV
Broyden Rank-1 Update After Newton-Raphson Iteration Point 4 of
Table IV ($E^3\Sigma_u^-$ of O₂ at 2.10 a.u.): Initial G is from Iteration Point 4
of Table IV

Iteration point ^a	$E - E^{\text{CONV}}$ ^c	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\alpha}\ $
0 ^b	0.0000311601	2.06×10^{-3}	2.66×10^{-1}
1	0.0000073514	4.34×10^{-3}	1.06×10^{-1}
2	0.0000010598	8.14×10^{-4}	2.79×10^{-2}
3	0.0000000067	2.82×10^{-4}	5.76×10^{-3}
4	0.0000000005	1.17×10^{-4}	1.98×10^{-3}
5	0.0000000000	1.97×10^{-5}	2.79×10^{-4}

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\alpha}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bInitializing Newton-Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy – 149.2781477108 a.u.

TABLE XXVI
Broyden Rank-1 Update After Newton-Raphson Iteration Point 3 of
Table IV ($E^3\Sigma_u^-$ of O₂ at 2.10 a.u.): Initial G is from Iteration Point 3
of Table IV

Iteration point ^a	$E - E^{\text{CONV}}$ ^c	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ \sim \ \underline{\alpha}\ $
0 ^b	0.0000947284	1.91×10^{-1}	2.76×10^{-1}
1	0.0000311601	2.06×10^{-2}	2.84×10^{-1}
2	0.0000141971	1.82×10^{-2}	1.98×10^{-1}
3	0.0000034206	2.76×10^{-3}	5.55×10^{-2}
4	0.0000002208	3.60×10^{-4}	9.39×10^{-3}
5	0.0000000165	2.56×10^{-4}	4.47×10^{-3}
6	0.0000000084	9.42×10^{-5}	3.45×10^{-3}

^aAt iteration point n , the energy and \mathbf{F} are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|\underline{\lambda}^{n+1} - \underline{\lambda}^n\|$ is an approximation to $\|\underline{\alpha}\| = \|\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bInitializing Newton-Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at this step and the converged total energy – 149.2781477108 a.u.

TABLE XXVII

BFGS and DFP Update After Newton–Raphson Iteration Point 4 of Table IV ($E^3\Sigma_u^-$ of O₂ at 2.10 a.u.): Initial G is from Iteration Point 4 of Table IV

Iteration point ^a	BFGS			DFP		
	$E - E^{\text{CONV}c}$	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\ \sim \ {}^n\mathbf{e}\ $	$E - E^{\text{CONV}c}$	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\ \sim \ {}^n\mathbf{e}\ $
0 ^b	0.0000311601	2.06×10^{-3}	2.66×10^{-1}	0.0000311601	2.06×10^{-3}	2.66×10^{-1}
1	0.0000073514	4.34×10^{-3}	9.62×10^{-2}	0.0000073514	4.34×10^{-3}	9.61×10^{-2}
2	0.0000004867	7.34×10^{-4}	2.06×10^{-2}	0.0000004674	6.30×10^{-4}	2.00×10^{-2}
3	0.0000000085	4.12×10^{-4}	1.31×10^{-2}	0.0000000006	3.82×10^{-4}	1.93×10^{-2}
4	0.0000000651	7.34×10^{-4}	9.62×10^{-3}	0.0000001973	1.26×10^{-3}	1.69×10^{-2}
5	0.0000000017	3.50×10^{-5}	8.22×10^{-4}	0.0000000000	5.72×10^{-6}	2.87×10^{-4}
6	0.0000000002	1.88×10^{-5}	5.58×10^{-4}			
7	0.0000000000	3.16×10^{-6}	6.36×10^{-5}			

^aAt iteration point n , the energy and \mathbf{F} are evaluated and ${}^{n+1}\underline{\lambda}$ is determined. Thus at iteration point n , iteration $n + 1$ is performed; $\|{}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\|$ is an approximation to $\|{}^n\mathbf{e}\| = \|{}^n\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bInitializing Newton–Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy – 149.2781477108 a.u.

from the Newton–Raphson iteration points 4 or 3 (Tables XXV and XXVI) of the Newton–Raphson calculation of Table IV. In Table XXVII we report the DFP and BFGS update calculations initiated using \mathbf{G} from iteration point 4 of Table IV. The Broyden rank-1 (Table XXV) and the DFP calculations converge in five iterations, whereas the BFGS calculation uses seven iterations to converge. All three update calculations converge monotonically and with about the same convergence characteristics. In Table XXVIII we report the DFP and BFGS update calculations initiated after Newton–Raphson iteration point 3 of Table IV. Both the DFP and BFGS calculations diverge. The Broyden rank-1 calculation was reported in Table XXVI and converges in seven iterations. The Broyden rank-1 update approach thus has a larger region of convergence in these excited state calculations than do the DFP and BFGS update approaches.

To understand why Broyden rank-1 update converges when the DFP and BFGS methods diverge, we analyze the Hessian updates that result in the three update methods. In the Broyden rank-1 update calculation in Table XXVI Hessian update corrections are in general very small. In the initial Broyden rank-1 update iteration the largest change in an approximate Hessian element is 3.16×10^{-3} , corresponding to a Hessian element that is changed from -1.78×10^{-2} to -1.46×10^{-2} . The exact Hessian element (when a nonlinear transformation of variables is carried out) is -1.16×10^{-2} . The Broyden rank-1 update thus significantly improves the agreement be-

TABLE XXVIII

BFGS and DFP Updates After Newton-Raphson Iteration Point 3 of Table IV ($E^3\Sigma_u^-$ of O₂ at 2.10 a.u.): Initial G is from Iteration Point 3 of Table IV

Iteration point ^a	BFGS			DFP		
	$E - E^{\text{CONV}c}$	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\ \sim \ {}^n\mathbf{e}\ $	$E - E^{\text{CONV}c}$	$\ \mathbf{F}\ $	$\ {}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\ \sim \ {}^n\mathbf{e}\ $
0 ^b	0.0000947284	1.91×10^{-1}	2.76×10^{-1}	0.0000947284	1.91×10^{-1}	2.76×10^{-1}
1	0.0000311601	2.06×10^{-2}	2.85×10^{-1d}	0.0000311601	2.06×10^{-2}	2.85×10^{-1d}
2	0.0000993762	2.78×10^{-2}	2.04×10^{-1d}	0.0000934239	2.78×10^{-2}	2.02×10^{-1d}
3	0.0000140309	4.22×10^{-2}	2.16×10^{-1d}	0.0000407694	9.44×10^{-3}	2.00×10^{-1d}
4	0.0002170358	6.20×10^{-2}	2.40×10^{-1d}	0.0000902366	2.72×10^{-2}	2.01×10^{-1d}
5	0.0001324723	6.60×10^{-2}	3.30×10^{-1d}	0.0000858022	4.34×10^{-2}	1.95×10^{-1}
6	0.0000614033	8.96×10^{-2}	2.86×10^{-1d}	0.0000964965	2.82×10^{-2}	3.64×10^{-2}
7	0.0017914279	1.08×10^{-1}	3.51×10^{-1d}	0.0000880822	2.64×10^{-2}	2.00×10^{-1d}
8	0.0024049328	1.28×10^{-1}	5.79×10^{-1d}			

^aAt iteration point n , the energy and \mathbf{F} are evaluated and ${}^{n+1}\underline{\lambda}$ is determined. Thus at iteration point n , iteration $n+1$ is performed; $\|{}^{n+1}\underline{\lambda} - {}^n\underline{\lambda}\|$ is an approximation to $\|{}^n\mathbf{e}\| = \|{}^n\underline{\lambda} - \underline{\alpha}\|$ [see Eq. (71)].

^bInitializing Newton-Raphson iteration. Updates start at iteration point 1.

^cHere, $E - E^{\text{CONV}}$ is the difference in atomic units between the total energy at the present step and the converged total energy – 149.2781477108 a.u.

^dThe step size and sign control algorithm¹⁷ is applied in this iteration.

tween the approximate Hessian matrix and the exact Hessian matrix. The Hessian updates can be justified further through the information about the exact Hessian that is contained in the quasi-Newton condition. The approximate Hessian element changes from -1.78×10^{-2} to -1.48×10^{-2} corresponds to a change of a variable ($\underline{\lambda}$ element change) of 0.23. The aforementioned trends are found in all the consecutive iterations. The success of the Broyden rank-1 update method can thus be ascribed to the fact that relative small changes occur in the Hessian updates and that these small changes are directly related to information about the exact Hessian matrix that is available in the quasi-Newton condition.

In the first BFGS iterations of Table XXVIII (the iterations using an initial $\mathbf{G}(0)$ of Newton-Raphson iteration point 3 of Table IV) very large changes occur in the approximate Hessian. For example a diagonal element of the Hessian matrix is changed from 2.08 to –34.8. The corresponding exact Hessian element is 2.50. Fourteen elements are changed by more than 2.0. These changes cannot at all be justified on the basis of information contained in the quasi-Newton condition. The aforementioned change of 36.8 is related to a change in a variable ($\underline{\lambda}$ element change) of 0.005. The large changes in the approximate Hessian can be traced back to very small denominators in the BFGS updates [see Eq. (194)]. The denominators $\mathbf{F}^T \mathbf{K}$ and

$\mathbf{K}^T \mathbf{H} \mathbf{K}$ of the BFGS update are -3.8×10^{-6} and 3.5×10^{-5} , respectively, in the initial update iteration of Table XXVIII. The wild changes of the Hessian approximation result in an undesired negative Hessian eigenvalue of -48. The same behavior continues during the update sequence of iterations, and the Hessian approximation fast loses any connection to the exact Hessian matrix. Therefore the BFGS update calculation diverges.

In the DFP update sequence of iterations we expect a behavior analogous to the one observed in the BFGS update calculation, since the denominators of a DFP update [see Eq. (195)] are the same as the denominators of a BFGS update. As an example of this behavior we note that in the first DFP update a diagonal element of the approximate Hessian is changed from 2.08 to 354, where the exact Hessian element is 2.50. We thus conclude that when updates are applied on excited states, the Broyden rank-1 and the Powell updates have larger regions of convergence than BFGS and DFP updates. This conclusion is contrary to the experience of numerical analysis. We should point out here that the experience of numerical analysis is obtained basically in connection with function minimization. The foregoing conclusions are obtained when optimizations are carried out to saddle points, and our conclusions therefore may differ from those in numerical analysis.

An investigation of the efficiency of the update methods can be obtained by comparing Broyden rank-1 and Powell update calculations with Newton-Raphson calculations. We consider the calculations on the $B^3\Sigma_u^-$ state at 2.13 a.u. In Tables XXI and XXII the Broyden rank-1 and Powell update calculations are reported initiated after Newton-Raphson iteration points 2 and 1, respectively, of Table III. Using the energy difference $E - E^{\text{CONV}}$ as a measure of the distance to the desired stationary point, in Table XXIX we have tabulated the number of update iterations of Tables XXI and XXII that are equivalent to a given number of Newton-Raphson iterations of Table III. The actual numbers in Table XXIX have been determined by requiring that the update give an energy approximation which is at least as good as the one of the equivalent number of Newton-Raphson iterations. When the update calculation is started out relatively early (after Newton-Raphson iteration point 1 of Table III), the update calculation fluctuates in the initial update iterations and a large number of update iterations (five) is required to reach a point that is at least as good as the one obtained by carrying out one additional Newton-Raphson iteration. After the fluctuations in the initial update iterations, the update calculation converges rapidly and monotonically and some two or three update iterations are in general found to simulate one Newton-Raphson iteration.

If the update calculation is started out after Newton-Raphson iteration point 2 (initial \mathbf{G} from iteration point 2), the update calculation converges rapidly and monotonically and some two or three update iterations simulate

TABLE XXIX

The Number of Powell Update Iterations That Is Equivalent to a Given Number of
Newton-Raphson Iterations ($B^3\Sigma_u^-$ of O₂ at 2.13 a.u.)

Powell update starts after Newton-Raphson iteration point 1 of Table III		Powell update starts after Newton-Raphson iteration point 2 of Table III	
Number of Newton-Raphson iterations	Equivalent number of Powell updates	Number of Newton-Raphson iterations	Equivalent number of Powell updates
1	5	1	3
2	5	2	4
3	7	3	6
4	11		

one Newton-Raphson iteration. A similar analysis may be performed on the $E^3\Sigma_u^-$ calculation at 2.10 a.u. for the Broyden rank-1 update calculations (Table XXVI) initiated after Newton-Raphson iteration point 3 of Table IV. This calculation gives a result similar to that of the $B^3\Sigma_u^-$ calculation. If the update is started close to convergence, the update results indicate that the number of update iterations required to converge is about twice the number required in the corresponding Newton-Raphson sequence. If the update is started earlier and fluctuations occur in the initial update iterations, the ratio may be as large as 3.

The computational work involved in carrying out an update iteration is very similar to the work involved in a fixed Hessian iteration. The convergence characteristics of fixed Hessian calculations have previously been described (Section VI). In the local region the Hessian updates do not significantly improve the convergence rate in the initial few iterations compared to a fixed Hessian approach. However, after these few iterations, Hessian updates become extremely important for obtaining efficient and rapid convergence. It is also of importance to note that the region of convergence of Hessian update methods is larger than the region of convergence of a fixed Hessian approach. Hessian update methods therefore generally can be used much more reliably and efficiently than fixed Hessian approaches to obtain local convergence of an MCSCF calculation. The Hessian update methods even have some promising global convergence properties.

G. Other Applications of Update Methods

Update methods have only recently been introduced in wavefunction optimization.^{21, 73} We anticipate that they will also ease the optimization of many types of wavefunctions besides MCSCF wavefunctions.

The methods can be modified so that one part of the Hessian is calculated exactly and other parts are approximated through update procedures. For

example, this variant can be used to approximate coupling elements (the block $\mathbf{A}^{\text{CO}} - \mathbf{B}^{\text{CO}}$) while the other parts of the Hessian are calculated exactly. Another conceivable use is to calculate exactly the part of the Hessian that can be constructed from integrals with one free index, while approximating other parts of the Hessian. Investigations of the usefulness of new update procedures^{76, 77} of, say, the Oren-Lauenberg⁷⁷ type may also be of interest.

VIII. CUBIC CONTRIBUTIONS IN MCSCF OPTIMIZATION

A. Theory

When the energy expansion in Eq. (53) is truncated after terms which are quadratic in $\underline{\lambda}$, the nonlinear Newton-Raphson iterative function [Eq. (60)] is straightforwardly derived. If a cubic expansion of the energy function is considered, the cubic iterative function is determined

$$F_i + G_{ij}\lambda_j + \frac{1}{2}K_{ijk}\lambda_j\lambda_k = 0 \quad (208)$$

When the Hessian matrix contains small eigenvalues, a serious slowdown of the global convergence of a Newton-Raphson calculation may be observed^{9, 15, 17, 18, 22} (see Section V). Such a deficiency may not show up when an iterative cubic function is used. The cubic term may dominate the modes that cause problems in the Newton-Raphson approach. To clarify this point, we transform Eq. (208) to the basis where the Hessian matrix is diagonal [Eq. (114)]. We then obtain

$$\bar{F}_i + \epsilon_{ij}\bar{\lambda}_j + \frac{1}{2}\bar{C}_{ij}\bar{\lambda}_j = 0 \quad (209)$$

where

$$\bar{\mathbf{C}} = \mathbf{U}^+ \mathbf{K} \underline{\lambda} \mathbf{U} \quad (210)$$

and hence

$$\bar{\lambda}_i = -(\epsilon + \frac{1}{2}\bar{\mathbf{C}})^{-1} \bar{F}_i \quad (211)$$

Equations (115) and (211) have the same structure and differ formally only in the inverse matrix where the ϵ matrix of Eq. (115) is changed to $\epsilon + \frac{1}{2}\bar{\mathbf{C}}$. Peculiarities in the convergence behavior that occur in the Newton-Raphson method due to small Hessian eigenvalues (e.g., near an inflection point) may be of less importance when Eq. (211) is used as an iterative function, since the term $\frac{1}{2}\bar{\mathbf{C}}$ is often of such magnitude that it will dominate the

very small ϵ elements. The modes of the Newton–Raphson approach that are dominated by the small eigenvalues of \mathbf{G} may therefore be expected to be controlled by the term $\frac{1}{2}\bar{\mathbf{C}}$ in Eq. (211), and the large step length amplitudes should be much less troublesome with the use of an iterative cubic MCSCF.

A solution to Eq. (208) may only be determined in terms of an iterative procedure, the solution of which then will satisfy Eq. (208). Each determination of the step length vector $\underline{\lambda}$ which solves Eq. (208) is called a *macro-iteration*. When the cubic term dominates certain modes, a solution to Eq. (208) cannot be determined perturbatively.

We demonstrate now how a Newton–Raphson technique may be used to solve Eq. (208). We define a vector function

$$f_i(\underline{\lambda}) = F_i + G_{ij} \lambda_j + \frac{1}{2} K_{ijk} \lambda_j \lambda_k \quad (212)$$

and attempt to determine iteratively a set of $\underline{\lambda}$ parameters such that the vector $\mathbf{f}(\underline{\lambda})$ becomes equal to zero. Carrying out a Taylor expansion of $\mathbf{f}(\underline{\lambda})$ about the point ${}^0\underline{\lambda}$ gives

$$f_i(\underline{\lambda}) = f_i({}^0\underline{\lambda}) + f'_{ij}({}^0\underline{\lambda})(\lambda_j - {}^0\lambda_j) + \dots \quad (213)$$

The first derivative matrix is determined from Eq. (212) to be

$$f'_{ij}(\underline{\lambda}) = G_{ij} + K_{ijk} \lambda_k \quad (214)$$

When a Newton–Raphson technique is used to obtain a solution to Eq. (208), only the linear terms in the Taylor expansion are kept and $\mathbf{f}(\underline{\lambda})$ is set equal to zero:

$$f_i({}^0\underline{\lambda}) + f'_{ij}({}^0\underline{\lambda})(\lambda_j - {}^0\lambda_j) = 0 \quad (215)$$

Thus the Newton–Raphson step length formally becomes²⁰

$$\Delta^0 \lambda_j = \lambda_j - {}^0\lambda_j = -f'({}^0\underline{\lambda})^{-1} f_j({}^0\underline{\lambda}) \quad (216)$$

The step length correction to ${}^0\underline{\lambda}$ is thus obtained by solving Eq. (216). An updated set of $\underline{\lambda}$ parameters is determined as

$${}^1\underline{\lambda} = {}^0\underline{\lambda} + \Delta^0 \underline{\lambda} \quad (217)$$

These ${}^1\underline{\lambda}$ parameters can then be used for the next application of Eq. (216) [i.e., ${}^1\underline{\lambda}$ replaces ${}^0\underline{\lambda}$ in Eqs. (212) and (214)], and the iterative sequence is continued until the correction term $\Delta \underline{\lambda}$ is smaller than a certain tolerance.

Each single iteration of such an iterative procedure [Eqs. (216) and (217)] will be referred to as a *micro-iteration*. A converged series of micro-iterations thus gives a set of λ parameters that satisfies Eq. (208). When a Newton–Raphson sequence of micro-iterations is performed, the step length corrections of a micro-iteration will, of course, show second-order convergence.

When the cubic iterative function of Eq. (208) is applied far from convergence, the cubic term cannot usually be considered as a small correction to the second-order function for the modes corresponding to the small eigenvalues of the Hessian matrix. If a Newton–Raphson (unconstrained) MCSCF step [Eq. (60)] is used as initial guess for a series of micro-iterations [Eqs. (216) and (217)], such a series of iterations may diverge. To determine a solution of the cubic iterative function in Eq. (208) when far from convergence, we have used the Newton–Raphson MCSCF step size and sign controlled rotational parameters¹⁷ (see Section V) as an initial guess of the series of micro-iterations and have obtained convergence in the sequence of micro-iterations in few iterations. However, in the calculations we discuss in the next subsection (VIII.B), the *final* λ from the series of micro-iterations for a certain macro-iteration is unconstrained even though the initial guess is the constrained Newton–Raphson λ .

We expect that the iterative cubic approach is particularly effective and reliable when far from convergence (the global convergence problem) since one more term is included in the expansion of the energy [Eq. (208)]. This is, in fact, borne out by the results to be presented in Section VIII.B. Hence, often the iterative cubic approach does not require that constraints are applied far from convergence. In that region one iterative cubic iteration can replace two to five constrained Newton–Raphson iterations. In addition, we show that iterative cubic step length amplitudes are often mimicked fairly well by the mode-controlled Newton–Raphson values.¹⁷ The constraint values for both mode damping¹⁵ and mode controlling¹⁷ have been determined by a somewhat “trial-and-error” process (often euphemistically called “numerical experience”), and this agreement is very encouraging.²⁰

When the cubic iterative functions of Eq. (208) are used in the local region, a solution to Eq. (208) may be determined perturbatively. Such a procedure is easily described by rearranging Eq. (208) as

$$\lambda_i = -G_{ij}^{-1}F_j - \frac{1}{2}G_{il}^{-1}K_{ljk}\lambda_j\lambda_k \quad (218)$$

A solution to Eq. (218) is then obtained by carrying out a sequence of iterations in which we initially set λ equal to zero on the right-hand side of Eq. (218). We then obtain the set of λ values of a Newton–Raphson MCSCF iteration [Eq. (60)]. The Newton–Raphson set of parameters are then in-

serted back into the right-hand side of Eq. (218), and the Chebyshev formula of Eq. (137) is determined. The iterative procedure may be continued until a self-consistent set of $\underline{\lambda}$ parameters are determined, which then constitutes a solution to Eq. (208). We should point out that this sequence of iterations will only converge linearly.

To understand the local convergence characteristics of the cubic iterative function we will now derive the error term of the cubic iterative function. To do so we insert the Chebyshev formula into the right-hand side of Eq. (218), keeping terms through third order in the Newton-Raphson step length:

$$\begin{aligned}\lambda_i = & {}^{\text{NR}}\lambda_i - \frac{1}{2}G_{ij}^{-1}K_{jkl}{}^{\text{NR}}\lambda_k{}^{\text{NR}}\lambda_l \\ & + \frac{1}{2}G_{ij}^{-1}K_{jkl}{}^{\text{NR}}\lambda_kG_{lp}^{-1}K_{pqr}{}^{\text{NR}}\lambda_q{}^{\text{NR}}\lambda_r + O(({}^{\text{NR}}\lambda)^4)\end{aligned}\quad (219)$$

where ${}^{\text{NR}}\lambda$ is the Newton-Raphson step length of Eq. (60). The error term of one Newton-Raphson iteration has previously been derived [see, e.g., Eq. (31)]. To understand the structure of the error term of the cubic iterative function it is necessary to determine the error term of the Newton-Raphson iteration to one higher order and therefore carry out the expansions in Eqs. (27) and (28) to one higher order in

$${}^0e = {}^0\lambda - \alpha \quad (220)$$

(the error vector for the initial guess of orbitals and states for the iterative cubic iteration). We obtain

$$\begin{aligned}{}^{\text{NR}}\lambda_i = & \alpha_i + \frac{1}{2}G_{ij}^{-1}(\underline{\alpha})K_{jmp}(\underline{\alpha}){}^0e_m{}^0e_p \\ & + \frac{1}{3}G_{ij}^{-1}(\underline{\alpha})M_{jmpq}(\underline{\alpha}){}^0e_m{}^0e_p{}^0e_q \\ & - \frac{1}{2}G_{ij}^{-1}(\underline{\alpha})K_{jrm}(\underline{\alpha}){}^0e_mG_{rk}^{-1}(\underline{\alpha})K_{kpq}(\underline{\alpha}){}^0e_p{}^0e_q + O({}^0e^4)\end{aligned}\quad (221)$$

The \mathbf{K} supermatrix may also be expanded around the stationary point $\underline{\alpha}$:

$$K_{ijk}({}^0\underline{\lambda}) = K_{ijk}(\underline{\alpha}) + M_{ijkl}(\underline{\alpha})[{}^0\lambda_l - \alpha_l] + \dots \quad (222)$$

Hence

$$\begin{aligned}-\frac{1}{2}(G({}^0\underline{\lambda}))_{ij}^{-1}K_{jkl}({}^0\underline{\lambda}) = & -\frac{1}{2}G_{ij}^{-1}(\underline{\alpha})K_{jkl}(\underline{\alpha}) - \frac{1}{2}G_{ij}^{-1}(\underline{\alpha})M_{ijkl}{}^0e_l \\ & + \frac{1}{2}G_{ij}^{-1}(\underline{\alpha})K_{jmp}(\underline{\alpha}){}^0e_mG_{pq}^{-1}(\underline{\alpha})K_{qkl}(\underline{\alpha}) + \dots\end{aligned}\quad (223)$$

Inserting Eqs. (221) and (223) into Eq. (219) and retaining only terms through third order in ${}^0\mathbf{e}$ gives the error term of one iteration of the cubic iterative function:

$$e_i^{IC} = \lambda_i - \alpha_i = -\frac{1}{6} G_{ij}^{-1}(\underline{\alpha}) M_{jmpq}(\underline{\alpha})^0 e_m^0 e_p^0 e_q \quad (224)$$

Equation (224) explicitly demonstrates third-order convergence of the iterative cubic MCSCF procedure.

We will now show that if an iterative procedure is used where only the initial Newton–Raphson and one additional micro-iteration of Eq. (216) are performed, then the errors that are introduced by not converging the sequence of micro-iterations will be of fourth order in the error vector ${}^0\mathbf{e}$. The set of rotational parameters obtained when a Newton–Raphson and a micro-iteration of Eq. (216) is carried out becomes

$${}^1\lambda_i = {}^{NR}\lambda_i - (f'({}^{NR}\underline{\lambda}))_{ij} f({}^{NR}\underline{\lambda})_j \quad (225)$$

where

$${}^{NR}\underline{\lambda} = -\mathbf{G}^{-1}\mathbf{F} \quad (226)$$

Equation (213) and (214) may be inserted into Eq. (225)

$${}^1\lambda_i = {}^{NR}\lambda_i - (G_{pq} + K_{pqr} {}^{NR}\lambda_r)_{ij}^{-1} (F_j + G_{jk} {}^{NR}\lambda_k + \frac{1}{2} K_{jkl} {}^{NR}\lambda_k {}^{NR}\lambda_l) \quad (227)$$

and expansion of the inverse matrix in Eq. (227) gives, when terms are kept through third order in ${}^{NR}\underline{\lambda}$,

$$\begin{aligned} {}^1\lambda_i = & {}^{NR}\lambda_i - \frac{1}{2} G_{ij}^{-1} K_{jkl} {}^{NR}\lambda_k {}^{NR}\lambda_l \\ & + \frac{1}{2} G_{ij}^{-1} K_{jkl} {}^{NR}\lambda_k G_{lp}^{-1} K_{pqr} {}^{NR}\lambda_p {}^{NR}\lambda_r + O(({}^{NR}\lambda)^4) \end{aligned} \quad (228)$$

Subtracting Eq. (228) from Eq. (219) (the iterative function of the cubic iterative procedure through third order in ${}^{NR}\underline{\lambda}$) and denoting $\underline{\lambda}$ of Eq. (219) ${}^{IC}\lambda$ then gives

$${}^{IC}\underline{\lambda} - {}^1\lambda = O(({}^{NR}\lambda)^4) \quad (229)$$

which shows that the errors that are introduced by not carrying out the sequence of micro-iterations to convergence are of fourth order in the error

The error vector of one iteration of the perturbative (Chebyshev) cubic MCSCF technique may be obtained by only keeping the first two terms in Eq. (219). Denoting the error term of the Chebyshev formula \mathbf{e}^{CB} we obtain

The error vector of one iteration of the perturbative (Chebyshev) cubic MCSCF technique may be obtained by only keeping the first two terms in Eq. (219). Denoting the error term of the Chebyshev formula \mathbf{e}^{CB} we obtain

$$e_i^{\text{CB}} = e_i^{\text{IC}} + \frac{1}{2} G_{ij}^{-1}(\underline{\alpha}) K_{jkl}(\underline{\alpha})^0 e_k G_{lp}^{-1}(\underline{\alpha}) K_{pqr}(\underline{\alpha})^0 e_q^0 e_r \quad (230)$$

A cubic calculation may also be carried out using two steps of a fixed Hessian approach. Subsequently, a new \mathbf{G} is constructed, two fixed Hessian steps performed, and so on. We will refer to such a calculation as a recursive cubic calculation. In Eq. (143) we demonstrated that the error term of a recursive cubic iteration is

$$e_i^{\text{RC}} = \frac{1}{2} L_{ijk}^0 e_k L_{jlp}^0 e_p^0 e_q \quad (231)$$

$$= \frac{1}{2} G_{ij}^{-1}(\underline{\alpha}) K_{jkl}(\underline{\alpha})^0 e_k G_{lp}^{-1}(\underline{\alpha}) K_{pqr}(\underline{\alpha})^0 e_p^0 e_r \quad (232)$$

We note that

$$e_i^{\text{CB}} = e_i^{\text{IC}} + e_i^{\text{RC}} \quad (233)$$

and that, hence, if either the iterative cubic or the recursive cubic procedures diverge locally, it is expected that the perturbative cubic (Chebyshev) will in general also diverge. In Section VI.C we demonstrated that the radius of convergence of a recursive cubic calculation is generally smaller than the radius of convergence for a Newton–Raphson calculation. Both the Chebyshev and the recursive cubic approaches therefore are useful only in the local region of an MCSCF calculation. In Table XXX we show the total

TABLE XXX
Total Order of Convergence for a Sequence of Second- and Third-Order Iterations

Iteration	Second-order MCSCF	Third-order MCSCF
1	2	3
2	4	9
3	8	27
4	16	81
5	32	243

local order of convergence of a cubic procedure versus a second-order procedure.

The only previous analysis of cubic contributions was by Yaffe and Goddard.²⁴ However, in their analysis, they did not include coupling with the CI states. Coupling terms enter in second order (the NR equation) and are particularly important when converging to excited states.¹³⁻¹⁵ Calculations that neglect coupling terms therefore cannot be expected to demonstrate either third-order or second-order convergence in general. The analysis by Yaffe and Goddard was primarily for the Chebyshev procedure. Hence, applicability without additional constraints is limited to the local region. Yaffe and Goddard performed SCF and GVB calculations using the Chebyshev procedure.

In Appendix E we discuss the explicit construction of the cubic contributions for both the iterative and Chebyshev cubic procedures. By explicitly constructing third-derivative terms multiplied by a vector, efficient procedures have been developed.¹⁶ In addition, large amounts of computer storage are not required since a matrix (and not a supermatrix) is calculated and stored. In Appendix F we discuss the various two-electron integral transformations required for Newton-Raphson, fixed Hessian, and cubic procedures. A new transformation technique is presented in Section IX.A and Appendix F that is extremely efficient for the transformations of the integrals required for the cubic procedures. In Section IX we discuss the overall efficiency of Newton-Raphson and cubic procedures.

B. Numerical Results

To demonstrate the global convergence characteristics of the iterative cubic approach and the local convergence properties of the iterative, the perturbative, and the recursive cubic approaches we first describe an O₂ calculation on the E³Σ_u⁻ state at 2.16 a.u.²⁰ In Table XXXI the results of carrying out an iterative cubic calculation are reported. For a comparison, the results of the corresponding step size and sign controlled Newton-Raphson calculation are also given in Table XXXI. The Newton-Raphson calculation converges in six iterations and is thus of average complexity.

We first consider the global convergence problem. The initial Newton-Raphson iteration gives a step length of 1.95, which the step size control automatically constrains to 0.72. The iterative cubic step length is 0.58. In the Newton-Raphson calculation there is one constraint applied which reduces a step length amplitude from 1.86 to 0.43. The corresponding step length amplitude with the (unconstrained) iterative cubic procedure is also 0.43. It is, of course, somewhat fortuitous that this amplitude is constrained with the mode-controlled Newton-Raphson approach to be exactly the value from the iterative cubic MCSCF. However, we have observed that the constrained

TABLE XXXI

Convergence Characteristics of the Iterative Cubic and a Newton-Raphson Calculation of the $E^3\Sigma_u^-$ State of O₂ at 2.16 a.u.

Iteration point ^a	Iterative Cubic			Newton-Raphson		
	$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $	$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $
0	0.0614068393	1.04×10^0	5.78×10^{-1}	0.0614068393	1.04×10^0	1.95×10^0
						$(7.21 \times 10^{-1})^c$
1	-0.0004083693	1.26×10^{-1}	1.68×10^{-1}	0.0028461916	1.22×10^{-1}	4.74×10^{-1}
2	-0.0000158313	5.38×10^{-3}	1.58×10^{-2}	-0.0001981711	6.54×10^{-2}	2.62×10^{-1}
3	0.0000000000	6.64×10^{-6}	2.61×10^{-5}	0.0000485658	1.50×10^{-2}	2.55×10^{-2}
4	0.0	< 10^{-8}	< 10^{-8}	0.0000000086	2.26×10^{-4}	5.39×10^{-4}
5				0.0000000000	1.30×10^{-6}	1.32×10^{-7}
6				0.0	< 10^{-8}	< 10^{-8}

^aAt iteration point n , the energy and all derivative matrices are evaluated and λ^{n+1} is determined. Thus, at iteration point n iteration $n + 1$ is performed; $\|\lambda^{n+1}\|$ is an approximation to $\|\lambda^n\| = \|\lambda^n - \alpha\|$ [see Eq. (71)].

^bHere, $E - E^{\text{CONV}}$ denotes the difference between the total energy of the considered step of the iterative procedure minus the total energy of the converged calculation. The total energy of the converged calculation is -149.2537057620 a.u.

^cThis is the value of the norm after one constraint with the mode controlling procedure¹⁷ is used to reduce the size of one step length amplitude.

values of the step length amplitudes of the step size and sign controlled Newton-Raphson approach are often very close to the iterative cubic values.

It is seen that the norms of the error vectors in the cubic iterative procedure are decreasing approximately cubically. The ratios $\|\lambda^{n+1}\|/\|\lambda^n\|^3$ (note that λ^{n+1} is determined at iteration point n) for $n = 1, 2$, and 3 are 0.87, 3.48, and 6.26, respectively, and hence are fairly constant, demonstrating the third-order convergence characteristics. In the Newton-Raphson sequence of iterations the ratios $\|\lambda^{n+1}\|/\|\lambda^n\|^2$ for $n = 2, 3, 4$, and 5 (1.17, 0.37, 0.83, and 0.45) are also fairly constant as expected of a second-order procedure.

In all cases, the series of micro-iterations which is required to obtain a solution to the cubic iterative function of Eq. (208) is carried out using the Newton-Raphson technique of Eq. (216). If a Newton-Raphson (unconstrained) MCSCF step [Eq. (60)] is used as initial guess for the series, in many cases such a series is found not to converge. To obtain a solution of the cubic iterative function when far from convergence (e.g., in the first macro-iteration) requires the Newton-Raphson step size and sign controlled parameters as an initial guess of the sequence. With such an initial guess, the sequence of micro-iterations converges in few iterations. An example of the

convergence of a sequence of micro-iterations is given in Table XXXII. Step size control is only applied in the initial micro-iteration of the first macro-iteration. Since we are using a Newton-Raphson technique to determine the step lengths of the micro-iterations, these step lengths, as expected, are diminishing quadratically.

The error vector obtained after the first iterative cubic iteration is carried out is $\sim 1.68 \times 10^{-1}$. When two Newton-Raphson steps are carried out the error vector norm is $\sim 2.62 \times 10^{-1}$, and when three NR steps are performed the error vector norm is $\sim 2.55 \times 10^{-2}$ (see Table XXXI). The initial macro-iteration of the iterative cubic approach thus is able to replace some number between 2 and 3 Newton-Raphson iterations. It is seen from Table XXXII that the norm of the error of the second micro-iteration in the first macro-iteration is 2.37×10^{-2} . If the second micro-iteration is discarded, it is seen from Table XXXII that the error after the first macro-iteration is $\sim 1.92 \times 10^{-1}$ ($1.68 \times 10^{-1} + 2.37 \times 10^{-2}$), which is *smaller* than the error after the second Newton-Raphson iteration (2.62×10^{-1}). Thus a first macro-iteration consisting of an initial step size and sign controlled Newton-Raphson iteration and one micro-iteration has a smaller approximate error vector norm than the first two Newton-Raphson iterations.

The first derivative of the function $f(\underline{\lambda})$ given in Eq. (212) is

$$f'_{ij}(\underline{\lambda}) = G_{ij} + K_{ijk}\lambda_k \quad (234)$$

Equation (234) is a first-order approximation to the Hessian matrix (using a linear transformation of variables) at the point defined by the step length parameters $\underline{\lambda}$ of the micro-iteration; that is,

$$G_{ij}(\underline{\lambda}) = G_{ij} + K_{ijk}\lambda_k + O(\underline{\lambda})^2 \quad (235)$$

TABLE XXXII

The Norm of the Correction to the Step Length $\|\underline{\lambda}\|$ of Each Micro-iteration for the Cubic Iterative Approach of Table XXXI

Micro-iteration	Macro-iteration			
	1	2	3	4
0 ^a	7.21×10^{-1}	1.68×10^{-1}	1.58×10^{-2}	2.61×10^{-5}
1	3.14×10^{-1}	6.71×10^{-2}	1.64×10^{-4}	5.47×10^{-10}
2	2.37×10^{-2}	2.82×10^{-2}		

^aMicro-iteration 0 gives the step length of the step size and sign controlled Newton-Raphson iteration used as the initial guess for solving Eq. (208).

The stability of the cubic iterative procedure may be understood by examining how the lowest eigenvalues of $f'(\lambda)$ develop during a sequence of micro-iterations. In Table XXXIII we report the lowest four eigenvalues of $f'(\lambda)$ in the sequence of micro-iterations of the first macro-iteration of Table XXXI. The eigenvalues of the micro-iteration point 0 denote the Hessian eigenvalues [$\lambda = \mathbf{0}$ in Eq. (214)]. It may be seen that a large change occurs in the small eigenvalues during the iterative sequence. The lowest eigenvalues of the Hessian matrix at macro-iteration 2 are -0.13904 , 0.03028 , 0.03536 , and 0.05078 , in fairly good agreement with the converged eigenvalues of the micro-iteration sequence of Table XXXIII. $f'(\lambda)$ is a first approximation to the Hessian matrix of macro-iteration 2 and therefore also contains one negative eigenvalue.

The unconstrained Newton-Raphson and also the perturbative cubic and recursive cubic approaches [see Eq. (151)] diverge for this case.²⁰ Only the iterative cubic approach and, of course, the step size and sign controlled Newton-Raphson approach have attractive global convergence properties. The details of constrained Newton-Raphson procedures are usually determined by numerical experience. Thus, when far from convergence, the iterative cubic approach is preferable. One iteration of the iterative cubic approach can replace two to five constrained Newton-Raphson steps when far from convergence. In Section IX, Appendices E and F, and Ref. 20, we demonstrate that the iterative cubic procedure is computationally of equal efficiency when applied far from convergence compared to a step sign and size controlled Newton-Raphson approach. The iterative cubic approach therefore has very attractive global convergence characteristics. More calculations, however, have to be carried out on difficult MCSCF cases before a full understanding of the global convergence characteristics of the iterative cubic approach is obtained.

TABLE XXXIII
Lowest Four Eigenvalues (a.u.) of $f'(\lambda)$ in Eq. (214) for the Initial
(Macro-iteration 1) Series of Micro-iterations of Table XXXI

Micro-iteration	Lowest four eigenvalues of $f'(\lambda)$			
	1	2	3	4
0	-0.21250	0.00348	0.00788	0.01594
1	-0.19264	0.02984	0.04206	0.05584
2	-0.24320	0.03058	0.04344	0.04768
⋮	⋮	⋮	⋮	⋮
Converged	-0.26002	0.02778	0.03800	0.04356

To illustrate the local convergence properties of cubically convergent approaches we also have carried out a recursive cubic and a perturbative cubic calculation on the $E^3\Sigma_u^-$ state at 2.16 a.u., using the orbitals and states obtained from the first iterative cubic macro-iteration as an initial guess. These calculations are reported in Table XXXIV together with the corresponding iterative cubic and the Newton–Raphson calculation results. Since the “initial guess” of orbitals and states is in the local region, the cubic approaches all show third-order convergence characteristics and as expected converge faster than the (unconstrained) Newton–Raphson calculation. As expected from Table XXX, approximately one iteration is saved in the local region by using cubic rather than quadratic MCSCF approaches. However, substantial differences are observed in the convergence rate of the three cubic approaches. This difference is caused by the different error terms of the three cubic approaches. The perturbative cubic approach has both the error term of the iterative cubic and the recursive cubic approaches [see Eq. (233)] and converges slowest of the three cubic procedures. The iterative cubic approach converges faster than the recursive cubic approach. An explanation may be that the error term of the iterative cubic and the recursive cubic approaches contains, respectively, one and two inverse Hessian matrices and that the small eigenvalues of the Hessian matrix dominate the error term.

To further demonstrate the local convergence properties of these cubic procedures, we report calculations on the $E^3\Sigma_u^-$ state of O₂ at 2.2 a.u. The basis set used in these calculations is given in Table XXXV, and the configurations are given in Table II. The initial guess for the Newton–Raphson part of these calculations is a set of grand canonical Hartree–Fock orbitals with occupation $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^{3.4} 3\sigma_g^{1.7} 1\pi_g^{0.6} 3\sigma_u^{0.3}$. The initial set of orbitals used in these MCSCF calculations was obtained at iteration point 4 of a sign and size controlled Newton–Raphson calculation for the $E^3\Sigma_u^-$ state (the first point where no constraints are required). Hence, the reported calculations are in the local region.

In Table XXXVI we show the results of the Newton–Raphson, iterative cubic, Chebyshev (perturbative) cubic, and recursive cubic MCSCF procedures. For these calculations convergence is obtained for all the cubic procedures in three iterations. The Newton–Raphson approach requires four iterations. As expected (see Table XXX) one iteration is saved in the local region by use of a cubic procedure.

For this case the recursive cubic is converging slightly faster than the other cubic procedures. The smallest (in magnitude) Hessian eigenvalue at iteration point 0 of 0.052 is relatively large, and only relatively modest fluctuations are observed in the smallest Hessian eigenvalues during the sequence of iterations. The ratios $\|\lambda^{n+1}\|/\|\lambda^n\|^3$ for the cubic procedures (see Table

TABLE XXXIV

Convergence Characteristics of the Iterative Cubic, Perturbative (Chebyshev) Cubic, a.u., Starting with Orbitals Obtained After One Macro-iteration of the Iterative Cubic

Iteration point ^a	Iterative cubic			Perturbative cubic		
	$E - E^{\text{CONV} b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \overset{n+1}{\lambda} - \overset{n}{\lambda}\ $	$E - E^{\text{CONV} b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \overset{n+1}{\lambda} - \overset{n}{\lambda}\ $
0	-0.0004083693	1.26×10^{-1}	1.68×10^{-1}	-0.0004083693	1.26×10^{-1}	1.73×10^{-1}
1	-0.000158313	5.38×10^{-3}	1.65×10^{-2}	-0.0001005968	1.24×10^{-2}	4.07×10^{-2}
2	0.0000000000	6.64×10^{-6}	2.81×10^{-5}	0.0000000055	6.88×10^{-5}	1.97×10^{-4}
3	0.0	$< 10^{-8}$	$< 10^{-8}$	0.0000000000	$< 10^{-8}$	1.30×10^{-8}
4				0.0	$< 10^{-8}$	$< 10^{-8}$

^aAt iteration point n , the energy and all derivative matrices are evaluated and $\overset{n+1}{\lambda}$ is determined. Thus, at iteration point n , iteration $n + 1$ is performed; $\|\overset{n+1}{\lambda}\|$ is an approximation to $\|\overset{n}{\lambda}\| = \|\overset{n}{\lambda} - \overset{n}{\alpha}\|$ [see Eq. (71)].

XXXVI) are almost constant. As explained previously, we expect

$$\|\overset{n+1}{\lambda}\| \leq K \|\overset{n}{\lambda}\|^3 \quad (236)$$

Thus these calculations demonstrate what might be called "classic" behavior; that is; $K \sim 1$ and the equality approximately holds in Eq. (236).

All the cubic procedures efficiently bring the MCSCF calculation to convergence when the "initial guess" of orbitals and states is in the local region. The three cubic approaches considered thus all have attractive local convergence properties. For further discussions of the efficiency of these procedures see Section IX, Appendixes E and F, and Ref. 20.

C. Summary and Conclusions

We have derived three cubic approaches—the iterative cubic, the perturbative cubic, and the recursive cubic:

1. The iterative cubic approach determines the set of rotational parameters from a cubic energy function. A solution of the cubic iterative function is defined in terms of an iterative solution, and we describe how a Newton-Raphson technique may be used to determine such a solution.
2. In the perturbative (Chebyshev) cubic approach the rotational parameters are determined such as to include the first-order perturbation correction to a Newton-Raphson set of rotational parameters.
3. The rotational parameters of the recursive cubic approach are obtained by carrying out a two-point fixed Hessian sequence of iterations. Such a sequence of iterations has previously been shown to have cubic convergence characteristics.

Recursive Cubic, and Newton-Raphson MCSCF for the $E^3\Sigma_u^-$ State of O₂ at 2.16 MCSCF

Recursive cubic			Newton-Raphson		
$E - E^{\text{CONV}}_b$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $	$E - E^{\text{CONV}}_b$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $
-0.0004083693	1.26×10^{-1}	1.68×10^{-1}	-0.0004083693	1.26×10^{-1}	1.68×10^{-1}
0.0000220099	6.48×10^{-3}	3.65×10^{-2}	0.0000811349	1.12×10^{-2}	4.43×10^{-2}
0.0000000000	1.80×10^{-5}	6.12×10^{-5}	0.00000000111	6.72×10^{-4}	2.57×10^{-3}
0.0	$< 10^{-8}$	$< 10^{-8}$	0.0000000000	1.89×10^{-6}	4.01×10^{-6}
			0.0	$< 10^{-8}$	$< 10^{-8}$

^bHere, $E - E^{\text{CONV}}$ denotes the difference between the total energy of the considered step of the iterative procedure minus the total energy of the converged calculation. The total energy of the converged calculation is -149.2537057620 a.u.

An error term analysis has been carried out for these three cubic approaches. To demonstrate the cubic convergence of the three cubic models, calculations were carried out with an "initial guess" of orbital and states in the local region. Differences in the rate of convergence of the three cubic procedures can be ascribed to the different structure of the error terms of the three approaches. All three cubic procedures have attractive local convergence properties.

We also carried out calculations with the three cubic procedures when the initial guess of orbitals was far from convergence. Only the iterative cubic approach has attractive global convergence characteristics. The iterative cubic procedure by itself (with no constraints) is able to correctly bring a calculation into the local region even when the Hessian matrix of the initial iteration has spurious negative eigenvalues. Large step lengths occur in a

TABLE XXXV
26 STO O₂ Basis Set ($R_{AB} = 2.2$ a.u.)

Function	Exponent
1s	7.65781
2s	2.68801
	1.67543
	0.9
2p	3.69445
	1.65864
	0.9

TABLE XXXVI
 Convergence Characteristics in the Local Region of the Iterative Cubic, Perturbative (Chebyshev)

Iteration point ^a	Iterative Cubic				Perturbative Cubic			
	$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ $	$\frac{\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ }{\ \underline{\lambda}^n\ ^3}$	$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ $	$\frac{\ \underline{\lambda}^{n+1} - \underline{\lambda}^n\ }{\ \underline{\lambda}^n\ ^3}$
0	0.0017425033	4.26×10^{-2}	3.14×10^{-1}		0.0017425033	4.26×10^{-2}	2.69×10^{-1}	
1	0.0000144510	3.20×10^{-3}	2.59×10^{-2}	0.84	0.0000088059	6.08×10^{-3}	2.10×10^{-2}	1.08
2	0.0000000000	2.88×10^{-6}	2.35×10^{-5}	1.35	0.0000000000	3.16×10^{-6}	7.21×10^{-6}	0.78
3	0.0000000000	$< 10^{-10}$	$< 10^{-10}$		0.0000000000	$< 10^{-10}$	$< 10^{-10}$	
4								

^aAt iteration point n , the energy and all derivative matrices are evaluated and $\underline{\lambda}^{n+1}$ is determined. Thus, at iteration point n , iteration $n+1$ is performed; $\|\underline{\lambda}^{n+1}\|$ is an approximation to $\|\mathbf{e}\| = \|\underline{\lambda} - \underline{g}\|$ [see Eq. (71)].

Newton-Raphson MCSCF approach because the Hessian matrix has small eigenvalues. When the iterative function is based on a third-order expansion of the energy function, the convergence of a sequence of iterations is not to the same extent dominated by the small eigenvalues of the Hessian matrix. Large step lengths of the Newton-Raphson procedure often are artifacts of truncating the energy function after second-order terms. In all our calculations, step length amplitudes from the iterative cubic procedure have been found to be very moderate.

When far from convergence (the global problem), the cubic iterative procedure therefore may constitute a real alternative to constrained Newton-Raphson MCSCF. The cubic calculations, however, indicate that the step size and sign controlled Newton-Raphson approaches often approximate fairly well the steps of an iterative cubic procedure.

IX. EFFECTIVE IMPLEMENTATION AND COMBINATION OF MCSCF PROCEDURES

In this review we have made no attempt to discuss in depth other MCSCF approaches such as Fock operator or super-CI procedures. We also have heretofore somewhat emphasized our own recent work. This, of course, is because we are most familiar with our own contributions and because we have performed similar MCSCF calculations using constrained and unconstrained Newton-Raphson, fixed Hessian, updated Hessian, and cubic procedures so that the nature of MCSCF convergence may be viewed as a more coherent whole.

Cubic, Recursive Cubic, and the Newton-Raphson MCSCF for the $E^3\Sigma_u^-$ State of O₂ at 2.20 a.u.

Recursive Cubic				Newton-Raphson			
$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $	$\frac{\ \lambda^{n+1} \sim \lambda^n\ }{\ \lambda^n\ ^2}$	$E - E^{\text{CONV}b}$ (a.u.)	$\ \mathbf{F}\ $	$\ \lambda^{n+1} \sim \lambda^n\ $	$\frac{\ \lambda^{n+1} \sim \lambda^n\ }{\ \lambda^n\ ^2}$
0.0017425033	4.26×10^{-2}	2.19×10^{-1}		0.0017425033	4.26×10^{-2}	2.19×10^{-1}	
0.0000077516	4.82×10^{-3}	2.39×10^{-2}	2.28	0.0000683846	8.24×10^{-3}	7.14×10^{-2}	1.49
0.0000000000	2.30×10^{-6}	7.94×10^{-6}	0.58	0.0000001390	1.30×10^{-3}	1.62×10^{-3}	0.32
0.0000000000	< 10^{-10}	< 10^{-10}		0.0000000000	9.60×10^{-7}	2.71×10^{-6}	1.03
					< 10^{-10}	< 10^{-10}	

^bHere, $E - E^{\text{CONV}}$ denotes the difference between the total energy at iteration point n minus the total energy of the converged calculation. The total energy of the converged calculation is -149.9712386378 a.u. Note that these calculations use the 26 STO basis set of Table XXXV. The initial guess is the set of orbitals from after iteration 4 of a constrained Newton-Raphson calculation.¹⁷

For super-CI and Fock operator approaches we, in general, refer the interested reader to the literature.^{1,2} The super-CI approach^{2,78-80} does not contain all second-order terms. It is expected that these contributions are particularly important for difficult cases. In the local region, super-CI converges only linearly. Fock operator procedures also are usually formulated without all second order contributions.^{1,2} In recent work by Das³³ and Werner and Meyer,²⁹⁻³⁰ complete second order Fock operator procedures have been formulated and studied.

Unitary exponential operators have been used in nuclear physics^{52,53,81} for many years. The first attempts to use unitary exponential operators in MCSCF in chemical physics were by Levy⁸² and Yaffe and Goddard.²⁴ Yaffe and Goddard²⁴ also developed both approximate second-order and approximate Chebyshev (third-order) techniques for the SCF and the MCSCF methods. They performed several SCF and GVB calculations with their approximate second and third order methods. For the MCSCF case Yaffe and Goddard did not include coupling with the MCSCF CI space, so their calculations were not fully second or third order. This can be critical, particularly for convergence to excited states. Further early work on second-order SCF optimization has been done by Douady et al.²⁵ (submission date, July 25, 1977; acceptance, July, 27 1977).

The first theoretical treatment of the general second-order MCSCF case using unitary exponential operators including coupling was by Dalgaard and Jørgensen.¹⁰ They formulated the two-step MCSCF approach in the language of second quantization. Yeager and Jørgensen¹¹ and Dalgaard¹² made further theoretical developments on the second-order procedures. Dalgaard

derived the one-step second-order procedure in the Brillouin formulation and presented some one-step calculations using the model of Pariser and Parr and of Pople.⁸³ The Brillouin formulation results with an unsymmetric "Hessian" matrix. Yeager and Jørgensen derived the one-step second-order procedure in the energy formulation and demonstrated relations between the one- and the two-step second-order procedures (some of which are described in Section III.C). Yeager and Jørgensen performed the first calculations with both the one- and the two-step second-order techniques.

At the 1979 Sanibel symposia D. Hopper and C. C. J. Roothaan lectured on new MCSCF procedures.²⁶ They commented there that their presentation was largely a reformulation in first quantized (wavefunction) form of the earlier work of Dalgaard and Jørgensen.¹⁰ In addition, they derived the one-step MCSCF procedure. Roothaan, Detrich, and Hopper subsequently submitted a manuscript on their developments.²⁶ To date they have reported no calculational results with their procedures. At the same 1979 Sanibel meeting Yeager presented work based on the already submitted paper of Yeager and Jørgensen.¹¹ Yeager discussed the theoretical development of both the one-step and the two-step MCSCF procedures as well as calculational results.

Since 1979 there have been several outstanding contributions in second-order MCSCF. It is impossible to adequately discuss them all in detail. Except for certain specific examples, we refer interested readers to the original literature.^{2, 10–37, 84–85}

The previous sections discussed primarily theoretical aspects of the iterative procedures. The present section complements this discussion with descriptions of how the procedures can be further modified so computational efforts are minimized. We first discuss (Section IX.A) how the procedures can be modified further in order to simplify integral transformations. Simplifications of the handling of the CI space are then indicated in Section IX.B. In Section IX.C we link different procedures together. It is shown that knowledge at one point can make computations at other points more effective. We then summarize the convergence characteristics of the methods mentioned and outline the methods' relative complexity. On the basis of this, efficient combinations of MCSCF procedures are discussed. In Section IX.D, we discuss two other ways of including higher order terms.

A. Simplifications of Transformations

Internal rotations in the subspace of completely occupied orbitals or internal rotations in the subspace of completely unoccupied orbitals do not change the total energy of the wavefunction. We shall now study a technique where transformations in these spaces are used to simplify two-electron transformations.

To do this, we introduce a formal partitioning of the set of properly symmetrized basis orbitals, b_k^+ . Three subsets, I, A, and S, of basis orbitals are defined so subset I matches the set of completely occupied (inactive) molecular orbitals, subset A matches the set of partly occupied (active) molecular orbitals, and subset S matches the set of completely unoccupied (secondary) molecular orbitals. Two sets of orbitals are said to match if for all symmetry types the number of orbitals of a given symmetry in one set equals the number of orbitals of the same symmetry in the other set. In this section basis orbitals from set I and inactive molecular orbitals are denoted $i, j, k \dots$; basis orbitals from set A and active molecular orbitals are denoted a, b, c ; and basis orbitals from set S and secondary molecular orbitals are denoted s, t, u, v . General indices are denoted o, p, q, r . The matrix \mathbf{C} which expands $\{a_k^+\}$ in terms of $\{b_k^+\}$ is partitioned thus:

$$\begin{pmatrix} \mathbf{a}_I^+ \\ \mathbf{a}_A^+ \\ \mathbf{a}_S^+ \end{pmatrix} = \begin{pmatrix} \mathbf{C}^{II} & \mathbf{C}^{IA} & \mathbf{C}^{IS} \\ \mathbf{C}^{AI} & \mathbf{C}^{AA} & \mathbf{C}^{AS} \\ \mathbf{C}^{SI} & \mathbf{C}^{SA} & \mathbf{C}^{SS} \end{pmatrix} \begin{pmatrix} \mathbf{b}_I^+ \\ \mathbf{b}_A^+ \\ \mathbf{b}_S^+ \end{pmatrix} \quad (237)$$

The energy truncated to second order is a function of the parameters $\underline{\lambda}$

$$E(\underline{\lambda}) = E_0 + \underline{\lambda}^T \mathbf{F} + \frac{1}{2} \underline{\lambda}^T \mathbf{G} \underline{\lambda}. \quad (238)$$

Introducing a new set of variables $\underline{\lambda}'$,

$$\underline{\lambda}' = \mathbf{P}'^{-1} \underline{\lambda} \quad (239)$$

where \mathbf{P}'^{-1} is nonsingular, the terms in $E(\underline{\lambda})$ are transformed to

$$E(\underline{\lambda}') = E_0 + \underline{\lambda}'^T \mathbf{P}'^T \mathbf{F} + \frac{1}{2} \underline{\lambda}'^T (\mathbf{P}'^T \mathbf{G} \mathbf{P}') \underline{\lambda}' \quad (240)$$

The set of $\underline{\lambda}$ parameters that correspond to the stationary point of Eq. (238) may be determined directly from Eq. (238). Alternatively this set of $\underline{\lambda}$ parameters may be evaluated through determining the set of $\underline{\lambda}'$ parameters that make Eq. (240) stationary and then using the transformation in Eq. (239).

We will now investigate the possibility of choosing \mathbf{P}' so $\mathbf{P}'^T \mathbf{F}$ and $\mathbf{P}'^T \mathbf{G} \mathbf{P}'$ can be calculated easier than \mathbf{F} and \mathbf{G} . We are only interested in orbital rotations, so \mathbf{P}' (the partition of \mathbf{P}' corresponding to orbital and state parameters) is confined to be of the form

$$\mathbf{P}' = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \quad (241)$$

The matrix Π may be defined as a direct product matrix

$$\Pi_{pp', qq'} = P_{pq} P_{p'q'} \quad (242)$$

where pp' (or qq') is an orbital excitation index. The matrix \mathbf{P} may be furthermore required to be a block diagonal with blocks dimensioned as the partitioning used in Eq. (237):

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}^{\text{II}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{\text{AA}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}^{\text{SS}} \end{pmatrix} \quad (243)$$

We are now ready to study elements of $\mathbf{P}'^T \mathbf{G} \mathbf{P}'$. The element $(\mathbf{P}'^T \mathbf{G} \mathbf{P}')_{si, t j}$ becomes (see Appendix C for further notation):

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G} \mathbf{P}')_{si, t j} &= \mathbf{P}'_{uk, si} G_{uk, vl} \mathbf{P}'_{vl, t j} \\ &= P_{ki} P_{us} \{ 4[4(uk|vl) - (uv|kl) - (ul|vk)] \\ &\quad + 4\delta_{kl} (D_{uv}^I + D_{uv}^A) - 4\delta_{uv} (D_{kl}^I + D_{kl}^A) \} P_{lj} P_{vt} \end{aligned} \quad (244)$$

Introducing the notation

$$\tilde{a}_r^+ = P_{r'r} a_r^+ \quad (245)$$

and corresponding integrals and \mathbf{D} elements

$$\begin{aligned} \tilde{D}_{pq} &= P_{p'p} P_{q'q} D_{p'q'} \\ (\tilde{o}\tilde{p}|\tilde{q}\tilde{r}) &= P_{o'o} P_{p'p} P_{q'q} P_{r'r} (o'p'|q'r') \end{aligned} \quad (246)$$

one obtains

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G}' \mathbf{P}')_{si, t j} &= 4[4(\tilde{i}\tilde{s}|\tilde{j}\tilde{l}) - (\tilde{s}\tilde{l}|\tilde{i}\tilde{j}) - (\tilde{s}\tilde{j}|\tilde{t}\tilde{i})] \\ &\quad + 4(\mathbf{P}^T \mathbf{P})_{ij} (\tilde{D}_{st}^I + \tilde{D}_{st}^A) - 4(\mathbf{P}^T \mathbf{P})_{st} (\tilde{D}_{ij}^I + \tilde{D}_{ij}^A) \end{aligned} \quad (247)$$

Equation (247) shows that $(\mathbf{P}'^T \mathbf{G}' \mathbf{P}')_{si, t j}$ has the same structure as $G_{si, t j}$, but with transformed integrals and \mathbf{D} elements. Furthermore, δ_{qr} has been replaced with $(\mathbf{P}^T \mathbf{P})_{qr}$. The construction of $(\mathbf{P}'^T \mathbf{G}' \mathbf{P}')_{si, t j}$ from $(\tilde{o}\tilde{p}|\tilde{q}\tilde{r})$ and $\tilde{\mathbf{D}}$ is thus basically identical to the construction of $G_{si, t j}$ from $(op|qr)$ and \mathbf{D} .

An element $(\mathbf{P}'^T \mathbf{GP}')_{ai, bj}$ becomes

$$\begin{aligned}
 (\mathbf{P}'^T \mathbf{GP}')_{ai, bj} = & 4 \left[\rho_{b\tilde{a}, cd}^{(2)} (cd|\tilde{i}\tilde{j}) + 2\rho_{b\tilde{d}, \tilde{a}c}^{(2)} (\tilde{c}\tilde{i}|\tilde{d}\tilde{j}) \right] \\
 & + 2 \left[\mathbf{P}_{ac}^T - (\mathbf{P}^T \boldsymbol{\rho}^{(1)})_{ac} \right] \left[4(c\tilde{i}|\tilde{b}\tilde{j}) - (\tilde{b}\tilde{i}|c\tilde{j}) - (\tilde{b}\tilde{i}|\tilde{i}\tilde{j}) \right] \\
 & + 2 \left[\mathbf{P}_{bc}^T - (\mathbf{P}^T \boldsymbol{\rho}^{(1)})_{bc} \right] \left[4(c\tilde{j}|\tilde{a}\tilde{i}) - (\tilde{a}\tilde{j}|c\tilde{i}) - (\tilde{a}\tilde{c}|\tilde{i}\tilde{j}) \right] \\
 & + 2(\mathbf{P}^T \boldsymbol{\rho}^{(1)} \mathbf{P})_{ab} \tilde{D}_{ij}^I + 2(\mathbf{P}^T \mathbf{P})_{ij} (2\tilde{D}_{ab}^I + 2\tilde{D}_{ab}^A - \tilde{D}_{ba}) \\
 & - 4(\mathbf{P}^T \mathbf{P})_{ab} (\tilde{D}_{Ij}^I + \tilde{D}_{ij}^A) \tag{248}
 \end{aligned}$$

It is seen that the formula for $(\mathbf{P}'^T \mathbf{GP}')_{ai, bj}$ is not, in general, obtained from the similar formula for $G_{ai, bj}$ by replacing $(op|qr)$, D_{qr} , δ_{qr} with $(\tilde{o}\tilde{p}|\tilde{q}\tilde{r})$, \tilde{D}_{qr} , $(\mathbf{P}^T \mathbf{P})_{qr}$. In order to avoid complications, \mathbf{P}^{AA} is set equal to 1. The transformation to orbitals of class A can be simplified by other means, but for reasons of clarity we will not discuss this here.

With \mathbf{P}^{AA} as the unit matrix, all elements of $\mathbf{P}'^T \mathbf{F}$ and $\mathbf{P}'^T \mathbf{GP}'$ are reported in Appendix C. All elements of $\mathbf{P}'^T \mathbf{GP}'$ and $\mathbf{P}'^T \mathbf{F}$ are seen to correspond to elements of \mathbf{G} and \mathbf{F} with the changes

$$\begin{aligned}
 (op|qr) &\rightarrow (\tilde{o}\tilde{p}|\tilde{q}\tilde{r}) \\
 D_{qr} &\rightarrow \tilde{D}_{qr} \\
 \delta_{qr} &\rightarrow (\mathbf{P}^T \mathbf{P})_{qr} \tag{249}
 \end{aligned}$$

This is due to the absence of indices corresponding to completely occupied and completely unoccupied orbitals in two-electron density matrices in the formulas for \mathbf{F} and \mathbf{G} (see Appendix C).

So far we have only introduced a new set of variables λ' [Eq. (239)] and used this set of variables to rewrite the total energy expression as in Eq. (240). Corresponding to this new set of variables, formulas have been derived for the “gradient” $\mathbf{P}'^T \mathbf{F}$ and the “Hessian” $\mathbf{P}'^T \mathbf{GP}'$. We will now show that with an appropriate choice of the matrix \mathbf{P} in Eq. (243) the two-electron integral transformations that are required for constructing $\mathbf{P}'^T \mathbf{GP}$ and $\mathbf{P}'^T \mathbf{F}$ are simpler than those required for constructing \mathbf{G} and \mathbf{F} by themselves. To calculate $\mathbf{P}'^T \mathbf{GP}'$ and $\mathbf{P}'^T \mathbf{F}$, integrals must be transformed to the tilde ($\tilde{}$) basis [Eq. (237)]. The corresponding orbital expansion is

$$\begin{pmatrix} \tilde{\mathbf{a}}_I^+ \\ \tilde{\mathbf{a}}_A^+ \\ \tilde{\mathbf{a}}_S^+ \end{pmatrix} = \begin{pmatrix} \mathbf{P}^{II} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}^{SS} \end{pmatrix} \begin{pmatrix} \mathbf{C}^{II} & \mathbf{C}^{IA} & \mathbf{C}^{IS} \\ \mathbf{C}^{AI} & \mathbf{C}^{AA} & \mathbf{C}^{AS} \\ \mathbf{C}^{SI} & \mathbf{C}^{SA} & \mathbf{C}^{SS} \end{pmatrix} \begin{pmatrix} \mathbf{b}_I^+ \\ \mathbf{b}_A^+ \\ \mathbf{b}_S^+ \end{pmatrix} \tag{250}$$

A choice that simplifies the construction of $(\delta\tilde{p}|\tilde{q}\tilde{r})$ is

$$\begin{aligned}\mathbf{P}^{\text{II}} &= \mathbf{C}^{\text{II}^{-1}} \\ \mathbf{P}^{\text{SS}} &= \mathbf{C}^{\text{SS}^{-1}}\end{aligned}\quad (251)$$

Then

$$\begin{pmatrix} \tilde{\mathbf{a}}_I^+ \\ \tilde{\mathbf{a}}_A^+ \\ \tilde{\mathbf{a}}_S^+ \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{C}'^{\text{IA}} & \mathbf{C}'^{\text{IS}} \\ \mathbf{C}'^{\text{AI}} & \mathbf{C}'^{\text{AA}} & \mathbf{C}'^{\text{AS}} \\ \mathbf{C}'^{\text{SI}} & \mathbf{C}'^{\text{SA}} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{b}_I^+ \\ \mathbf{b}_A^+ \\ \mathbf{b}_S^+ \end{pmatrix} \quad (252)$$

The set of orbitals $\{\tilde{a}_k^+\}$ are nonorthogonal with the choice of the matrix \mathbf{P} in Eq. (251). The basis $\{\tilde{a}_k^+\}$ is, however, only used to construct the modified two-electron integrals (etc.) that are required in evaluating $\mathbf{P}'^T \mathbf{G} \mathbf{P}'$ [e.g., see Eq. (247) and Appendix C] and $\mathbf{P}'^T \mathbf{F}$, so that the nonorthogonality of the (\cdot) orbitals is of no consequence. The two-electron integrals in $\mathbf{P}'^T \mathbf{G} \mathbf{P}'$ and $\mathbf{P}'^T \mathbf{F}$ are easier to evaluate than the ones required for constructing \mathbf{G} and \mathbf{F} since the transformation matrix \mathbf{C}' that is obtained through the choice of \mathbf{P} in Eq. (251) reduces the transformations needed in second-order procedures. This is because the usual largest block \mathbf{C}^{SS} now becomes a unit matrix, $\mathbf{C}^{\text{SS}} = \mathbf{1}$. The advantages of introducing the reduced transformations are discussed in detail in Appendix F.

The additional work required to be able to perform these reduced transformations are inversion and multiplication of matrices [see Eqs. (251) and (239)]. Due to the direct product character of \mathbf{P}' , only matrices with the dimensions equal to the number of orbitals are involved. These additional matrix operations are therefore of negligible complexity.

The simplified reduced matrix Eq. (252) can also be used in connection with gradient based methods and one-point cubic methods. In gradient calculations iterations are of the type

$${}^{i+1}\underline{\lambda} = {}^i\underline{\lambda} - \mathbf{H}^{-1}{}_i\mathbf{F}({}^i\underline{\lambda}) \quad (253)$$

By using the reduced matrix \mathbf{C}' , the vector $\mathbf{P}'^T \mathbf{F}$ is calculated. Then \mathbf{F} is calculated by multiplying $\mathbf{P}'^T \mathbf{F}$ with $(\mathbf{P}'^T)^{-1}$, which again is simple because of the direct product nature of \mathbf{P}' .

The use of the reduced expansion matrix [Eq. (251)] in the iterative cubic procedure is less obvious. The variables in the cubic energy expansion

$$E(\underline{\lambda}) = E_0 + \underline{\lambda}^T \mathbf{F} + \frac{1}{2} \underline{\lambda}^T \mathbf{G} \underline{\lambda} + \frac{1}{6} \mathbf{K} \underline{\lambda} \underline{\lambda} \underline{\lambda} \quad (254)$$

are transformed according to Eq. (239)

$$\mathbf{E}(\underline{\lambda}') = E_0 + \underline{\lambda}'^T (\mathbf{P}'^T \mathbf{F}) + \frac{1}{2} \underline{\lambda}'^T (\mathbf{P}'^T \mathbf{G} \mathbf{P}') \underline{\lambda}' + \frac{1}{6} \mathbf{K} (\mathbf{P}' \underline{\lambda}') (\mathbf{P}' \underline{\lambda}') (\mathbf{P}' \underline{\lambda}') \quad (255)$$

This cubic approximation can be minimized with respect to $\underline{\lambda}'$. In Appendix E a method to construct the terms necessary for minimizing Eq. (243) is outlined. The cubic contributions $\mathbf{K} \underline{\lambda}$ in Eq. (254) are introduced by calculating a “Hessian” with transformed integrals and density matrices. The transformed integrals are of the type

$$\begin{aligned} (op|\tilde{qr}) &= \kappa_{oo'}(o'p|qr) + \kappa_{pp'}(op'|qr) \\ &\quad + \kappa_{qq'}(op|q'r) + \kappa_{rr'}(op|qr') \end{aligned} \quad (256)$$

An analysis very similar to the analysis in Appendix E can be carried out for the modified cubic function in Eq. (255). The cubic terms are now introduced by calculation of a “Hessian” of the type $\mathbf{P}'^T \mathbf{K} (\mathbf{P}' \underline{\lambda}') \mathbf{P}'$. Instead of transformed integrals of the type of Eq. (256), the transformed integrals become

$$\begin{aligned} (\tilde{o}\tilde{p}|\tilde{qr}) &= (\tilde{o}'\tilde{p}|\tilde{qr})(\mathbf{P}^T \mathbf{P} \kappa')_{oo'} + (\tilde{o}\tilde{p}'|\tilde{qr})(\mathbf{P}^T \mathbf{P} \kappa')_{pp'} \\ &\quad + (\tilde{o}\tilde{p}|\tilde{q'r})(\mathbf{P}^T \mathbf{P} \kappa')_{qq'} + (\tilde{o}\tilde{p}|\tilde{qr'})(\mathbf{P}^T \mathbf{P} \kappa')_{rr'} \end{aligned} \quad (257)$$

From Eq. (255) it is easily seen that only integrals in the tilde ($\tilde{\cdot}$) basis are needed. The corresponding transformations are discussed in Appendix F.

The aforementioned modifications can result in numerical instabilities. These can be eliminated by choosing \mathbf{P}^{II} and \mathbf{P}^{SS} so that $\mathbf{P}^{\text{II}} \mathbf{C}^{\text{II}}$ and $\mathbf{P}^{\text{SS}} \mathbf{C}^{\text{SS}}$ are general diagonal matrices instead of unit matrices. The reduction of transformation times is not affected by this.

Thus, by redefining our MCSCF approaches in terms of the tilde ($\tilde{\cdot}$) basis set above, we are able to significantly reduce transformation times (see Appendix F). Some small additional work is required with this technique, that is, matrix inversion [Eq. (251)] and multiplication [to obtain $\underline{\lambda}$ from Eq. (239)]. These two matrix multiplications are $\propto N^3$ and hence, in general, will not significantly increase computer times.

B. Effective Treatment of Large CI Expansions

When the number n_{CI} of variables in the CI expansion of the MCSCF state is large, it is important to have effective algorithms to calculate terms connected with changes of the CI expansion. The purpose of this section is to discuss such algorithms. We first show how a basis for the CI space orthogonal to $|0\rangle$ can be defined through $n_{\text{CI}} - 1$ variables. This parametrization is

then used to construct the G^{CO} and G^{CC} blocks of the Hessian in a simple manner. We then discuss a direct second order MCSCF method of use when the complete Hessian cannot be constructed and stored in central memory.

1. Basis for the Orthogonal Complement States to $|0\rangle$

In the ansatz [Eq. (9)] for the MCSCF wavefunction a basis for the space $\langle|\Phi_g\rangle\rangle/\langle|0\rangle\rangle$ is needed. To define such a basis let $|\Phi_0\rangle$ be an arbitrary configuration state function which has non-zero overlap with $|0\rangle$. The CI expansion of $|0\rangle$ is

$$|0\rangle = C_{00}|\Phi_0\rangle + \sum_{g \neq 0} C_{g0}|\Phi_g\rangle \quad (258)$$

A set of parameters $S'_g (g \neq 0)$ is defined so

$$e^{i\hat{S}'}|\Phi_0\rangle = \exp i \left[i \sum_{g \neq 0} S'_g (\langle|\Phi_g\rangle\langle\Phi_0| - |\Phi_0\rangle\langle\Phi_g|) \right] |\Phi_0\rangle = |0\rangle \quad (259)$$

Since¹²

$$\exp \left[- (S'_g |\Phi_g\rangle\langle\Phi_0| - |\Phi_0\rangle\langle\Phi_g|) \right] |\Phi_0\rangle = \cos d |\Phi_0\rangle - \frac{\sin d}{d} S'_g |\Phi_g\rangle \quad (260)$$

where

$$d = \sqrt{\sum_{g \neq 0} S'^2_g} \quad (261)$$

and the parameters S'_g can be chosen as

$$S'_g = \frac{-C_{g0}d}{\sin d}, \quad d = \cos^{-1} C_{00} \quad (262)$$

With this choice, Eq. (259) is fulfilled and the subspace $\langle|k\rangle\rangle$ orthogonal to $|0\rangle$ may be constructed

$$|k\rangle = \exp \left[- S'_g (|\Phi_g\rangle\langle\Phi_0| - |\Phi_0\rangle\langle\Phi_g|) \right] |\Phi_k\rangle, \quad k \neq 0 \quad (263)$$

By expanding the exponential operator in Eq. (263) one obtains¹²

$$|k\rangle = |\Phi_k\rangle + \frac{\sin d}{d} S'_k |\Phi_0\rangle - S'_k \left(\frac{\cos d - 1}{d^2} \right) |*\rangle \quad (264)$$

where

$$|* \rangle = \sum_{g \neq 0} S'_g |\Phi_g \rangle \quad (265)$$

The simple expansion Eq. (264) is due to the rank-2 structure of \hat{S}' .

The \mathbf{G}^{CC} matrix

$$G_{kl}^{\text{CC}} = 2\langle k | H | l \rangle - 2\delta_{kl} E_0 \quad (266)$$

requires the transformation

$$\langle k | H | l \rangle = C_{gk} \langle \Phi_g | H | \Phi_l \rangle C_{g'l} \quad (267)$$

The direct transformation Eq. (267) requires, in general, a number of multiplications proportional to n_{CI}^3 , where n_{CI} is the dimension (e.g., number of configuration state functions) of the MCSCF CI space. Such a transformation can be extremely time consuming, for example, where $n_{\text{CI}} \approx 10^3$. The transformation in Eq. (267) is, however, not an n_{CI}^3 algorithm when $|k\rangle$ is defined as in Eq. (264). Then we have

$$\begin{aligned} \langle k | H | l \rangle &= \langle \Phi_k | H | \Phi_l \rangle + \frac{\sin^2 d}{d^2} S'_k S'_l \langle \Phi_0 | H | \Phi_0 \rangle \\ &\quad + \frac{\sin d}{d} (S'_l \langle \Phi_k | H | \Phi_0 \rangle + S'_k \langle \Phi_0 | H | \Phi_l \rangle) \\ &\quad - \frac{(\cos d - 1)}{d^2} (S'_l \langle \Phi_k | H | * \rangle + S'_k \langle * | H | \Phi_l \rangle) \\ &\quad - \frac{(\cos d - 1) \sin d}{d^3} (S'_k S'_l) \langle * | H | \Phi_0 \rangle \\ &\quad + \frac{(\cos d - 1)^2}{d^4} (S'_k S'_l) \langle * | H | * \rangle \end{aligned} \quad (268)$$

where the Einstein summation convention is not used. The basis Eq. (263) thus enables us to construct $\langle k | H | l \rangle$ from $\langle \Phi_g | H | \Phi_l \rangle$ with a number of multiplications proportional to n_{CI}^2 . The construction of \mathbf{G}^{CO} requires a transformation from $\langle 0 | [\mathbf{Q}^+, H] | \Phi_g \rangle$. This transformation is also simplified significantly by defining $|k\rangle$ as in Eq. (263).

Other ways of bypassing what was thought to be a n_{CI}^3 bottleneck in constructing $\langle k | H | l \rangle$ have been introduced by other workers.^{2, 30, 84} In these methods the complete set of CSFs $\langle |\Phi_g\rangle \rangle$ is used to span the CI manifold.

(The "Hessian" thereby obtained is related to the augmented Hessian⁸⁶ by similarity transformations.) This introduces a redundant variable which has been counteracted in different ways. Werner and Meyer^{29, 30} have introduced the normalization condition as a constraint via a Lagrange multiplier. Lengsfield and Liu⁸⁴ have eliminated the extra degree of freedom by introducing projection operators. The extra degree of freedom can also be eliminated by using the methods developed for the augmented Hessians since the "Hessian" obtained by Lengsfield and Liu⁸⁴ actually is an augmented Hessian.

2. Direct MCSCF

If the size of the CI expansion is very large, \mathbf{G}^{CO} and \mathbf{G}^{CC} cannot conveniently be stored in central memory, while \mathbf{G}^{OO} usually is of manageable dimensions. In this case a direct second-order MCSCF procedure (i.e., a method that solves a set of Newton-Raphson equations without setting up the Hessian) is of interest. The Newton-Raphson equations [Eq. (72)] can be solved iteratively in several ways. Many techniques for the iterative solution of linear equations have been developed.^{2, 87-89} The Newton-Raphson equations can also be solved by the update procedures discussed in Section VII. In either case it is necessary in each step to calculate a vector of the type

$$\mathbf{f}(\underline{\lambda}) = \mathbf{F} + \mathbf{G}\underline{\lambda} \quad (269)$$

The vector $\mathbf{G}\underline{\lambda}$ is

$$\begin{aligned} & \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \underline{\kappa} \\ \mathbf{s} \end{pmatrix} \\ &= \left(\begin{array}{l} \langle \mathbf{G}^{OO} \underline{\kappa} + \langle \delta 0 | [\mathbf{Q}^+ - \mathbf{Q}, H] | 0 \rangle + \langle 0 | [\mathbf{Q}^+ - \mathbf{Q}, H] | \delta 0 \rangle \\ \langle 0 | [\mathbf{R}^+ - \mathbf{R}, [(\mathbf{Q}_i^+ - \mathbf{Q}_i) \underline{\kappa}_i, H]] | 0 \rangle + \langle \delta 0 | [\mathbf{R}^+ - \mathbf{R}, H] | 0 \rangle + \langle 0 | [\mathbf{R}^+ - \mathbf{R}, H] | \delta 0 \rangle \end{array} \right) \end{aligned} \quad (270)$$

where

$$|\delta 0\rangle = -S_{n0}|n\rangle \quad (271)$$

If \mathbf{G}^{OO} cannot be stored either, one obtains

$$\begin{aligned} & \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \underline{\kappa} \\ \mathbf{s} \end{pmatrix} \\ &= \left(\begin{array}{l} \langle 0 | [\mathbf{Q}^+ - \mathbf{Q}, [(\mathbf{Q}_i^+ - \mathbf{Q}_i) \underline{\kappa}_i, H]] | 0 \rangle + \langle \delta 0 | [\mathbf{Q}^+ - \mathbf{Q}, H] | 0 \rangle + \langle 0 | [\mathbf{Q}^+ - \mathbf{Q}, H] | \delta 0 \rangle \\ \langle 0 | [\mathbf{R}^+ - \mathbf{R}, [(\mathbf{Q}_i^+ - \mathbf{Q}_i) \underline{\kappa}_i, H]] | 0 \rangle + \langle \delta 0 | [\mathbf{R}^+ - \mathbf{R}, H] | 0 \rangle + \langle 0 | [\mathbf{R}^+ - \mathbf{R}, H] | \delta 0 \rangle \end{array} \right) \end{aligned} \quad (272)$$

The operator $[(Q_i^+ - Q_i)\kappa_i, H]$ corresponds to H with replaced integrals

$$\begin{aligned} h_{rs} \rightarrow \tilde{h}_{rs} &= \kappa_{rp} h_{ps} + \kappa_{sp} h_{rp} \\ (rs|tu) \rightarrow (rs|\tilde{tu}) &= \kappa_{rp}(ps|tu) + \kappa_{sp}(rp|tu) + \kappa_{tp}(rs|pu) + \kappa_{up}(rs|tp) \end{aligned} \quad (273)$$

If Eq. (270) is used only integrals $(rs|\tilde{tu})$ with all four indices corresponding to occupied orbitals are needed. These integrals can be constructed from $(rs|tu)$ by a transformation that requires about $\frac{1}{2}n^4N$ multiplications (n is the number of occupied orbitals and N is the total number of orbitals). If Eq. (272) is used, then integrals $(rs|\tilde{tu})$ with one unrestricted index are needed. These integrals can be constructed from $(rs|tu)$ in about $\frac{3}{2}N^2n^3$ multiplications.

Introducing the notation

$$\begin{aligned} H' &= H + [\kappa, H] \\ |0'\rangle &= |0\rangle + |\delta 0\rangle \end{aligned} \quad (274)$$

one obtains to first order in $\underline{\lambda}$

$$\mathbf{F} + \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \underline{\kappa} \\ \mathbf{S} \end{pmatrix} = \begin{pmatrix} \mathbf{G}^{OO}\underline{\kappa} + \langle 0' | [\mathbf{Q}^+ - \mathbf{Q}, H] | 0' \rangle \\ \langle 0' | [\mathbf{R}^+ - \mathbf{R}, H'] | 0' \rangle \end{pmatrix} \quad (275)$$

or alternatively

$$\mathbf{F} + \begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \underline{\kappa} \\ \mathbf{S} \end{pmatrix} = \begin{pmatrix} \langle 0' | [\mathbf{Q}^+ - \mathbf{Q}, H'] | 0' \rangle \\ \langle 0' | [\mathbf{R}^+ - \mathbf{R}, H'] | 0' \rangle \end{pmatrix} \quad (276)$$

From Eqs. (275) and (276) it is seen that $\mathbf{F} + \mathbf{G}\underline{\lambda}$ can be calculated with a gradient routine with small modifications. The hamiltonian H is (partly in Eq. (275)) replaced by H' and $|0\rangle$ is replaced by $|0'\rangle$. The replacement of $|0\rangle$ with $|0'\rangle$ corresponds to the optimization of the CI coefficients. The orbital optimization corresponds to the replacement of H with H' .

If a usual direct CI iteration^{88,89} is performed H' and $\langle 0' | [\mathbf{Q}^+ - \mathbf{Q}, H'] | 0' \rangle$ are not calculated. The construction of $\langle 0' | [\mathbf{Q}^+ - \mathbf{Q}, H] | 0' \rangle$ requires about Nn^4 multiplications. The extra work for a direct second-order MCSCF iteration compared to a direct CI iteration is thus about $\frac{3}{2}Nn^4$ operations if Eq. (275) is used. A direct second-order MCSCF iteration is thus not expected to be significantly more complicated than a direct CI iteration.

The direct second-order MCSCF is, however, initiated by a two-electron transformation and [if Eq. (275) is used] a construction of \mathbf{G}^{OO} . Further-

more, convergence in direct second-order MCSCF results in a second-order approximation to a stationary point, whereas convergence in direct CI calculations gives the stationary point.

C. Combination and Overview of Methods

In this section we show how different methods can be linked together to give an effective and sophisticated polyalgorithm. We first outline how second- or third-order information at one point can be used together with first-order information at other points. A summary of the methods is then given, and convergence characteristics are compared. By weighing the convergence characteristics of a given method with the complexity of a method, relative efficiencies are then obtained. This information is used finally to sketch the optimal combination of the procedures.

1. Combination of Gradient-Based Methods with Second-Order Methods

The fixed Hessian method (Section VI.B) combines second-order information of one point with first-order information of other points. This yields an algorithm that is effective only if a few iterations with a fixed Hessian are carried out before the Hessian is reevaluated.

It was noted in Section VII.E that the first two iterations of a fixed Hessian sequence in the local region seem as effective as the corresponding first two update iterations. When the initial Hessian approximation is an exact Hessian of a point in the local region, the imposition of the quasi-Newton condition does not usually improve the Hessian approximation in the first few update iterations. Since

$$\mathbf{G}^{(k)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}} = \mathbf{G}^{(k+1)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}} + O^{(k+1)\underline{\lambda} - k\underline{\lambda}}^2 \quad (277)$$

it is seen that using $\mathbf{G}^{(k)}(\underline{\lambda})$ as a Hessian approximation, \mathbf{H}_{k+1} is consistent with Eq. (161). Since this equation is the rationale behind the quasi-Newton condition Eq. (160), it is obvious that fixed Hessian methods and update methods based on Eq. (160) often behave similarly in the first iteration where the Hessian is updated.

In order to obtain new information about $\mathbf{G}^{(k+1)}(\underline{\lambda})$ from $\mathbf{F}^{(k+1)}(\underline{\lambda})$, when iteration $k+1$ is a Newton-Raphson iteration, $\mathbf{F}^{(k+1)}(\underline{\lambda})$ and $\mathbf{G}^{(k+1)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}}$ are expanded through second order

$$\begin{aligned} \mathbf{F}^{(k+1)}(\underline{\lambda}) &= \mathbf{F}^{(k)}(\underline{\lambda}) + \mathbf{G}^{(k)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}} + \frac{1}{2}\mathbf{K}^{(k)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}}(k+1)\underline{\lambda} - k\underline{\lambda} \\ \mathbf{G}^{(k+1)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}} &= \mathbf{G}^{(k)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}} + \mathbf{K}^{(k)}(\underline{\lambda})^{(k+1)\underline{\lambda} - k\underline{\lambda}}(k+1)\underline{\lambda} - k\underline{\lambda} \end{aligned} \quad (278)$$

Thus, the exact Hessian $\mathbf{G}^{(k+1)}\underline{\lambda}$ satisfies through second order

$$\mathbf{G}^{(k+1)}\underline{\lambda}(\underline{\lambda} - \underline{\lambda}) = -\mathbf{G}^{(k)}\underline{\lambda}(\underline{\lambda} - \underline{\lambda}) + 2\mathbf{F}^{(k+1)}\underline{\lambda} - 2\mathbf{F}^{(k)}\underline{\lambda} \quad (279)$$

A condition for an Hessian approximation \mathbf{H}_{k+1} , when the exact Hessian $\mathbf{G}^{(k)}\underline{\lambda}$ is known, is thus

$$\mathbf{H}_{k+1}(\underline{\lambda} - \underline{\lambda}) = -\mathbf{G}^{(k)}\underline{\lambda}(\underline{\lambda} - \underline{\lambda}) + 2\mathbf{F}^{(k+1)}\underline{\lambda} - 2\mathbf{F}^{(k)}\underline{\lambda} \quad (280)$$

If iteration $k+1$ is an unmodified Newton–Raphson iteration [Eq. (20)], Eq. (280) becomes

$$\mathbf{H}_{k+1}(\underline{\lambda} - \underline{\lambda}) = 2\mathbf{F}^{(k+1)}\underline{\lambda} - \mathbf{F}^{(k)}\underline{\lambda} \quad (281)$$

The close resemblance between Eqs. (281) and (160) is noted.

If an update procedure based on Eqs. (280) and (281) is carried out, the corresponding approximate Hessian \mathbf{H}_{k+1} fulfills

$$[\mathbf{H}_{k+1} - \mathbf{G}^{(k+1)}\underline{\lambda}] (\underline{\lambda} - \underline{\lambda}) = O(\underline{\lambda} - \underline{\lambda})^3 \quad (282)$$

We thus expect that updates based on Eqs. (280) and (281) will converge faster than the corresponding fixed Hessian iteration, even in the initial iterations.

The update condition [Eq. (280) or (281)] must be supplemented with other conditions. Based on the results discussed in Section VII.F, we suggest that \mathbf{H}_{k+1} can be chosen so that $\|\mathbf{H}_{k+1} - \mathbf{G}^{(k)}\underline{\lambda}\|$ is minimal.

2. Combination of Gradient-Based Methods with the Iterative Cubic Method

If iteration k is an iterative cubic iteration at $\underline{\lambda}^{k-1}$, we obtain a first-order approximation \mathbf{G}'_k to the exact Hessian at the resulting iteration point $\underline{\lambda}^k$:

$$\begin{aligned} \mathbf{G}^{(k)}\underline{\lambda} &= \mathbf{G}'_k + O(\underline{\lambda}^k - \underline{\lambda}^{k-1})^2 \\ &= \mathbf{G}^{(k-1)}\underline{\lambda} + \mathbf{K}^{(k-1)}\underline{\lambda}(\underline{\lambda}'^k - \underline{\lambda}^{k-1}) + O(\underline{\lambda}^k - \underline{\lambda}^{k-1})^2 \end{aligned} \quad (283)$$

Typically $\underline{\lambda}'^k$ is the last point in the sequence of micro-iterations of macro-iteration k at which the first-order Hessian [Eq. (235)] is calculated. The reason we allow $\underline{\lambda}'^k$ to differ from $\underline{\lambda}^k$ is that the first-order Hessian [Eq. (235)] is usually not evaluated at the final point $\underline{\lambda}^k$. In Eq. (283) $\underline{\lambda}'^k$ is at least a first-order approximation to $\underline{\lambda}^k$ in the following sense

$$\underline{\lambda}'^k = \underline{\lambda}^k + \underline{\delta}, \quad \|\underline{\delta}\| = C \|\underline{\lambda}^{k-1} - \underline{\alpha}\|^2 \quad (284)$$

where $\underline{\alpha}$ is the stationary point of interest. Later we will prove a method which requires that ${}^k\underline{\lambda}' - {}^k\underline{\lambda}$ have third-order (rather than second-order) errors, that is, ${}^k\underline{\lambda}'$ is then closer to ${}^k\underline{\lambda}$.

We can now perform a sequence of "modified fixed Hessian" iterations

$$\begin{aligned} {}^k\underline{\lambda}^{\text{MFH}} &= {}^k\underline{\lambda} \\ {}^{k+i}\underline{\lambda}^{\text{MFH}} &= {}^{k+i-1}\underline{\lambda}^{\text{MFH}} - \mathbf{G}_k'^{-1} \mathbf{F}({}^{k+i-1}\underline{\lambda}^{\text{MFH}}), \quad i = 1, 2, \dots, n \end{aligned} \quad (285)$$

The method defined by Eq. (285) uses third-order information at ${}^{k-1}\underline{\lambda}$. We now show that this increases the convergence rate compared to usual fixed Hessian methods (Section VI.B).

The Newton-Raphson iteration at ${}^k\underline{\lambda}$ is

$${}^{k+1}\underline{\lambda}^{\text{NR}} = {}^k\underline{\lambda} - (\mathbf{G}({}^k\underline{\lambda}))^{-1} \mathbf{F}({}^k\underline{\lambda}) \quad (286)$$

A procedure consisting of one iterative cubic iteration and one Newton-Raphson iteration has order 6 (3·2):

$$\|{}^{k+1}\underline{\lambda}^{\text{NR}} - \underline{\alpha}\| = \|{}^{k-1}\underline{\lambda} - \underline{\alpha}\|^6 \quad (287)$$

Since iteration k is cubic, the resulting gradient $\mathbf{F}({}^k\underline{\lambda})$ is of third order in $\|{}^{k-1}\underline{\lambda} - \underline{\alpha}\|$. We then have

$$\begin{aligned} \|{}^{k+1}\underline{\lambda}^{\text{MFH}} - \underline{\alpha}\| &= \|{}^k\underline{\lambda} - [\mathbf{G}({}^k\underline{\lambda}) + O({}^{k-1}\underline{\lambda} - \underline{\alpha})^2]^{-1} \mathbf{F}({}^k\underline{\lambda}) - \underline{\alpha}\| \\ &= \|{}^{k+1}\underline{\lambda}^{\text{NR}} + O({}^{k-1}\underline{\lambda} - \underline{\alpha})^5 - \underline{\alpha}\| \\ &= O({}^{k-1}\underline{\lambda} - \underline{\alpha})^5 \end{aligned} \quad (288)$$

The procedure consisting of one iterative cubic iteration and one modified fixed Hessian iteration is a fifth-order procedure.

The foregoing result can be generalized inductively. For an arbitrary value of i in Eq. (285), one obtains

$$\|{}^{k+i}\underline{\lambda}^{\text{MFH}} - \underline{\alpha}\| = C \|{}^{k-1}\underline{\lambda} - \underline{\alpha}\|^{3+2i} \quad (289)$$

By calculating 1, 2, 3, ... gradients after one iterative cubic iteration, Eq. (285) gives a procedure of order 5, 7, 9,

The values of n in Eq. (285) corresponding to an optimal combination of the two iteration types involved can easily be found. The optimal number of gradient evaluations per cubic iteration is obtained when approximately 30%

of computer processing time is used to construct gradients.¹⁶ It is thus not advisable to keep the same "Hessian" \mathbf{G}'_k for very many iterations.

The fast convergence obtained with the algorithm Eq. (285) is due to the good approximation \mathbf{G}'_k one has to $\mathbf{G}({}^k\lambda)$. It is easily shown that \mathbf{G}'_k is a Hessian approximation that satisfies Eqs. (161) and (282) to relevant order. If one wants to obtain new information about $\mathbf{G}({}^k\lambda)$ from $\mathbf{F}({}^k\lambda)$ and a first-order Hessian [Eq. (235)], update conditions Eqs. (161) and (282) thus can not be used. We now derive an update condition that gives new information about $\mathbf{G}({}^k\lambda)$ from Eq. (235) and $\mathbf{F}({}^k\lambda)$. Such a procedure should usually give faster convergence than the modified fixed Hessian procedure hitherto discussed. $\mathbf{F}({}^k\lambda)$ and $\mathbf{G}({}^k\lambda)({}^k\lambda - {}^{k-1}\lambda)$ are expanded to third order:

$$\begin{aligned}\mathbf{F}({}^k\lambda) &= \frac{1}{3!} \mathbf{M}({}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda) + \dots \\ \mathbf{G}({}^k\lambda)({}^k\lambda - {}^{k-1}\lambda) &= \mathbf{G}''_k({}^k\lambda - {}^{k-1}\lambda) \\ &\quad + \frac{1}{2} \mathbf{M}({}^k\lambda - {}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda) + \dots \quad (290)\end{aligned}$$

In the expansion of $\mathbf{F}({}^k\lambda)$ we used the fact that iteration k is an iterative cubic iteration, so the first three terms in the expansion of \mathbf{F} [Eq. (59)] vanish. In Eq. (290)

$$\mathbf{G}''_k = \mathbf{G}({}^{k-1}\lambda) + \mathbf{K}({}^{k-1}\lambda)({}^k\lambda - {}^{k-1}\lambda) + O({}^k\lambda - {}^{k-1}\lambda)^3 \quad (291)$$

contrary to the previous requirement [Eq. (284)]. A third-order condition to $\mathbf{G}({}^k\lambda)$ is thus

$$\mathbf{G}({}^k\lambda)({}^k\lambda - {}^{k-1}\lambda) = \mathbf{G}''({}^k\lambda - {}^{k-1}\lambda) + 3\mathbf{F}({}^k\lambda) \quad (292)$$

When \mathbf{G}''_k is known, a condition for the Hessian approximation in the following iteration is thus

$$\mathbf{H}_k({}^k\lambda - {}^{k-1}\lambda) = \mathbf{G}''_k({}^k\lambda - {}^{k-1}\lambda) + 3\mathbf{F}({}^k\lambda) \quad (293)$$

If \mathbf{H}_k is in accord with Eq. (293), it satisfies the equation

$$[\mathbf{H}_k - \mathbf{G}({}^k\lambda)]({}^k\lambda - {}^{k-1}\lambda) = \mathbf{0} + O({}^k\lambda - {}^{k-1}\lambda)^4 \quad (294)$$

When the new error ${}^k\lambda - \underline{\alpha}$ is nearly parallel with ${}^k\lambda - {}^{k-1}\lambda$, an iteration with this choice of \mathbf{H}_k must be very similar to a Newton-Raphson iteration. If the error ${}^k\lambda - \underline{\alpha}$ is nearly orthogonal to ${}^k\lambda - {}^{k-1}\lambda$, the use of update condi-

tion Eq. (293) does not improve the rate of convergence significantly compared to Eq. (285) with $n = 1$.

The third-order condition Eq. (293) must be supplemented with other requirements in order to define \mathbf{H}_k unambiguously. Based on experience with conventional update methods (Section VII), we suggest that \mathbf{H}_k should be chosen so the norm of $\mathbf{H}_k - \mathbf{G}_k''$ is minimal.

3. Overview of Methods and Optimal Combinations of Methods

The convergence characteristics of the iterative cubic method, the step controlled Newton–Raphson method, update methods, and fixed Hessian methods are summarized in Table XXXVII. The typical regions of convergence and the nonlocal and local convergence rates are described for each method. For clarity, in this section we do not include analyses of other cubic procedures or “infinite order” techniques. Of course, it is straightforward to extend the analysis of this section to the other cubic procedures.

The iterative cubic method requires the smallest number of iterations to converge. The small number of iterations required with this method is partly gained at the expense of increased complexity of a single iterative cubic iteration. In order to study relative efficiencies of different methods, one must study the amount of work involved in a single iteration of a given method. The basis for this analysis is given in Table XXXVIII. In Table XXXVIII the most time-consuming parts in an iteration of the different methods are stated.

Table XXXVII stresses the need for different methods for local and global convergence. Only the iterative cubic method and the step controlled Newton–Raphson method can reliably be used to bring the iteration point to a point relatively close to the stationary point of interest. While the non-local part of the optimization can be reliably performed by just two methods, all four methods mentioned in Tables XXXVII and XXXVIII can, in principle, be used for the local optimization. We thus discuss separately the optimal methods for the nonlocal and the local optimization.

Typically some two or three constrained Newton–Raphson iterations are required to move the iteration point into the local region when the initial point is defined by a CI eigenvector with Hartree–Fock orbitals. Compared to this, only one iterative cubic iteration is typically required to bring the iteration point into the local region. It was demonstrated in Section VIII that usually only one micro-iteration is needed to follow the initial Newton–Raphson iteration in an iteration cubic iteration. This number of micro-iterations is sufficient when two Newton–Raphson iterations bring the calcula-

TABLE XXXVII
Summary of Convergence Characteristics of Different Optimization Methods

Characteristics	Method			
	Step size and sign controlled (constrained) Newton-Raphson	Iterative cubic	Update procedures	Fixed Hessian
Region of convergence	Almost always converges even when the Hessian has an incorrect structure	Usually converges even when the Hessian has an incorrect structure. May not need constraints	Usually converges from a region where Hessian has correct number of negative eigenvalues	Converges only from a point relatively close to stationary point
Effectiveness of method when far from convergence compared to Newton-Raphson	Unconstrained Newton-Raphson often converges very slowly or diverges. Properly constrained NR approach usually converges	One iterative cubic iteration approximates typically 2–5 constrained Newton-Raphson iterations	3–5 update iterations approximate typically 1 constrained Newton-Raphson iteration	Fixed Hessian techniques diverge if Newton-Raphson diverges
Local convergence	Second order convergence (in region with no constraints)	Third-order convergence	Superlinear convergence	Linear convergence
Effectiveness of method in local region compared to (unconstrained) Newton-Raphson	Second-order convergence in local region (no constraints applied)	One iterative cubic iteration approximates 1.3–1.6 Newton-Raphson iterations	2 update iterations approximate typically 1 Newton-Raphson iteration	$2^N - 1$ fixed Hessian iterations approximate typically N Newton-Raphson iterations

TABLE XXXVIII
Summary of Requirements of an Iteration in Different Optimization Methods

Parts of an Iteration	Method			
	Step controlled Newton-Raphson	Iterative Cubic	Update	Fixed Hessian
Type of transformation (see Appendix F for explanation of notation)	T_2	$T_3 + T_\mu$	T_1	T_1
Number of constructions of Hessian	1	2	0	0
Number of constructions of gradient	1	1	1	1
Number of triangularizations used for solving linear equations	1	2	0	0

tion to a region with pronounced local character. This assumption about the number of micro-iterations is tacitly used in Table XXXVIII.

We now compare these two nonlocal procedures: two Newton-Raphson iterations and one iterative cubic iteration consisting of the above "1+1" micro-iterations. One such iterative cubic iteration and two Newton-Raphson iterations is seen from Table XXXVIII to have different requirements only with respect to transformations. The iterative cubic iteration requires a T_3 transformation plus a T_μ transformation. Two Newton-Raphson iterations require two T_2 transformations. We refer to Appendix F for more information about the different transformations. From Tables F.1 and F.4 it is seen that the number of multiplications in $T_3 + T_\mu$ is about $0.75-0.86N^5$ when the number of occupied orbitals is half the total number of orbitals N . When $\frac{1}{3}$ of the set of orbitals corresponds to occupied orbitals the corresponding number of multiplications is about $0.23-0.24N^5$. For two T_2 transformations; one obtains operation counts $1.10N^5$ and $0.28N^5$ for the ratios $\frac{1}{2}$ and $\frac{1}{3}$, respectively. The number of operations in the transformations of an iterative cubic iteration is thus about 30% lower than the number of operations for the transformations of two Newton-Raphson iterations.

If it is of importance to reduce the integral transformations in the MCSCF optimization, the iterative cubic may be a real alternative to step controlled Newton-Raphson method for the global optimization. The usefulness of the iterative cubic method is furthermore increased by the use of subsequent "modified fixed Hessian" iterations [Eq. (285)] or "third-order

updates" corresponding to Eq. (293). The difference in efficacy between the two global methods is not great, and the step controlled Newton-Raphson method may be preferred as a global method because it is conceptually simpler and in many ways more flexible.

However, as we pointed out in Section VII, the constraints imposed on Newton-Raphson step-length amplitudes are usually imposed based somewhat on numerical experience. With the iterative cubic procedure our experience to date has indicated that with reasonable starting orbitals and CI coefficients often no constraints are necessary. Hence, particularly in very difficult cases, the iterative cubic approach may be useful since it may, in fact, replace several additional constrained Newton-Raphson iterations.

The optimal combination of optimization methods in the local region is not finally settled. All four methods in Tables XXXVII and XXXVIII will now be compared, and it will be shown that at least three of the methods discussed in Tables XXXVII and XXXVIII can be ingredients of a very fast local algorithm.

Our numerical experience indicates an iterative cubic iteration is typically equivalent to 1.3–1.6 Newton-Raphson iterations in the local region. One Newton-Raphson iteration is typically equivalent to two conventional update iterations. The number of fixed Hessian iterations corresponding to one Newton-Raphson depends on how "old" is the Hessian in use. The first two fixed Hessian iterations after a Newton-Raphson iteration correspond roughly to one Newton-Raphson iteration. A significantly larger number of fixed Hessian iterations are required to simulate more Newton-Raphson iterations as discussed in Section VI.C.

From Tables F.2 and F.3 in Appendix F it is seen that the two T_1 transformations required in two update or fixed Hessian iterations are more time consuming than one T_2 transformation which is required in one Newton-Raphson iteration. From Tables E.3 and E.4 (Appendix E) it is seen that the transformations $T_3 + T_\mu$ required in an iterative cubic iteration have a higher operation count than the 1.3 Newton-Raphson T_2 transformations required in the similar 1.3 Newton-Raphson iterations. In the local region the Newton-Raphson method thus is seen to require the fewest number of multiplications in the required integral transformation if one cubic iterative replaces 1.3 Newton-Raphson iterations. If one iterative cubic iteration replaces 1.6 Newton-Raphson iterations, then the transformations in the iterative cubic method appear slightly more efficient.

The differences in requirements of the integral transformations are not very large, and so other parts of an iteration must be studied in order to get reasonable assignments of the various methods' local efficiency. A conventional Newton-Raphson iteration requires the construction and triangulation of an Hessian. These are absent in fixed Hessian and update

methods, whereas they are required two times in an iterative cubic iteration. The iterative cubic method is thus often not an efficient local method, and we leave it out in the following discussion. The fixed Hessian and update methods are effective local methods. The slightly increased transformation times in these methods are usually outweighed by the absence of Hessian constructions and triangularizations.

The update methods are usually more efficient than fixed Hessian methods. If update methods are used in the local region it will often be efficient to use these methods alone to obtain convergence. Only in cases where slow convergence is observed should the Hessian approximation be reset by a new Newton–Raphson calculation.

Update methods usually are superior to the fixed Hessian methods. However, even the most stable update method can encounter instabilities in extreme cases. In these situations the fixed Hessian method is preferred, since the behavior of the method is determined by the initial Newton–Raphson iteration. The Newton–Raphson method is very stable in the local region.

The optimal number of fixed Hessian iterations per Newton–Raphson corresponds to the use of about 30% of cpu time on fixed Hessian iterations (transformations included) (see Section VI.C). Since a fixed Hessian iteration usually is at least three times faster than a Newton–Raphson iteration, the Newton–Raphson method should not be used alone in the local region. The procedure with a large number of fixed Hessian iterations is also not an effective choice.

The only remaining feature is how and when one should go from a global method to a local method. We now discuss the case where the iterative cubic method is the global method in use. It is our experience that the transition to a local method can be performed after the iterative cubic method converges to a point where the first-order corrected Hessian [Eq. (235)] has the correct structure, that is, the correct number of negative eigenvalues. The transition to a local method can therefore usually be done after one iterative cubic iteration. We then propose the use of a modified fixed Hessian iteration [Eq. (285)] or a third-order update iteration [Eq. (293)] as the initial local iteration. These methods are the only methods that reuse the third-order information obtained in an iterative cubic iteration.

D. Higher (Infinite) Order Procedures

In Section VIII it was demonstrated that an algorithm that optimizes a cubic energy approximation [Eq. (208)] is very stable. Other methods that include certain cubic, quartic, ... terms in κ have been developed.^{19, 29, 30, 90} We will now describe these. Since these methods' basic aim is to improve the orbital optimization, we will focus on this and for simplicity delete the coupling terms connected to the state optimization.

Werner and Meyer^{29,30} have introduced a method in which the energy is expanded through second order in $e^{i\kappa} - 1$:

$$\begin{aligned}
 E(\kappa) &= \langle 0 | e^{-i\hat{K}} H e^{i\hat{K}} | 0 \rangle \\
 &= h_{rs} \langle 0 | e^{-i\hat{K}} a_{r\sigma}^+ e^{i\hat{K}} e^{-i\hat{K}} a_{s\sigma} e^{i\hat{K}} | 0 \rangle \\
 &\quad + \frac{1}{2} (\langle pq | rs \rangle \langle 0 | e^{-i\hat{K}} a_{p\sigma}^+ e^{i\hat{K}} e^{-i\hat{K}} a_{r\sigma}^+ e^{i\hat{K}} e^{-i\hat{K}} a_{s\sigma} e^{i\hat{K}} e^{-i\hat{K}} a_{q\sigma} e^{i\hat{K}} | 0 \rangle) \\
 &= E^{\text{WM}} = E_0 + 2(e^{i\kappa} - 1)_{rk} h_{rs} \rho_{sk}^{(1)} + 2(e^{i\kappa} - 1)_{pk} (\langle pq | rs \rangle \rho_{kq, rs}^{(2)}) \\
 &\quad + (e^{i\kappa} - 1)_{rk} (e^{i\kappa} - 1)_{sl} \rho_{kl}^{(1)} h_{rs} + 2(e^{i\kappa} - 1)_{pk} (e^{i\kappa} - 1)_{rl} \rho_{kq, ls}^{(2)} (\langle pq | rs \rangle) \\
 &\quad + (e^{i\kappa} - 1)_{pk} (e^{i\kappa} - 1)_{ql} \rho_{kl, rs}^{(2)} (\langle pq | rs \rangle)
 \end{aligned} \tag{295}$$

In Eq. (295) the usual nonrelativistic Hamiltonian is assumed, and $\rho^{(1)}$ and $\rho^{(2)}$ are the one-electron density matrix and symmetric two-electron density matrix.

Werner and Meyer express E^{WM} [Eq. (295)] by introducing the following matrices

$$\begin{aligned}
 F_{rs}^{pq} &= \rho_{pq}^{(1)} h_{rs} + \rho_{pq, kl}^{(2)} (kl | rs) \\
 G_{rs}^{pq} &= F_{rs}^{pq} + 2\rho_{pk, ql}^{(2)} (kr | ls),
 \end{aligned} \tag{296}$$

where p and q correspond to occupied orbitals. By denoting

$$(e^{i\kappa} - 1)_{kl} = \Delta_{kl}$$

Eq. (295) becomes

$$E^{\text{WM}} = E_0 + 2F_{rs}^{ps}\Delta_{rp} + G_{rs}^{pq}\Delta_{rp}\Delta_{sq} \tag{297}$$

The energy approximation Eq. (297) includes terms cubic, quartic, ... in κ due to the expansion of Δ :

$$\Delta = i\kappa - \frac{1}{2}\kappa^2 - \frac{i}{6}\kappa^3 + \frac{1}{24}\kappa^4 + \dots \tag{298}$$

However, there are also cubic, quartic, ... terms in Eq. (295) which are missing. Since the basic variable κ occurs as $e^{i\kappa}$, E^{WM} is periodical in κ . The approximate energy E^{WM} therefore always has a minimum. The second-order energy functional Eq. (19) defining the uncontrolled Newton–Raphson iteration does not always have a minimum, but the restricted Newton–Raphson problem [Eq. (110)] always does.

In a given iteration the \mathbf{F}^{rs} and \mathbf{G}^{rs} matrices are constructed and a "point" κ is found where E^{WM} is stationary; that is,

$$\frac{\partial E^{\text{WM}}}{\partial \kappa_{ij}} = 2F_{rs}^{ps} \frac{\partial \Delta_{rp}}{\partial \kappa_{ij}} + 2G_{rs}^{pq} \frac{\partial \Delta_{rp}}{\partial \kappa_{ij}} \Delta_{sq} = 0 \quad (299)$$

The derivatives of Δ occurring in Eq. (299) can be expressed by the generators \mathbf{E}_{kl} of the orthogonal matrix group.⁹¹ The only nonzero elements of \mathbf{E}_{kl} are

$$(\mathbf{E}_{kl})_{kl} = 1, \quad (\mathbf{E}_{kl})_{lk} = -1 \quad (300)$$

By performing algebraic manipulations very similar to those in Appendix B, one obtains

$$\frac{\partial}{\partial \kappa_{ij}} (e^{i\kappa} - 1) = e^{i\kappa} \left(i\mathbf{E}_{ij} - \frac{1}{2} [\kappa, \mathbf{E}_{ij}] - \frac{i}{3!} [\kappa, [\kappa, \mathbf{E}_{ij}]] + \dots \right) \quad (301)$$

In Eq. (301) commutators between matrices are introduced. The exact derivative Eq. (301) can easily be obtained by diagonalizing κ in analogy with derivatives in Appendix B. However, since E^{WM} has errors starting in third order, Eq. (301) does not need to be evaluated to high accuracy. Direct use of the first three or four terms of Eq. (301) therefore gives a satisfactory approximation to the derivative of Δ . Werner and Meyer solve Eq. (299) with approximate derivatives of Δ by a modified Gauss-Seidel procedure.²⁹ They report that satisfactory convergence is obtained if different "tricks" to improve convergence are used.

After having outlined the basic ideas of the method proposed by Werner and Meyer, we now discuss how the construction of \mathbf{F}^{rs} and \mathbf{G}^{rs} [Eq. (296)] can be replaced by the construction of matrices that are easier to construct.

The construction of the \mathbf{F}^{rs} and \mathbf{G}^{rs} matrices [Eq. (296)] requires a two-electron integral transformation equal to the transformation required by other second-order procedures. In Section IX.A it was shown that the two-electron integral transformation used in the Newton-Raphson method can be simplified. A similar simplification in the summations over secondary orbitals can be obtained in connection with the method of Werner and Meyer. Introduce the tilde ($\tilde{\cdot}$) basis as in Eq. (250) with

$$\mathbf{P}^{\text{II}} = \mathbf{1}, \quad \mathbf{P}^{\text{SS}} = \mathbf{C}^{\text{SS}-1} \quad (302)$$

and perform the corresponding integral transformation. This gives rise to

modified **F** and **G** matrices

$$\begin{aligned}\tilde{F}_{rs}^{pq} &= \rho_{pq}^{(1)} \tilde{h}_{rs} + \rho_{pq,kl}^{(2)} (\tilde{k}\tilde{l}| \tilde{r}\tilde{s}) \\ \tilde{G}_{rs}^{pq} &= \tilde{F}_{rs}^{pq} + 2\rho_{pk,ql}^{(2)} (\tilde{k}\tilde{r}| \tilde{l}\tilde{s})\end{aligned}\quad (303)$$

Since the upper indices in \tilde{F} and \tilde{G} correspond to occupied orbitals [see Eq. (296)], which are the same in the molecular orbital basis and in the tilde basis, one obtains

$$\begin{aligned}F_{rs}^{pq} &= \tilde{F}_{kl}^{pq} (\mathbf{P}^{-1})_{kr} (\mathbf{P}^{-1})_{ls} \\ G_{rs}^{pq} &= \tilde{G}_{kl}^{pq} (\mathbf{P}^{-1})_{kr} (\mathbf{P}^{-1})_{ls}\end{aligned}\quad (304)$$

where

$$\mathbf{P} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}^{\text{SS}} \end{pmatrix} \quad (305)$$

If Eq. (304) is introduced in the energy approximation [Eq. (297)], one obtains

$$E^{\text{WM}} = E_0 + 2\tilde{F}_{mn}^{ps} (\mathbf{P}^{-1})_{mr} (\mathbf{P}^{-1})_{ns} \Delta_{rp} + \tilde{G}_{mn}^{pq} (\mathbf{P}^{-1})_{mr} (\mathbf{P}^{-1})_{ns} \Delta_{rp} \Delta_{sq} \quad (306)$$

In the first summation in Eq. (306) $(\mathbf{P}^{-1})_{ns}$ equals δ_{sn} , since s corresponds to an occupied orbital (s occurs as upper index in \tilde{F}_{mn}^{ps}). Equation (306) then becomes

$$E^{\text{WM}} = E_0 + 2\tilde{F}_{ms}^{ps} \tilde{\Delta}_{mp} + \tilde{G}_{mn}^{pq} \tilde{\Delta}_{mp} \tilde{\Delta}_{nq} \quad (307)$$

where

$$\tilde{\Delta}_{lk} = (\mathbf{P}^{-1} \Delta)_{lk} \quad (308)$$

By introducing a new set of variables [Eq. (308)], integrals are thus only needed in the tilde basis. As discussed in Section IX.A and Appendix F, this simplifies transformations significantly.

The method of Werner and Meyer is thus not significantly more involved than the Newton–Raphson method. Higher order terms are included. However, it is not obvious that the cubic and higher order terms included are the

most essential cubic and higher order terms, so theoretical predication about the method are not easy to make.

The reported numerical experience^{29,30} with the Werner–Meyer method is now summarized. In the nonlocal region, Werner and Meyer report improved convergence can be sometimes obtained by using their procedure without coupling. However, similar improvements are sometimes observed in the Newton–Raphson technique by not including coupling for the first few iterations.^{13–15} Several iterations with the Werner–Meyer procedure are usually required to enter the local region.^{29,30} This is in contrast with the iterative cubic method discussed in Section VII. In the local region, the inclusion of Werner and Meyer's higher order terms are reported to often slow convergence.^{29–30}

Line search methods can exactly optimize a function in a subspace defined by a single variable. These methods are very stable, and only minor modifications have to be introduced in order to guarantee global convergence. We now discuss a method proposed by Igawa et al.^{19,90} that optimizes the exact function in a subspace defined by *several* variables.

A diagonal $\hat{\kappa}$ operator is defined

$$\hat{\kappa}^d = \sum_A \gamma_A \hat{n}_A, \quad \hat{n}_A = \sum_\sigma a_{A\sigma}^\dagger a_{A\sigma} \quad (\text{no Einstein summation}) \quad (309)$$

If the basis in which $\hat{\kappa}$ is assumed diagonal were the usual molecular orbital basis, the *Ansatz* [Eq. (309)] introduces only a phase factor in the trial function. By choosing another basis, that is, another set of number operators $\{\hat{n}_A\}$, nontrivial variations are introduced by Eq. (309). The maximal dimension of the space

$$|0^d\rangle = e^{i\hat{\kappa}^d} |0\rangle \quad (310)$$

can be shown¹⁹ to be

$$K = \min(N, 2n) \quad (311)$$

where N is the total number of orbitals and n is the number of occupied orbitals. Since K is usually smaller than the dimension of the complete orbital optimization manifold, the *Ansatz* Eq. (310) is only a subspace of the complete MCSCF manifold.

The energy corresponding to the ansatz Eq. (310) is

$$\begin{aligned} E^d &= \langle 0 | e^{-i\hat{\kappa}^d} H e^{i\hat{\kappa}^d} | 0 \rangle \\ &= \langle 0 | H | 0 \rangle + \langle 0 | e^{i\hat{\kappa}^d} [e^{-i\hat{\kappa}^d}, H] | 0 \rangle \end{aligned} \quad (312)$$

It will now be shown that a compact expression for the energy Eq. (312) can

be derived for an arbitrary \hat{K}^d [Eq. (309)]. Since the number operators commute, the exponential operator in Eq. (310) can be factorized

$$e^{i\sum_A \gamma_A \hat{n}_A} = \prod_{A=1, K} e^{i\gamma_A \hat{n}_A} \quad (\text{no Einstein summation}) \quad (313)$$

The relation

$$e^{i\gamma_r \hat{n}_r} = (1 - \hat{n}_r) + e^{i\gamma_r} \hat{n}_r \quad (\text{no sum}) \quad (314)$$

is easily derived. By using Eqs. (313) and (314) and commutators for fermions, the following expressions are derived:

$$e^{-i\sum_B \gamma_B \hat{n}_B} \left[\sum_{\sigma} a_{r\sigma}^+ a_{s\sigma}, e^{i\sum_A \gamma_A \hat{n}_A} \right] = (e^{-i(\gamma_r - \gamma_s)} - 1) \sum_{\sigma} a_{r\sigma}^+ a_{s\sigma} \quad (\text{no Einstein summation}) \quad (315a)$$

$$\begin{aligned} & e^{-i\sum_B \gamma_B \hat{n}_B} \left[\sum_{\sigma, \sigma'} a_{r\sigma}^+ a_{s\sigma'}^+, a_{t\sigma'} a_{u\sigma}, e^{i\sum_A \gamma_A \hat{n}_A} \right] \\ &= (e^{-i(\gamma_r + \gamma_s - \gamma_t - \gamma_u)} - 1) \sum_{\sigma, \sigma'} a_{r\sigma}^+ a_{s\sigma'}^+ a_{t\sigma'} a_{u\sigma} \quad (\text{no Einstein summation}) \end{aligned} \quad (315b)$$

The exact energy (Eq. (312)) with the nonrelativistic Born–Oppenheimer Hamiltonian expressed in the basis where \hat{K}^d is diagonal becomes

$$\begin{aligned} E^d &= \sum_{A, B}^K h_{AB} e^{-i(\gamma_A - \gamma_B)} \rho_{AB}^{(1)} \\ &+ \frac{1}{2} \sum_{A, B, C, D}^K (AB|CD) e^{-i(\gamma_A + \gamma_C - \gamma_B - \gamma_D)} \rho_{AB, CD}^{(2)} \end{aligned} \quad (316)$$

With the conventional MCSCF *Ansatz* [Eq. (11)] the energy dependence of orbital rotations is

$$\begin{aligned} E &= \langle 0 | e^{-i\hat{K}} H e^{i\hat{K}} | 0 \rangle \\ &= h_{rs} (e^{i\kappa})_{r'r} (e^{i\kappa})_{s's} \rho_{r's}^{(1)} \\ &+ \frac{1}{2} (rs|tu) (e^{i\kappa})_{r'r} (e^{i\kappa})_{s's} (e^{i\kappa})_{t't} (e^{i\kappa})_{u'u} \rho_{r't', s'u'}^{(2)} \end{aligned} \quad (317)$$

While the general energy expression, Eq. (317), involves eightfold sum-

mations for $\hat{\kappa} \neq \hat{0}$, a nonzero $\hat{\kappa}^d$ in Eq. (316) only introduces fourfold summations and phasefactors. The exact energy [Eq. (316)] for an arbitrary operator $\hat{\kappa}^d$ [Eq. (309)] is thus easily obtained. The same is true for derivatives with respect to the γ parameters. The exact energy can thus easily be optimized in the subspace defined by Eqs. (309) and (310) once the Hamiltonian is known in this basis. In the method of Igawa et al.¹⁹ a given iteration first defines a basis where $\hat{\kappa}^d$ is diagonal, then the transformation of integrals to this basis is performed, and finally the exact energy is optimized in this space.

The basis in which $\hat{\kappa}^d$ is diagonal is chosen as follows. First an usual MCSCF iteration with, for example, the constrained Newton-Raphson method or the conjugate gradient technique is performed. This gives a reference operator κ^{ref} .

$$\hat{\kappa}^{\text{ref}} = \kappa_{rs}^{\text{ref}} a_{ro}^+ a_{so} \quad (318)$$

The matrix κ^{ref} defined by Eq. (318) is then diagonalized

$$\kappa^{\text{ref}} \mathbf{U} = i \mathbf{U} \gamma^{\text{ref}} \quad (319)$$

The eigenvectors of κ^{ref} are then used to define the basis in which $\hat{\kappa}^d$ is constructed

$$a_A^+ = a_r^+ U_{rA} \quad (320)$$

where $\langle a_r^+ \rangle$ denotes the molecular orbital basis. The $\hat{\kappa}^d$ [Eq. (312)] is then defined as

$$\hat{\kappa}^d = (\Delta\gamma_A + \gamma_A^{\text{ref}}) a_{A\sigma}^+ a_{A\sigma} = \gamma_A \hat{n}_A \quad (321)$$

A subspace around the reference point is thus constructed.

The step lengths are defined only by the γ parameters, and these are optimized exactly (i.e., no constraints on the exact γ 's are required). It is thus expected that the method is particularly effective in the nonlocal region, where most other methods need an empirically based constraint scheme. It is not likely that the exact optimization in the subspace will improve local convergence, just as line search is not advisable locally.⁹

Experience with the method¹⁹ stresses these points. No constraints for the γ 's are needed in the nonlocal region in order to obtain convergence. The actual rate of convergence in the nonlocal region depends often on the selected subspace. If the eigenvectors determining the subspace are obtained from an undamped Newton-Raphson iteration, convergence is often slow but steady. If a κ matrix from a constrained Newton-Raphson iteration is used to define the subspace, convergence is faster. The exact optimi-

zation in a subspace can then reduce the total number of iterations in the nonlocal region by approximately 1 or more. In the local region, the exact optimization in a subspace does not improve convergence.

The extra computational work due to the exact optimization in a subspace is the extra transformation, which is required to construct the Hamiltonian in the "diagonal basis." This transformation required about $(K/2)N^4$ operations. Since K is usually equal to $2n$ [Eq. (311)], the operation count for this transformation is about nN^4 . This transformation has a higher multiplication count than the reduced transformations needed in a second-order iteration (Section IX.A).

This infinite order method is very time consuming. However, the exact optimization of the energy in a subspace makes the method extremely stable and reliable. The method is thus expected to be useful when the energy function changes so abruptly that exact information about the energy at many points in a subspace is needed in order to properly define a step.

Conventional line search in the nonlocal region has been investigated with promising results.⁹ Comparison studies between the optimization in a subspace as discussed and more conventional line search procedures are currently under way and will be reported soon.⁹²

X. SUMMARY AND CONCLUSIONS

Current theoretical research in MCSCF primarily involves the study of proper characterization of a multiconfigurational state and the development of efficient and reliable optimization procedures. Until recent developments with second- and higher order MCSCF schemes, it was often very difficult to converge to the correct state of interest.^{1,2} Previously, the characteristics of a multiconfigurational state were usually not properly examined, and many reported calculations undoubtedly converged to an undesired stationary point on the energy hypersurface. Since we have already discussed at length, particularly in Section IX, effective implementation and combination of methods, in this section we will only briefly restate a few main points.

* We have discussed in Section IV some desired characteristics of an MCSCF state. If converging to the n th state in energy of a certain symmetry, these include the following:

1. The Hessian has $n - 1$ negative eigenvalues.
2. The MCTDHF calculation using the MCSCF state as a reference state is stable. In the symmetry block of the MCSCF reference state the MCTDHF should have $n - 1$ negative energy differences.
3. In the CI using the MCSCF configuration state functions, the state of interest is the n th state.

Since the MCSCF state includes a finite (usually small) number of configurations and the orbitals are usually expanded using a finite basis set, some (or, in rare cases, all) of these criteria may not be fulfilled for a converged MCSCF state that represents the n th state—in which case it is usually most important to fulfill the criteria that are of particular relevance to subsequent calculations which use the MCSCF state and orbitals. As has been demonstrated through both theory and calculations, characteristic 3 (above) is the one most often not fulfilled and in many cases is not essential (e.g., especially since reparameterization may cause a converged MCSCF state to change from being the $(n - 1)$ st to the n th CI state).

The MCSCF convergence problem differs significantly in the local and the nonlocal regions. When far from convergence (the global convergence problem) it is most important for a technique to move a calculation to a region on the energy hypersurface close to the proper stationary point. Either constrained Newton–Raphson or iterative cubic methods may be used for the global convergence problem. Constrained Newton–Raphson approaches usually seem to work well.^{13, 15, 18} Recent calculations with the Fletcher constraint algorithm which guarantees convergence for the lowest state of a certain symmetry are extremely promising.⁵⁵ The iterative cubic technique is apparently both efficient and reliable.²⁰ Iterative cubic calculations have also demonstrated the validity of parameters chosen by numerical experience for the mode-controlling technique. “Infinite order” techniques are interesting and appear useful.^{19, 29, 30} However, exactly what is occurring theoretically and calculationally with these is much less clear (e.g., it is not at all certain that the most important variables are taken to infinite order, and calculations are quently performed without coupling). Constraints are still usually required with these procedures when far from convergence. Augmented Hessian approaches^{31, 32, 86} often also require constraints when far from convergence and currently do not appear to offer any significant advantages over adequately and properly constrained Newton–Raphson approaches.¹⁸

When in the local region, the Hessian has the proper number of negative eigenvalues and Newton–Raphson step length amplitudes are not large. Convergence with the Newton–Raphson approach is rapid and reliable. In this region computational efficiency is most important. Current theory and numerical experience indicates that an update procedure such as the Broyden rank-1 method should be used.^{21, 75} A combination of fixed Hessian and Newton–Raphson approaches is also useful, provided no more than about 30% of the CPU time (including transformation time) is used on fixed Hessian approaches.²² Infinite order approaches are not, in general, advocated in the local region.

Often wavefunction optimization approaches have been based almost entirely on numerical experience. Occasionally this has led to disasterous consequences. We and others have recently explicitly demonstrated, in this paper

and elsewhere, how both detailed theoretical and computational studies should be combined in order to properly design the most efficient MCSCF procedures which reliably converge to the correct stationary point on the energy hypersurface.

APPENDIX A: REDUNDANT VARIABLES

For real orbitals, a general unitary transformation of the reference state $|0\rangle$ may be written as

$$|\tilde{0}\rangle = \exp i\hat{\kappa} \exp i\hat{S} |0\rangle \quad (\text{A.1})$$

where $\hat{\kappa}$ and \hat{S} are defined in Eqs. (5) and (7). The excitation operators $\langle a_r^+ a_s - a_s^+ a_r \rangle$ in $\hat{\kappa}$ and $\langle |k\rangle\langle 0| - |0\rangle\langle k| \rangle$ in \hat{S} may span a basis with linear dependencies. When optimizing a state or when evaluating the response of an MCSCF state to an external one-electron perturbation, such linear dependencies must be eliminated.

The generalized Brillouin's theorem (GBT)^{10–12, 78}

$$\langle 0 | [a_r^+ a_s, H] | 0 \rangle = 0 \quad (\text{A.2})$$

$$\langle 0 | [|k\rangle\langle 0|, H] | 0 \rangle = \langle 0 | H | k \rangle = 0 \quad (\text{A.3})$$

is derived by considering first-order variations of in the reference state $|0\rangle$. This variational space is embedded in the space spanned by

$$\langle |0\rangle; |k\rangle; (a_r^+ a_s - a_s^+ a_r) |0\rangle, \quad r > s \rangle = \langle |j\rangle; (a_r^+ a_s - a_s^+ a_r) |0\rangle, \quad r > s \rangle \quad (\text{A.4})$$

Because the orbitals are real, Eq. (A.2) may be written as

$$\langle 0 | H (a_r^+ a_s - a_s^+ a_r) | 0 \rangle = 0 \quad (\text{A.5})$$

Operators $a_{cl}^+ a_{cl'}$ or $a_p^+ a_{p'}$, where here cl, cl' refer to orbitals doubly occupied in all MCSCF CI configurations and p, p' refer to unoccupied orbitals, trivially fulfill the generalized Brillouin's theorem at all points on the energy hypersurface; that is,

$$\langle 0 | [a_{cl}^+ a_{cl'}, H] | 0 \rangle = 0 \quad (\text{A.6})$$

$$\langle 0 | [a_p^+ a_{p'}, H] | 0 \rangle = 0 \quad (\text{A.7})$$

These operators are trivially redundant and can be easily eliminated immediately from an MCSCF calculation.

It is more complicated to eliminate operators due to linear dependencies in the space spanned by the sets $\langle a_r^+ a_s - a_s^+ a_r \rangle$ and $\langle |k\rangle\langle 0| - |0\rangle\langle k| \rangle$ [Eq.

(A.4)]. A linear dependency in the set (A.4) may be determined if

$$\begin{aligned}
 (a_x^+ a_y - a_y^+ a_x) |0\rangle &= |0\rangle \langle 0| a_x^+ a_y - a_y^+ a_x |0\rangle + \sum_{k \neq 0} |k\rangle \langle k| a_x^+ a_y - a_y^+ a_x |0\rangle \\
 &\quad + \sum_{(r,s) \neq (x,y)} (a_r^+ a_s - a_s^+ a_r) |0\rangle \\
 &\quad \times \langle 0| (a_r^+ a_s - a_s^+ a_r) (a_x^+ a_y - a_y^+ a_x) |0\rangle
 \end{aligned} \quad (\text{A.8})$$

These operators, $a_x^+ a_y - a_y^+ a_x$ in Eq. (A.8), are thus redundant in the sense that, if the GBT in Eqs. (A.2) and (A.3) is satisfied for the nonredundant set of operators (i.e., without $a_x^+ a_y - a_y^+ a_x$), the GBT will be automatically satisfied for the redundant set of operators (i.e., the set which includes $a_x^+ a_y - a_y^+ a_x$). The operators $a_{cl}^+ a_{cl'} - a_{cl'}^+ a_{cl}$ and $a_p^+ a_{p'} - a_{p'}^+ a_p$ in Eqs. (A.6) and (A.7) trivially fulfill Eq. (A.8). However, there are other operators which may also fulfill Eq. (A.8) and should not be included in the calculations.

The states $\langle |0\rangle, |k\rangle \rangle = \langle |j\rangle \rangle$ are related to the configuration space $\langle |\phi_g\rangle \rangle$ through a unitary transformation. Hence, Eq. (A.8) may be reexpressed as

$$\begin{aligned}
 (a_x^+ a_y - a_y^+ a_x) |0\rangle &= \sum_g |\phi_g\rangle \langle \phi_g| a_x^+ a_y - a_y^+ a_x |0\rangle \\
 &\quad + \sum_{(r,s) \neq (x,y)} (a_r^+ a_s - a_s^+ a_r) |0\rangle \\
 &\quad \times \langle 0| (a_r^+ a_s - a_s^+ a_r) (a_x^+ a_y - a_y^+ a_x) |0\rangle
 \end{aligned} \quad (\text{A.9})$$

If the last term in Eq. (A.9) is neglected, we determine only linear dependences between the operator $a_x^+ a_y - a_y^+ a_x$ and $\langle |0\rangle, |k\rangle \rangle$. Only if both terms in Eq. (A.9) are included are all redundant operators of an MCSCF calculation eliminated. In Ref. 17 examples are given where the last term in Eq. (A.9) is required in order to eliminate all redundant variables of the MCSCF calculation.

An operator that fulfills Eq. (A.8) usually gives a zero eigenvalue only at a stationary point. This can easily be demonstrated since

$$\begin{aligned}
 \langle 0| [a_t^+ a_u - a_u^+ a_t, H, a_x^+ a_y - a_y^+ a_x] |0\rangle \\
 = \frac{1}{2} \langle 0| [[a_t^+ a_u - a_u^+ a_t, H], a_x^+ a_y - a_y^+ a_x] |0\rangle \\
 + \frac{1}{2} \langle 0| [a_t^+ a_u - a_u^+ a_t, [H, a_x^+ a_y - a_y^+ a_x]] |0\rangle
 \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned}
 &= \langle 0| [[a_t^+ a_u - a_u^+ a_t, H], a_x^+ a_y - a_y^+ a_x] |0\rangle \\
 &\quad - \frac{1}{2} \langle 0| [[a_t^+ a_u - a_u^+ a_t, a_x^+ a_y - a_y^+ a_x], H] |0\rangle \\
 &= 2\langle 0| [a_t^+ a_u - a_u^+ a_t, H] (a_x^+ a_y - a_y^+ a_x) |0\rangle \\
 &\quad - (\delta_{ux} \langle 0| [a_t^+ a_y, H] |0\rangle - \delta_{yt} \langle 0| [a_x^+ a_u, H] |0\rangle \\
 &\quad - \delta_{xt} \langle 0| [a_u^+ a_y, H] |0\rangle + \delta_{yu} \langle 0| [a_x^+ a_t, H] |0\rangle)
 \end{aligned} \quad (\text{A.11})$$

and

$$\begin{aligned} \langle 0 | [|k\rangle\langle 0| - |0\rangle\langle k|, H, a_x^+ a_y - a_y^+ a_x] | 0 \rangle \\ = -2 \langle 0 | [H, a_x^+ a_y - a_y^+ a_x] | k \rangle \end{aligned} \quad (\text{A.12})$$

$$= -2 \langle 0 | H (a_x^+ a_y - a_y^+ a_x) | k \rangle + 2 \langle 0 | (a_x^+ a_y - a_y^+ a_x) H | k \rangle \quad (\text{A.13})$$

At a stationary point (i.e., where the GBT, Eq. (A.2), is fulfilled) the last four terms in Eq. (A.11) are zero. If we assume that Eq. (A.8) is valid when $|0\rangle$ is replaced by $|j\rangle$ it can easily be seen that at a stationary point all the Hessian elements involving $a_x^+ a_y - a_y^+ a_x$ are composed of linear combinations of Hessian elements which do not involve $a_x^+ a_y - a_y^+ a_x$. Hence in this case, if the redundant operator $a_x^+ a_y - a_y^+ a_x$ is included in the operator set, a zero eigenvalue of the Hessian results. This is generally true *only* at convergence. An example of this behavior is given in Section V.B. If the redundant operator is included in the set, serious convergence problems may result since, for example, the Hessian eigenvalues may approach zero more rapidly than do the corresponding redundant GBT amplitudes and large step length amplitudes.

APPENDIX B: THE FIRST PARTIAL DERIVATIVE OF THE TOTAL ENERGY

We consider in the following how the first partial derivative (the gradient) of the energy

$$E\left({}^0\underline{\kappa} + {}^0\mathbf{S} + \mathbf{S}\right) = \langle 0 | \exp(-i\hat{S}) \exp(-i\hat{\kappa}) H \exp(i\hat{\kappa}) \exp(i\hat{S}) | 0 \rangle \quad (\text{B.1})$$

where

$$\hat{\kappa} = i \sum_{r > s} ({}^0\kappa_{rs} + \kappa_{rs})(a_r^+ a_s - a_s^+ a_r) \quad (\text{B.2})$$

$$\hat{S} = i \sum_{k \neq 0} ({}^0S_{k0} + S_{k0})(|k\rangle\langle 0| - |0\rangle\langle k|) \quad (\text{B.3})$$

may be evaluated at an arbitrary point (${}^0\underline{\kappa}, {}^0\mathbf{S}$). Such a derivative becomes important when deriving MCSCF iterative approaches that carry out a linear transformation among the variables between each step of the iterative procedure. Initially we consider how the differential $\exp[-(\hat{\lambda} + d\hat{\lambda})] -$

$\exp(-\hat{\lambda})$, where $\hat{\lambda}$ is an arbitrary operator, may be evaluated:

$$\begin{aligned} \exp[-(\hat{\lambda} + d\hat{\lambda})] - \exp(-\hat{\lambda}) &= \exp(-\hat{\lambda}) [\exp(\hat{\lambda}) \exp(-(\hat{\lambda} + d\hat{\lambda}))]^{-1} \\ &= \exp(-\hat{\lambda}) \exp(\hat{\lambda}z) \exp[-(\hat{\lambda} + d\hat{\lambda})z] |_{z=1} \\ &\quad - \exp(\hat{\lambda}z) \exp[-(\hat{\lambda} + d\hat{\lambda})z] |_{z=0} \\ &= \exp(-\hat{\lambda}) \int_0^1 dz \frac{d}{dz} [\exp(\hat{\lambda}z) \exp(-(\hat{\lambda} + d\hat{\lambda})z)] \\ &= \exp(-\hat{\lambda}) \int_0^1 dz \exp(\hat{\lambda}z) (\hat{\lambda} - \hat{\lambda} - d\hat{\lambda}) \exp[-(\hat{\lambda} + d\hat{\lambda})z] \end{aligned} \quad (\text{B.4})$$

Since we wish to determine the differential $d(\exp(-\hat{\lambda}))$, all terms which are not linear in $d\hat{\lambda}$ may be neglected. We get

$$d(\exp(-\hat{\lambda})) = -\exp(-\hat{\lambda}) \int_0^1 dz \exp(\hat{\lambda}z) d\hat{\lambda} \exp(-\hat{\lambda}z) \quad (\text{B.5})$$

Defining a superoperator $\hat{\lambda}$ corresponding to $\hat{\lambda}$

$$\hat{\lambda} = [\hat{\lambda}, f] \quad (\text{B.6})$$

(the $\hat{\lambda}$ notes the superoperator) for an arbitrary operator f gives

$$\begin{aligned} d(\exp(-\hat{\lambda})) &= -\exp(-\hat{\lambda}) \int_0^1 dz \left(\sum_{n=0}^{\infty} \frac{(z)^n}{n!} \hat{\lambda}^n d\lambda \right) \\ &= -\exp(-\hat{\lambda}) \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \hat{\lambda}^n d\lambda \end{aligned} \quad (\text{B.7})$$

If $\hat{\lambda}$ is a sum of certain operators O_i with expansion coefficients L_i

$$\hat{\lambda} = \sum_i (^0L_i + L_i) O_i$$

we have from Eq. (B.7) that

$$\frac{\partial \exp(-\hat{\lambda})}{\partial L_i} = -\exp(-\hat{\lambda}) \left(\sum_n \frac{1}{(n+1)!} \hat{\lambda}^n O_i \right) \quad (\text{B.8})$$

The relation

$$\frac{\partial \exp(\hat{\lambda})}{\partial L_i} = \left(\sum_n \frac{1}{(n+1)!} \hat{\lambda}^n O_i \right) \exp(\hat{\lambda}) \quad (\text{B.9})$$

may be derived in an equivalent way.

B.1. Derivative with Respect to Orbital Parameters

To determine the partial derivative of the total energy with respect to the orbital parameters, it becomes convenient to introduce the notation

$${}^0\hat{\kappa} = i \sum_{r>s} {}^0\kappa_{rs} (a_r^+ a_s - a_s^+ a_r) \quad (\text{B.10})$$

$${}^0\hat{S} = i \sum_{k \neq 0} {}^0S_{k0} (|k\rangle\langle 0| - |0\rangle\langle k|) \quad (\text{B.11})$$

$${}^0\hat{\kappa} a_r^+ a_s = [{}^0\hat{\kappa}, a_r^+ a_s] \quad (\text{B.12})$$

We then obtain

$$\begin{aligned} F({}^0\lambda)_{rs} &= \left. \frac{\partial E({}^0\kappa - {}^0\hat{\kappa}, {}^0\mathbf{S} + \hat{\mathbf{S}})}{\partial \kappa_{rs}} \right|_{{}^0\kappa = 0} \\ &= \langle 0 | \exp(-i{}^0\hat{S}) \left(\sum_n \frac{(-i)^n}{(n+1)!} ({}^0\hat{\kappa}^n (a_r^+ a_s - a_s^+ a_r)) \right) \\ &\quad \times \exp(-i{}^0\hat{\kappa}) H \exp(i{}^0\hat{\kappa}) \exp(i{}^0\hat{S}) | 0 \rangle \\ &\quad - \langle 0 | \exp(-i{}^0\hat{S}) \exp(-i{}^0\hat{\kappa}) H \exp(i{}^0\hat{\kappa}) \\ &\quad \times \left(\sum_n \frac{(-i)^n}{(n+1)!} ({}^0\hat{\kappa}^n (a_r^+ a_s - a_s^+ a_r)) \right) \exp(i{}^0\hat{S}) | 0 \rangle \\ &= \langle \tilde{0} | \left[\sum_n \frac{(-i)^n}{(n+1)!} ({}^0\tilde{\kappa}^n (\tilde{a}_r^+ \tilde{a}_s - \tilde{a}_s^+ \tilde{a}_r)), H \right] | \tilde{0} \rangle \\ &= 2 \langle \tilde{0} | \left[\sum_n \frac{(-i)^n}{(n+1)!} ({}^0\tilde{\kappa}^n \tilde{a}_r^+ \tilde{a}_s), H \right] | \tilde{0} \rangle \end{aligned} \quad (\text{B.13})$$

where

$$\tilde{a}_r^+ = \exp(i{}^0\hat{\kappa}) a_r^+ \exp(-i{}^0\hat{\kappa}) \quad (\text{B.14})$$

$$|\tilde{0}\rangle = \exp(i{}^0\hat{\kappa}) \exp(i{}^0\hat{S}) | 0 \rangle \quad (\text{B.15})$$

Since

$$(-i)^n {}^0\tilde{\kappa}^n \tilde{a}_r^+ \tilde{a}_s = \sum_{pq} C_{rs,pq}^n \tilde{a}_p^+ \tilde{a}_q \quad (\text{B.16})$$

we get

$$F({}^0\lambda)_{rs} = 2 \sum_n \frac{1}{(n+1)!} \sum_{pq} C_{rs,pq}^n \langle \tilde{0} | [\tilde{a}_p^+ \tilde{a}_q, H] | \tilde{0} \rangle \quad (\text{B.17})$$

and the energy gradient elements that correspond to a linear transformation among the variables thus become a sum of the ordinary gradient elements $2\langle \tilde{0}|[\tilde{a}_p^+ \tilde{a}_q, H]|\tilde{0}\rangle$. The evaluation of $F(^0\lambda)_{rs}$ through determining $C_{rs,pq}^n$ is very cumbersome if high accuracy is required, and we describe in the following how Eq. (B.16) may be rewritten to allow an easy high-accuracy evaluation.

Since the matrix $i^0\kappa$ is hermitian, it may be diagonalized of a unitary matrix

$$i^0\kappa = \mathbf{U}\boldsymbol{\varepsilon}\mathbf{U}^+ \quad (\text{B.18})$$

where $\boldsymbol{\varepsilon}$ is a real diagonal matrix. The ${}^0\tilde{\kappa}$ may therefore be rewritten as

$${}^0\tilde{\kappa} = \sum \epsilon_p \bar{n}_p \quad (\text{B.19})$$

where

$$\bar{n}_p = \bar{a}_p^+ \bar{a}_p \quad (\text{B.20})$$

$$\bar{a}_p^+ = \sum_j \tilde{a}_j^+ (\mathbf{U})_{jp} \quad (\text{B.21})$$

$$\bar{a}_p = \sum_j \tilde{a}_j (\mathbf{U}^+)_{pj} \quad (\text{B.22})$$

Transforming all operators in Eq. (B.13) into the overbar ($\bar{\cdot}$) basis gives

$$F(^0\lambda)_{rs} = 2 \sum_n \frac{(-i)^n}{(n+1)!} \langle \tilde{0} | \left[\sum_k \left((\epsilon_k \bar{n}_k)^n \bar{a}_p^+ \bar{a}_q \right), H \right] |\tilde{0}\rangle U_{pr}^+ U_{sq} \quad (\text{B.23})$$

Since

$$\left(\sum_k \epsilon_k \bar{n}_k \right) (\bar{a}_p^+ \bar{a}_q) = (\epsilon_p - \epsilon_q) \bar{a}_p^+ \bar{a}_q \quad (\text{B.24})$$

we may rewrite Eq. (B.23) as

$$\begin{aligned} F(^0\lambda)_{rs} &= 2 \sum_{pq} \frac{(-i)^n}{(n+1)!} (\epsilon_p - \epsilon_q)^n \langle \tilde{0} | [\bar{a}_p^+ \bar{a}_q, H] |\tilde{0}\rangle U_{pr}^+ U_{sq} \\ &= \sum_{pq} g(p, q) f(p, q) U_{pr}^+ U_{sq} \end{aligned} \quad (\text{B.25})$$

where

$$g(p, q) = \sum_n \frac{(-i)^n}{(n+1)!} (\epsilon_p - \epsilon_q)^n \quad (B.26)$$

$$\bar{f}(p, q) = 2\langle \tilde{0} | [\bar{a}_p^+ \bar{a}_q, H] | \tilde{0} \rangle \quad (B.27)$$

If $\epsilon_p \neq \epsilon_q$, then $g(p, q)$ may be written as

$$\frac{i}{\epsilon_p - \epsilon_q} \sum_{n=0}^{\infty} \frac{(-i)^{n+1}}{(n+1)!} (\epsilon_p - \epsilon_q)^{n+1} = \frac{i}{\epsilon_p - \epsilon_q} [\exp(-i(\epsilon_p - \epsilon_q)) - 1] \quad (B.28)$$

Equation (B.28) is numerically unstable for small values of $\epsilon_p - \epsilon_q$, and it then becomes advantageous to use Eq. (B.25) directly to evaluate $f(p, q)$. If, for example, Eq. (B.25) is applied when $\epsilon_p - \epsilon_q < 10^{-5}$, only the first few terms in the series in Eq. (B.25) are sufficient to give $g(p, q)$ to machine accuracy. The gradient elements in an iterative procedure based on both carrying out a linear and nonlinear transformation of the variables require evaluation of the elements $\langle \tilde{0} | [\tilde{a}_r^+ \tilde{a}_s, H] | \tilde{0} \rangle$. The additional work required for constructing the gradient elements corresponding to a linear transformation is the diagonalization of ${}^0\kappa$ and the matrix multiplications in Eq. (B.25).

B.2. Derivatives with Respect to State Parameters

The partial derivative of the total energy in Eq. (B.1) with respect to state parameters becomes using Eqs. (B.8), (B.9), and (B.10):

$$\begin{aligned} F({}^0\lambda)_n &= \left. \frac{\partial E({}^0\kappa + {}^0\kappa, {}^0\mathbf{S} + \mathbf{S})}{\partial S_{n0}} \right|_{\substack{\mathbf{S} = \mathbf{0} \\ \kappa = \mathbf{0}}} \\ &= \langle \tilde{0} | \left[\sum_m \frac{(-i)^m}{(m+1)!} {}^0\hat{S}^m R_{n0} \right] \exp(-{}^0i\hat{S}) \\ &\quad \times \exp(-{}^0i\hat{\kappa}) H \exp({}^0i\hat{\kappa}) \exp({}^0i\hat{S}) | \tilde{0} \rangle \\ &\quad - \langle \tilde{0} | \exp(-{}^0i\hat{S}) \exp(-{}^0i\hat{\kappa}) H \exp({}^0i\hat{\kappa}) \\ &\quad \times \exp({}^0i\hat{S}) \left[\sum_m \frac{(-i)^m}{(m+1)!} {}^0\hat{S}^m R_{n0} \right] | \tilde{0} \rangle \\ &= \langle \tilde{0} | \left[\sum_m \frac{(-i)^m}{(m+1)!} {}^0\tilde{\hat{S}}^m \tilde{R}_{n0}, H \right] | \tilde{0} \rangle \end{aligned} \quad (B.29)$$

where $|\tilde{0}\rangle$ is defined in Eq. (B.15) and the superoperator ${}^0\tilde{\hat{S}}$ is defined through

the relations

$${}^0\tilde{\hat{S}}f = \left[{}^0\tilde{S}, f \right], \quad {}^0\tilde{S} = i \sum_m {}^0S_{m0} \tilde{R}_{m0} \quad (\text{B.30})$$

$$\tilde{R}_{m0} = |\tilde{m}\rangle\langle\tilde{0}| - |\tilde{0}\rangle\langle\tilde{m}| \quad (\text{B.31})$$

The most straightforward way of evaluating Eq. (B.29) would be to reuse the method of the previous section and thus to diagonalize 0S . If the number of the CI states is up in the range of 100–1000, this part of the iterative procedure may be computationally rather demanding. Because of the simple rank-2 structure of 0S , an easier method may be developed.

Let us consider the series

$$\sum_m \frac{(-i)^m}{(m+1)!} {}^0\tilde{S}^m \tilde{R}_{n0} = \tilde{R}_{n0} - \frac{i}{2} \left[{}^0\hat{S}, \tilde{R}_{n0} \right] + \frac{1}{3!} \left[{}^0\tilde{S}, \left[{}^0\tilde{S}, \tilde{R}_{n0} \right] \right] + \dots \quad (\text{B.32})$$

By denoting $(|\tilde{n}\rangle\langle\tilde{m}| - |\tilde{m}\rangle\langle\tilde{n}|)$ with \tilde{R}_{nm} we may express the second and third terms of Eq. (B.32) as

$${}^0\tilde{S}\tilde{R}_{n0} = \left[{}^0\tilde{S}, \tilde{R}_{n0} \right] = \sum_i i {}^0S_{m0} \tilde{R}_{nm} \quad (\text{B.33})$$

$${}^0\tilde{S}^2\tilde{R}_{n0} = \left[{}^0\tilde{S}, \left[{}^0\tilde{S}, \tilde{R}_{n0} \right] \right] = i {}^0S_{n0} {}^0\tilde{S} + P \tilde{R}_{n0}; \quad P = \sum_k S_{k0}^2 \quad (\text{B.34})$$

Using Eq. (B.31), we obtain

$${}^0\tilde{S}^3\tilde{R}_{n0} = \left[{}^0\tilde{S}, \left[{}^0\tilde{S}, \left[{}^0\tilde{S}, \tilde{R}_{n0} \right] \right] \right] = P \left[{}^0\tilde{S}, \tilde{R}_{n0} \right] = P {}^0\tilde{S}\tilde{R}_{n0} \quad (\text{B.35})$$

and Eqs. (B.33) and (B.34) may be generalized

$${}^0\tilde{S}^{2n}\tilde{R}_{n0} = (P)^{n-1} \left(i {}^0S_{n0} {}^0\tilde{S} + P \tilde{R}_{n0} \right) \quad (\text{B.36})$$

$${}^0\tilde{S}^{2n+1}\tilde{R}_{n0} = i(P)^n \sum_m {}^0S_{m0} \tilde{R}_{nm} \quad (\text{B.37})$$

We thus have

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{(-i)^m}{(m+1)!} {}^0\tilde{S}^m \tilde{R}_{n0} &= \left(1 - \frac{P}{3!} + \frac{P^2}{5!} - \frac{P^3}{7!} + \dots \right) \tilde{R}_{n0} \\ &\quad + \left(\frac{1}{3!} - \frac{P}{5!} + \frac{P^2}{7!} + \dots \right) {}^0S_{n0} \left(\sum_m {}^0S_{m0} \tilde{R}_{m0} \right) \\ &\quad + \left(\frac{1}{2} - \frac{P}{4!} + \frac{P^2}{6!} + \dots \right) \sum_m {}^0S_{m0} \tilde{R}_{nm} \quad (\text{B.38}) \end{aligned}$$

Denoting the series

$$\psi = 1 - \frac{P}{3!} + \frac{P^2}{5!} - \frac{P^3}{7!} + \dots \quad (\text{B.39})$$

$$\phi = \frac{1}{3!} - \frac{P}{5!} + \frac{P^2}{7!} - \dots \quad (\text{B.40})$$

we may write the gradient elements $F(^0\lambda)_n$ as

$$F(^0\lambda)_n = \psi \langle \tilde{0} | [\tilde{R}_{n0}, H] | \tilde{0} \rangle + \phi ^0 S_{n0} \sum_m {}^0 S_{m0} \langle \tilde{0} | [\tilde{R}_{m0}, H] | \tilde{0} \rangle \quad (\text{B.41})$$

The calculation of ψ and ϕ is very simple. The additional work required for constructing the gradient elements in the configuration space corresponding to a linear transformation among the variables is thus the summations in the last term of Eq. (B.41), which are extremely simple to carry out even for very large configuration spaces.

APPENDIX C: FORMULAS FOR $(P'^T F(0))$ AND $(P'^T G(0) P')$

To simplify the computation of gradient-like and Hessian-like elements, we divide orbitals in three classes: class i , consisting of all inactive (completely occupied) orbitals; class a , consisting of all active (partly occupied) orbitals; and class s , consisting of all secondary (completely unoccupied) orbitals. Orbitals from class i are denoted i, j, k, \dots ; orbitals from class a are denoted a, b, c, \dots ; orbitals from class s are denoted s, t, u, \dots . General indices are denoted o, p, q, r .

Four types of orbital excitations exist:

$$\begin{aligned} & a_a^+ a_i - a_i^+ a_a \\ & a_s^+ a_i - a_i^+ a_s \\ & a_a^+ a_b - a_b^+ a_a \\ & a_s^+ a_a - a_a^+ a_s \end{aligned} \quad (\text{C.1})$$

By taking the specific nature of an orbital excitation into account, several simplifications can be obtained. The number of commutators in a given expression can be increased for operators of the type $a_s^+ a_i - a_i^+ a_s$:

$$\langle 0 | [a_s^+ a_i, H] | 0 \rangle = \langle 0 | [[a_s^+, H], a_i]_+ | 0 \rangle \quad (\text{C.2})$$

By introducing the anticommutator an extra delta function is introduced, so the summations are reduced. The occurrence of an inactive orbital in the one-

and two-electron density matrices allows simplifications:

$$\begin{aligned}\langle 0 | a_{i\sigma}^+ a_{r\sigma}^- | 0 \rangle &= 2\delta_{i,r} \\ \langle 0 | a_{i\sigma}^+ a_{o\sigma'}^+ a_{p\sigma'}^- a_{r\sigma}^- | 0 \rangle &= 2\delta_{i,r} \langle 0 | a_{o\sigma'}^+ a_{p\sigma'}^- | 0 \rangle - \delta_{i,p} \langle 0 | a_{o\sigma'}^+ a_{r\sigma'}^- | 0 \rangle\end{aligned}\quad (\text{C.3})$$

In Eq. (C.3) we have explicitly written spin indices. As pointed out by Siegbahn et al.,³⁵ density matrices are thus only needed for indices corresponding to active orbitals.

The construction of relevant formulas are also eased by the use of the following relations:

$$\begin{aligned}\langle 0 | [T_i^+ - T_i, H] | 0 \rangle &= 2\langle 0 | [T_i^+, H] | 0 \rangle \\ \langle 0 | [T_i^+ - T_i, T_j^+ - T_j, H] | 0 \rangle &= 2\langle 0 | [T_i^+, T_j^+ - T_j, H] | 0 \rangle \\ \langle 0 | [T_i^+, T_j^+ - T_j, H] | 0 \rangle &= \langle 0 | [T_i^+, [T_j^+ - T_j, H]] | 0 \rangle \\ &\quad + \frac{1}{2} \langle 0 | [[T_i^+, T_j^+ - T_j], H] | 0 \rangle\end{aligned}\quad (\text{C.4})$$

The computations are further facilitated with the use of *symmetric* two-electron density matrices

$$\rho_{ad, bc}^{(2)} = \frac{1}{2} \langle 0 | a_{a\sigma}^+ a_{b\sigma'}^+ a_{c\sigma'}^- a_{d\sigma}^- | 0 \rangle + \frac{1}{2} \langle 0 | a_{a\sigma}^+ a_{c\sigma'}^+ a_{b\sigma'}^- a_{d\sigma}^- | 0 \rangle \quad (\text{C.5})$$

and one-electron density matrices

$$\rho_{ab}^{(1)} = \langle 0 | a_{a\sigma}^+ a_{b\sigma}^- | 0 \rangle \quad (\text{C.6})$$

A combination of the aforementioned relations together with Eqs. (83) and (84) allows reasonable, easy derivations of the following. Siegbahn et al.³⁵ give formulas for the gradient $\mathbf{F}(\mathbf{0})$ and $\mathbf{G}(\mathbf{0})$. We list here the formulas for $\mathbf{P}'^T \mathbf{F}(\mathbf{0})$ and $\mathbf{P}'^T \mathbf{G}(\mathbf{0}) \mathbf{P}'$, where \mathbf{P}' is a direct-product matrix:

$$\mathbf{P}' = \begin{pmatrix} \mathbf{P}^{OO} & \mathbf{P}^{OC} \\ \mathbf{P}^{CO} & \mathbf{P}^{CC} \end{pmatrix} \begin{pmatrix} \mathbf{P}^{OO} & \mathbf{P}^{OC} \\ \mathbf{P}^{CO} & \mathbf{P}^{CC} \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \quad (\text{C.7})$$

In Eq. (C.7) the indices have been partitioned in an orbital excitation part (O), and a state excitation part (C). The matrix \mathbf{P} has the form

$$\begin{pmatrix} \mathbf{P}^{II} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}^{SS} \end{pmatrix} \quad (\text{C.8})$$

where the orbital space has been partitioned in its usual three parts. Construction of $\mathbf{P}'^T \mathbf{F}(\mathbf{0})$ and $\mathbf{P}'^T \mathbf{G}(\mathbf{0}) \mathbf{P}'$ is required when the reduced transformations discussed in Section IX.A are used. To compact expressions we introduce the following:

$$\begin{aligned} (\hat{o}\tilde{p}|\tilde{q}\tilde{r}) &= P_{o'o} P_{p'p} P_{q'q} P_{r'r} (o'p'|q'r') \\ \tilde{h}_{pq} &= P_{p'p} P_{q'q} h_{p'q'} \end{aligned} \quad (\text{C.9})$$

and modified Fock matrices

$$\begin{aligned} \tilde{D}_{ap} &= \rho_{ab}^{(1)} \tilde{D}_{pb}^I + \rho_{ab,cd}^{(2)} (\tilde{p}\tilde{b}|\tilde{c}\tilde{d}) \\ \tilde{D}_{ip} &= 2(\tilde{D}_{ip}^I + \tilde{D}_{ip}^A) \\ \tilde{D}_{pq}^I &= \tilde{h}_{pq} + [2(\tilde{p}\tilde{q}|ii) - (\tilde{p}i|\tilde{q}i)] \\ \tilde{D}_{pq}^A &= \rho_{ab}^{(1)} [2(\tilde{p}\tilde{q}|\tilde{a}\tilde{b}) - \frac{1}{2}(\tilde{p}\tilde{a}|\tilde{q}\tilde{b})] \end{aligned} \quad (\text{C.10})$$

The two-electron indices in \tilde{D}^I in Eq. (C.10) involves two different sets of orbitals. This does not introduce complications since \tilde{D}^I usually is constructed directly from integrals over basis orbitals. The formulas below are equivalent to the formulas presented by Siegbahn et al. for $\mathbf{P} = \mathbf{1}$.

C.1. Formulas for $\mathbf{P}'^T \mathbf{F}(\mathbf{0})$

Five types of elements $\mathbf{P}'^T \mathbf{F}(0)$ exist:

$$\begin{aligned} (\mathbf{P}'^T \mathbf{F}(0))_{ai} &= P'_{(bj)(ai)} \langle 0 | [a_{b\sigma}^+ a_{j\sigma} - a_{j\sigma}^+ a_{b\sigma}, H] | 0 \rangle \\ &= P_{ji} \langle 0 | [a_{a\sigma}^+ a_{j\sigma} - a_{j\sigma}^+ a_{a\sigma}, H] | 0 \rangle \\ &= 2(\tilde{D}_{ai} - \tilde{D}_{ia}) \\ &= 2\rho_{ac}^{(1)} D_{ac}^I + 2\rho_{ac,de}^{(2)} (\tilde{c}d|de) - 4(\tilde{D}_{ia}^I + \tilde{D}_{ia}^A) \end{aligned} \quad (\text{C.11a})$$

$$\begin{aligned} (\mathbf{P}'^T \mathbf{F}(0))_{si} &= 2(\tilde{D}_{is} - \tilde{D}_{si}) \\ &= -4(\tilde{D}_{is}^I + \tilde{D}_{is}^A) \end{aligned} \quad (\text{C.11b})$$

$$\begin{aligned} (\mathbf{P}'^T \mathbf{F}(0))_{ab} &= -2\rho_{bc}^{(1)} \tilde{D}_{ac}^I + 2\rho_{ac}^{(1)} \tilde{D}_{bc}^I \\ &\quad - 2\rho_{bc,de}^{(2)} (\tilde{a}\tilde{c}|\tilde{d}\tilde{e}) \\ &\quad + 2\rho_{ac,de}^{(2)} (\tilde{b}\tilde{c}|\tilde{d}\tilde{e}) \end{aligned} \quad (\text{C.11c})$$

$$(\mathbf{P}'^T \mathbf{F}(0))_{sa} = -2\rho_{ab}^{(1)} \tilde{D}_{sb}^I - 2\rho_{ab,cd}^{(2)} (\tilde{s}\tilde{b}|\tilde{c}\tilde{d}) \quad (\text{C.11d})$$

The “derivative” with respect to a state variable is

$$(\mathbf{P}'^T \mathbf{F}(0))_n = \langle 0 | [| n \rangle \langle 0 | - | 0 \rangle \langle n |, H] | 0 \rangle = -2 \langle 0 | H | n \rangle \quad (\text{C.11e})$$

The transformed gradient ($\mathbf{P}' \mathbf{F}(0)$) can thus be calculated from \tilde{D}^I , \tilde{D}^A and two-electron integrals ($\tilde{a}\tilde{b}|\tilde{c}\tilde{p}$). This is the basis for the reduced T_1 transformation discussed in Section IX.A.

C.2. Formulas for $(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')$

There are 10 formulas for the orbital–orbital Hessian, four formulas for the orbital–state part of the Hessian, and one formula for the state–state part of the Hessian:

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{ai,bj} &= 2[\rho_{ba,cd}^{(2)}(\tilde{c}\tilde{d}|\tilde{i}\tilde{j}) + 2\rho_{bd,ca}^{(2)}(\tilde{c}\tilde{i}|\tilde{d}\tilde{j})] \\ &\quad + 2(\delta_{ac} - \rho_{ac}^{(1)}) (4(\tilde{c}\tilde{i}|\tilde{b}\tilde{j}) - (\tilde{b}\tilde{i}|\tilde{c}\tilde{j}) - (\tilde{b}\tilde{c}|\tilde{i}\tilde{j})) \\ &\quad + 2(\delta_{bc} - \rho_{bc}^{(1)}) (4(\tilde{c}\tilde{i}|\tilde{a}\tilde{j}) - (\tilde{a}\tilde{j}|\tilde{c}\tilde{i}) - (\tilde{a}\tilde{c}|\tilde{i}\tilde{j})) \\ &\quad + 2\rho_{ab}^{(1)}\tilde{D}_{ab}^I + 2(\mathbf{P}^T \mathbf{P})_{ij} (2\tilde{D}_{ab}^I + 2\tilde{D}_{ab}^A - \tilde{D}_{ba}) \\ &\quad - 4\delta_{ab} (\tilde{D}_{ij}^I + \tilde{D}_{ij}^A) \end{aligned} \quad (\text{C.12a})$$

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{ai,sj} &= 2(2\delta_{ab} - \rho_{ab}^{(1)}) (4(\tilde{s}\tilde{j}|\tilde{b}\tilde{i}) - (\tilde{s}\tilde{b}|\tilde{i}\tilde{j}) - (\tilde{s}\tilde{i}|\tilde{b}\tilde{j})) \\ &\quad + 4(\mathbf{P}^T \mathbf{P})_{ij} (\tilde{D}_{sa}^I + \tilde{D}_{sa}^A) - (\mathbf{P}^T \mathbf{P})_{ij} \tilde{D}_{as} \end{aligned} \quad (\text{C.12b})$$

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{ai,bc} &= 4[(\tilde{d}\tilde{i}|\tilde{c}\tilde{e})\rho_{ad,be}^{(2)} - (\tilde{d}\tilde{i}|\tilde{b}\tilde{e})\rho_{ad,ce}^{(2)}] \\ &\quad + 2[(\tilde{c}\tilde{i}|\tilde{d}\tilde{e})\rho_{ab,de}^{(2)} - (\tilde{b}\tilde{i}|\tilde{d}\tilde{e})\rho_{ac,de}^{(2)}] \\ &\quad - 2\rho_{bd}^{(1)} [4(\tilde{a}\tilde{i}|\tilde{c}\tilde{d}) - (\tilde{a}\tilde{d}|\tilde{c}\tilde{i}) - (\tilde{a}\tilde{c}|\tilde{d}\tilde{i})] \\ &\quad + 2\rho_{cd}^{(1)} [4(\tilde{a}\tilde{i}|\tilde{b}\tilde{d}) - (\tilde{a}\tilde{d}|\tilde{b}\tilde{i}) - (\tilde{a}\tilde{b}|\tilde{d}\tilde{i})] \\ &\quad + 2\rho_{ab}^{(1)}\tilde{D}_{ic}^I - 2\delta_{ab}\tilde{D}_{ci} - 2\rho_{ac}^{(1)}\tilde{D}_{ib} + 2\delta_{ac}\tilde{D}_{bi} \end{aligned} \quad (\text{C.12c})$$

$$\begin{aligned} (\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{ai,sb} &= -2[\rho_{ab,cd}^{(2)}(\tilde{s}\tilde{i}|\tilde{c}\tilde{d}) + 2\rho_{ac,bd}^{(2)}(\tilde{s}\tilde{d}|\tilde{c}\tilde{i})] \\ &\quad + 2\rho_{bc}^{(1)} (4(\tilde{s}\tilde{c}|\tilde{a}\tilde{i}) - (\tilde{s}\tilde{i}|\tilde{a}\tilde{c}) - (\tilde{s}\tilde{a}|\tilde{c}\tilde{i})) \\ &\quad - 2\rho_{ab}^{(1)}\tilde{D}_{si}^I + 2\delta_{ab} (\tilde{D}_{si}^I + \tilde{D}_{si}^A) \end{aligned} \quad (\text{C.12d})$$

$$\begin{aligned}
(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{ab, cd} = & 4 \left[(\tilde{a}\tilde{e}|\tilde{c}\tilde{f}) \rho_{be, df}^{(2)} + (\tilde{b}\tilde{e}|\tilde{d}\tilde{f}) \rho_{ae, cf}^{(2)} \right. \\
& - (\tilde{b}\tilde{e}|\tilde{c}\tilde{f}) \rho_{ae, df}^{(2)} - (\tilde{a}\tilde{e}|\tilde{d}\tilde{f}) \rho_{be, df}^{(2)} \left. \right] \\
& + 2 \left[(\tilde{a}\tilde{c}|\tilde{e}\tilde{f}) \rho_{bd, ef}^{(2)} + (\tilde{b}\tilde{d}|\tilde{e}\tilde{f}) \rho_{ac, ef}^{(2)} \right. \\
& - (\tilde{a}\tilde{d}|\tilde{e}\tilde{f}) \rho_{bc, ef}^{(2)} - (\tilde{b}\tilde{c}|\tilde{e}\tilde{f}) \rho_{ad, ef}^{(2)} \left. \right] \\
& + 2 \left[\rho_{bd}^{(1)} \tilde{D}_{ac}^I + \rho_{ac}^{(1)} \tilde{D}_{bd}^I - \rho_{ad}^{(1)} \tilde{D}_{bc}^I - \rho_{bc}^{(1)} \tilde{D}_{ad}^I \right] \\
& + 2 \left[\delta_{ad} \tilde{D}_{cb} + \delta_{bc} \tilde{D}_{da} - \delta_{bd} \tilde{D}_{ca} - \delta_{ac} \tilde{D}_{db} \right]
\end{aligned} \tag{C.12e}$$

$$\begin{aligned}
(P'^T G(0) P')_{ab, sc} = & 4 \left[(\tilde{a}\tilde{d}|\tilde{e}\tilde{s}) \rho_{bd, ce}^{(2)} - (\tilde{b}\tilde{d}|\tilde{e}\tilde{s}) \rho_{ad, ce}^{(2)} \right] \\
& + 2 \left[(\tilde{a}\tilde{s}|\tilde{d}\tilde{e}) \rho_{bc, de}^{(2)} - (\tilde{b}\tilde{s}|\tilde{d}\tilde{e}) \rho_{ac, de}^{(2)} \right] \\
& + 2 \rho_{bc}^{(1)} \tilde{D}_{sa}^I - 2 \rho_{ac}^{(1)} \tilde{D}_{sb}^I
\end{aligned} \tag{C.12f}$$

$$\begin{aligned}
(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{si, t j} = & 4 \left((\tilde{s}\tilde{i}|\tilde{i}\tilde{j}) - (\tilde{s}\tilde{i}|\tilde{j}\tilde{i}) - (\tilde{s}\tilde{j}|\tilde{i}\tilde{i}) \right) \\
& + 4(\mathbf{P}^T \mathbf{P})_{ij} \left(\tilde{D}_{st}^I + \tilde{D}_{st}^A \right) \\
& - 4(\mathbf{P}^T \mathbf{P})_{st} \left(\tilde{D}_{ij}^I + \tilde{D}_{ij}^A \right)
\end{aligned} \tag{C.12g}$$

$$\begin{aligned}
(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{si, ab} = & 2 \rho_{bc}^{(1)} \left[4(\tilde{a}\tilde{c}|\tilde{i}\tilde{s}) - (\tilde{a}\tilde{i}|\tilde{c}\tilde{s}) - (\tilde{a}\tilde{s}|\tilde{c}\tilde{i}) \right] \\
& - 2 \rho_{ac}^{(1)} \left[4(\tilde{b}\tilde{c}|\tilde{i}\tilde{s}) - (\tilde{b}\tilde{i}|\tilde{c}\tilde{s}) - (\tilde{b}\tilde{s}|\tilde{c}\tilde{i}) \right]
\end{aligned} \tag{C.12h}$$

$$\begin{aligned}
(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{si, ta} = & 2 \rho_{ab}^{(1)} \left(4(\tilde{s}\tilde{i}|\tilde{i}\tilde{b}) - (\tilde{s}\tilde{b}|\tilde{i}\tilde{i}) - (\tilde{s}\tilde{i}|\tilde{b}\tilde{i}) \right) \\
& - 2(\mathbf{P}^T \mathbf{P})_{st} \left(2 \tilde{D}_{ai}^I + 2 \tilde{D}_{ai}^A - \tilde{D}_{ai} \right)
\end{aligned} \tag{C.12i}$$

$$\begin{aligned}
(\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}')_{sa, tb} = & 2 \left(\rho_{ab, cd}^{(2)} (\tilde{s}\tilde{t}|\tilde{c}\tilde{d}) + 2 \rho_{ad, bc}^{(2)} (\tilde{s}\tilde{d}|\tilde{t}\tilde{c}) \right) \\
& + 2 \rho_{ab}^{(1)} \tilde{D}_{st}^I - 2(\mathbf{P}^T \mathbf{P})_{st} \tilde{D}_{ba}
\end{aligned} \tag{C.12j}$$

Coupling-elements have the form

$$(\mathbf{P}'^T \mathbf{G}(0))_{pq, n} = -2 P_{p'p} P_{q'q} \langle 0 | [a_{p'\sigma}^+ a_{q'\sigma}^- - a_{q'\sigma}^+ a_{p'\sigma}^-, H] | n \rangle \tag{C.12k}$$

These elements have the same structure as $\mathbf{P}'^T \mathbf{F}(0)$ [Eq. (C.11)], but regular density matrices have been replaced by transition density matrices. The four types of coupling elements can thus be obtained from Eq. (C.11e) by replacing *symmetric* density matrices with symmetric transition density matrices.

The state-state part of $\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}'$ becomes

$$\mathbf{G}(0)_{nm} = +2\langle n|H|m\rangle - 2\delta_{nm}\langle 0|H|0\rangle \quad (\text{C.12l})$$

These elements can be calculated from integrals over active orbitals and \mathbf{D}^I .

It has thus been demonstrated that $\mathbf{P}'^T \mathbf{G}(0) \mathbf{P}'$, where \mathbf{P}' has the structure indicated in Eqs. (C.7) and (C.8), can be calculated from integrals in the tilde basis and a one-electron operator. This is the basis for the reduced integral transformations T_2 and T'_2 discussed in Section IX.A.

APPENDIX D: EIGENVALUES OF THE FULL HESSIAN AND THE REDUCED HESSIAN MATRIX

One desirable feature of the n th excited state of a given symmetry is that the Hessian matrix for that state has n negative eigenvalues.^{13, 15, 16} In the one-step second-order approach, the Hessian matrix

$$\begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OC} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix}$$

is part of the iterative function [see Eqs. (60), (75), and (76)] and the number of negative eigenvalues of the Hessian matrix may therefore be used to direct the calculation to the desired stationary point. In the two-step second order procedure, it is the reduced Hessian matrix $\mathbf{G}^{OO} - \mathbf{G}^{OC}(\mathbf{G}^{CC})^{-1}\mathbf{G}^{CO}$ and the configuration block \mathbf{G}^{CC} that appear explicitly in the iterative function [see Eq. (80)]. Therefore in a two-step calculation it becomes important, in order to direct the calculation to the desired stationary point, to know how the negative eigenvalues of the n th excited state Hessian matrix get distributed into the reduced Hessian matrix and the configuration interaction matrix. We show in this appendix that if the Hessian matrix has n negative eigenvalues, and \mathbf{G}^{CC} contains m negative eigenvalues, then the reduced Hessian matrix must contain $n - m$ negative eigenvalues. We use multidimensional partitioning technique to show this point.⁴¹ For the case where there are n total negative eigenvalues, all in the \mathbf{G}^{CC} block of the Hessian matrix, Shepard et al.³² have previously proved, using multidimensional partitioning theory, that the reduced Hessian matrix is positive definite.

The eigenvalue problem for the Hessian matrix

$$\begin{pmatrix} \mathbf{G}^{OO} & \mathbf{G}^{OO} \\ \mathbf{G}^{CO} & \mathbf{G}^{CC} \end{pmatrix} \begin{pmatrix} \mathbf{U}^O \\ \mathbf{U}^O \end{pmatrix} = \begin{pmatrix} \mathbf{U}^O \\ \mathbf{U}^C \end{pmatrix} E \quad (\text{D.1})$$

transforms, using partitioning theory, to an equation of the reduced

dimension

$$(\mathbf{G}^{OO} + \mathbf{G}^{OC}(\mathbf{E}\mathbf{1} - \mathbf{G}^{CC})^{-1}\mathbf{G}^{CO})\mathbf{U}^O = \mathbf{U}^O E \quad (\text{D.2})$$

If we introduce the matrix function

$$\mathbf{L}(E) = \mathbf{G}^{OO} + \mathbf{G}^{OC}(\mathbf{E}\mathbf{1} - \mathbf{G}^{CC})^{-1}\mathbf{G}^{CO} \quad (\text{D.3})$$

and define the multivalued function $\underline{\epsilon}(E)$ to contain the eigenvalues of $\mathbf{L}(E)$, then from the partitioned form of the Hessian eigenvalue equation [Eq. (D.3)] it is clear that the eigenvalues of the Hessian matrix occur at the values of $\mathbf{L}(E)$ for which

$$E = \underline{\epsilon}(E) \quad (\text{D.4})$$

In Fig. 1 (see p. 61) we have plotted the multivalued function $\underline{\epsilon}(E)$ as a function of E for the case where the configuration space \mathbf{G}^{CC} has a dimension of 4 and the orbital excitation space a dimension of 6. Generalization to arbitrary dimensions is straightforward. The branches of $\underline{\epsilon}(E)$ have horizontal asymptotes at the eigenvalues of \mathbf{G}^{OO} denoted O_1, O_2, \dots, O_6 and vertical asymptotes at the eigenvalues of \mathbf{G}^{CC} denoted C_1, C_2, C_3 , and C_4 . Each branch of $\underline{\epsilon}(E)$ is a nonincreasing function of E , and except when the matrices involved have a very special structure the individual branches of $\underline{\epsilon}(E)$ satisfy a noncrossing rule.⁴¹ The eigenvalues of the reduced Hessian matrix are values of $\underline{\epsilon}(E)$ for $E = 0$. A comprehensive discussion of the structural characteristics of a multidimensional function $\underline{\epsilon}(E)$ has been given by Löwdin.⁴¹ In Fig. 1 the Hessian matrix has two negative eigenvalues [points of intersection with $E = \underline{\epsilon}(E)$] and \mathbf{G}^{CC} has one negative eigenvalue. We thus assume that we are converging to the second excited state of a given symmetry and that root flipping occurs. From the figure it is clear that the one negative eigenvalue of \mathbf{G}^{CC} introduces one negative eigenvalue in the Hessian matrix. As a result of the root flipping the second branch of $\underline{\epsilon}(E)$, which results in the second negative eigenvalue of the Hessian matrix, has a positive asymptote (C_2). Because of its crossing with the line $E = \underline{\epsilon}(E)$ at the negative E value, this branch must cross the line at $E = 0$ for a negative function value of $\underline{\epsilon}(0)$ of more negative value (larger magnitude) than the crossing at $E = \underline{\epsilon}(E)$. The reduced Hessian therefore has one negative eigenvalue.

A generalization of this result to the case where the Hessian has n negative eigenvalues and \mathbf{G}^{CC} has m negative eigenvalues is simple and straightforwardly yields the conclusion that the reduced Hessian has $n - m$ negative eigenvalues. We should mention that of course \mathbf{G}^{OO} should have no more than n negative eigenvalues.

**APPENDIX E: CONSTRUCTION OF CUBIC CONTRIBUTIONS
FOR THE ITERATIVE CUBIC AND PERTURBATIVE
(CHEBYSHEV) CUBIC PROCEDURES**

E.1. The Cubic Formulas

In this appendix we describe in detail how to calculate the cubic contributions in Eqs. (137) and (208).¹⁶

The direct calculation of the cubic supermatrix K_{ijk} in Eq. (15) when ${}^0\lambda = \mathbf{0}$ is difficult for several reasons. First, there are approximately $N^3 \cdot n^3$ terms, where N is the total number of orbitals and n is the number of occupied orbitals. Second is the fact that all of the density matrices $\langle n | a_i^+ a_j^+ a_k a_l | m \rangle$, where $|n\rangle$ and $|m\rangle$ are the MCSCF CI states, are required. Third is the fact that the calculation of $K_{ijk}\lambda_k$ requires about $N^3 n^3$ multiplications. We develop here an alternative method based on the fact that cubic terms always occur as $K_{ijk}v_k$, where v is a vector and, hence, we calculate $K_{ijk}v_k$ directly. In Eq. (208) v is the vector λ , and in Eq. (137) v is ${}^{NR}\lambda$.

Writing out $K_{ijk}\lambda_k$ explicitly, we get

$$K_{ijk}\lambda_k = \langle 0 | [T_i^+ - T_i, T_j^+ - T_j, T_k^+ - T_k, H] | 0 \rangle \lambda_k \quad (\text{E.1})$$

where T^+ and λ are defined in Eqs. (52) and (51), respectively. Introducing the shorthand notation

$$T_i^+ - T_i = \begin{pmatrix} Q_i^+ - Q_i \\ R_i^+ - R_i \end{pmatrix} = X_i = \begin{pmatrix} U_i \\ V_i \end{pmatrix} \quad (\text{E.2})$$

we may write Eq. (E.1) as

$$K_{ijk}\lambda_k = \langle 0 | [X_i, X_j, X_k, H] | 0 \rangle \lambda_k \quad (\text{E.3})$$

We will evaluate Eq. (E.3) by considering the four possible cases: (1) X_i and X_j both refer to orbital excitation operators; (2) X_i refers to an orbital excitation operator and X_j to a state excitation operator; (3) X_i refers to a state excitation operator and X_j to an orbital excitation operator; and (4) both X_i and X_j refer to state excitation operators.

Case (1): i, j both orbital indices. Equation (E.1) may now be written as

$$K_{ijk}\lambda_k = \langle 0 | [U_i, U_j, U_k, H] | 0 \rangle \kappa_k + \langle 0 | [V_k, [U_i, U_j, H]] | 0 \rangle S_{k0} \quad (\text{E.4})$$

The first term in Eq. (E.4) may be written as

$$\begin{aligned} & \frac{1}{6} \left\{ \langle 0 | [U_i, [U_j, [U_k, H]]] | 0 \rangle + \langle 0 | [U_i, [U_k, [U_j, H]]] | 0 \rangle \right. \\ & + \langle 0 | [U_j, [U_i, [U_k, H]]] | 0 \rangle + \langle 0 | [U_j, [U_k, [U_i, H]]] | 0 \rangle \\ & \left. + \langle 0 | [U_k, [U_i, [U_j, H]]] | 0 \rangle + \langle 0 | [U_k, [U_j, [U_i, H]]] | 0 \rangle \right\} \kappa_k \quad (\text{E.5}) \end{aligned}$$

If we commute the operators in Eq. (E.5) such that in terms of the type $\langle 0 | [U, [U, H]] | 0 \rangle$, U_k becomes commuted with H , that is, $[U_k, H]$, we obtain

$$\begin{aligned} \langle 0 | [U_i, U_j, U_k, H] | 0 \rangle \kappa_k = & \frac{1}{2} \langle 0 | [U_i, [U_j, [U_k, H]]] | 0 \rangle \kappa_k \\ & + \frac{1}{2} \langle 0 | [U_j, [U_i, [U_k, H]]] | 0 \rangle \kappa_k \\ & + \frac{1}{3} \langle 0 | [U_i, [[U_k, U_j], H]] | 0 \rangle \kappa_k \\ & + \frac{1}{3} \langle 0 | [U_j, [[U_k, U_i], H]] | 0 \rangle \kappa_k \\ & + \frac{1}{6} \langle 0 | [[U_k, U_i], [U_j, H]] | 0 \rangle \kappa_k \\ & + \frac{1}{6} \langle 0 | [[U_k, U_j], [U_i, H]] | 0 \rangle \kappa_k \quad (\text{E.6}) \end{aligned}$$

Denoting

$$[\hat{\kappa}, H] = \tilde{H},$$

and commuting terms in Eq. (E.6), we obtain

$$\begin{aligned} \langle 0 | [U_i, U_j, U_k, H] | 0 \rangle \kappa_k = & \langle 0 | [U_i, U_j, \tilde{H}] | 0 \rangle + \frac{1}{2} \langle 0 | [[U_k, U_j], [U_i, H]] | 0 \rangle \kappa_k \\ & + \frac{1}{2} \langle 0 | [[U_k, U_i], [U_j, H]] | 0 \rangle \kappa_k \\ & + \frac{1}{2} \langle 0 | [[U_k, U_j], [U_i, H]] | 0 \rangle \kappa_k \\ & + \frac{1}{6} \langle 0 | [[U_i, [U_k, U_j]], H] | 0 \rangle \kappa_k \\ & + \frac{1}{6} \langle 0 | [[U_j, [U_k, U_i]], H] | 0 \rangle \kappa_k \quad (\text{E.7}) \end{aligned}$$

It may readily be seen that the term $\langle 0 | [U_i, U_j, \tilde{H}] | 0 \rangle$ corresponds to a Hessian element with a modified Hamiltonian \tilde{H} , which we now evaluate: H has the form

$$H = \sum_{rs} h_{rs} a_r^+ a_s + \frac{1}{2} \sum_{rstu} \langle rs | ut \rangle a_r^+ a_s^+ a_t a_u$$

and

$$\tilde{H} = [\hat{\kappa}, H] = \sum_{rs} \tilde{h}_{rs} a_r^+ a_s + \frac{1}{2} \sum_{rstu} \langle rs | \tilde{ut} \rangle a_r^+ a_s^+ a_t a_u \quad (\text{E.8})$$

where

$$\begin{aligned} \tilde{h}_{rs} &= \sum_p (h_{rp} \kappa_{sp} + h_{ps} \kappa_{rp}) \\ \langle rs | \tilde{tu} \rangle &= \sum_p (\kappa_{rp} \langle ps | tu \rangle + \kappa_{sp} \langle rp | tu \rangle + \kappa_{tp} \langle rs | pu \rangle + \kappa_{up} \langle rs | tp \rangle) \end{aligned} \quad (\text{E.9})$$

The modified Hamiltonian \tilde{H} is thus identical to H with transformed integrals. However, the integral transformation is a *one-index* transformation and may therefore easily be carried out. The residual terms of Eq. (E.7) are just straightforward Hessian and GBT matrix elements (i.e., the commutator of two one-body operators is a sum of one-body operators). We also point out that

$$\langle 0 | [[U_k, U_j], [U_i, H]] | 0 \rangle \kappa_k \quad (\text{E.10})$$

contains some terms of the form

$$\langle 0 | [a_{cl}^+ a_{cl'}, [U_i, H]] | 0 \rangle \quad \text{and} \quad \langle 0 | [a_{un}^+ a_{un'}, [U_i, H]] | 0 \rangle \quad (\text{E.11})$$

where cl and cl' refer to orbitals that are completely occupied and un and un' refer to orbitals which are completely unoccupied. These terms, however, become zero as $a_{cl}^+ a_{cl'}$ or $a_{un}^+ a_{un'}$ operates directly on the reference state. This is why the commutators $[U_k, U_j]$ are chosen to be the outer part of the commutator operator string.

To simplify the second term in Eq. (E.4), it is convenient to introduce the notation

$$|\delta 0\rangle = - \sum_n S_{n0} R_n^+ |0\rangle = - \sum_n S_{n0} |n\rangle \quad (\text{E.12})$$

With this notation we may rewrite the second term of Eq. (E.4) as

$$\langle 0 | [V_n, [U_i, U_j, H]] | 0 \rangle S_{n0} = \langle 0 | [U_i, U_j, H] | \delta 0 \rangle + \langle \delta 0 | [U_i, U_j, H] | 0 \rangle \quad (\text{E.13})$$

Equation (E.13) thus corresponds to a Hessian element with modified den-

sity matrix elements

$$\begin{aligned}\langle 0 | a_r^+ a_s | 0 \rangle &\rightarrow \langle 0 | a_r^+ a_s | \delta 0 \rangle + \langle \delta 0 | a_r^+ a_s | 0 \rangle \\ \langle 0 | a_r^+ a_s^+ a_t a_u | 0 \rangle &\rightarrow \langle \delta 0 | a_r^+ a_s^+ a_t a_u | 0 \rangle + \langle 0 | a_r^+ a_s^+ a_t a_u | \delta 0 \rangle\end{aligned}\quad (\text{E.14})$$

We have thus shown how the cubic term in Eq. (E.4) may be efficiently evaluated.

Case (2) i orbital index, j state index. In this case we may express the cubic term in Eq. (E.3) as

$$K_{ijk} \lambda_k = \langle 0 | [V_j, [U_i, U_k, H]] | 0 \rangle \kappa_k + \langle 0 | [V_j, V_k, [U_i, H]] | 0 \rangle S_{k0} \quad (\text{E.15})$$

We consider initially the first term of Eq. (E.15)

$$\begin{aligned}\langle 0 | [V_j, [U_i, U_k, H]] | 0 \rangle \kappa_k &= \langle 0 | [V_j, [U_i, [U_k, H]]] | 0 \rangle \kappa_k \\ &\quad + \tfrac{1}{2} \langle 0 | [V_j, [[U_k, U_i], H]] | 0 \rangle \kappa_k\end{aligned}\quad (\text{E.16})$$

The first term on the right-hand side of Eq. (E.16) is a Hessian element with the modified Hamiltonian

$$\langle 0 | [V_j, [U_i, \tilde{H}]] | 0 \rangle \quad (\text{E.17})$$

whereas the second term in Eq. (E.16) is a sum of standard Hessian elements.

Denoting

$$|\delta j\rangle = S_{j0}|0\rangle \quad (\text{E.18})$$

we may express the second term in Eq. (E.15) as

$$\langle 0 | [V_j, V_k, [U_i, H]] | 0 \rangle S_{k0} = -2\langle \delta j | [U_i, H] | 0 \rangle - 2\langle 0 | [U_i, H] | \delta 0 \rangle \quad (\text{E.19})$$

Eq. (E.19) describes coupling elements of the Hessian with modified density matrices

$$\begin{aligned}\langle j | a_r^+ a_s | 0 \rangle &\rightarrow \langle \delta j | a_r^+ a_s | 0 \rangle + \langle 0 | a_r^+ a_s | \delta 0 \rangle \\ \langle j | a_r^+ a_s^+ a_t a_u | 0 \rangle &\rightarrow \langle \delta j | a_r^+ a_s^+ a_t a_u | 0 \rangle + \langle j | a_r^+ a_s^+ a_t a_u | \delta 0 \rangle\end{aligned}\quad (\text{E.20})$$

We have now shown that Eq. (E.15) may be rewritten as two modified cou-

pling elements of the Hessian matrix plus an additional sum of ordinary Hessian coupling elements.

Case (3): i state index, j orbital index. Since $K_{ijk}\lambda_k = K_{jik}\lambda_k$, this term may be straightforwardly evaluated from Case (2), above.

Case (4): i, j both state indices. The cubic term in Eq. (E.3) may now be written as

$$K_{ijk}\lambda_k = \langle 0 | [V_i, V_j, [U_k, H]] | 0 \rangle \kappa_k + \langle 0 | [V_i, V_j, V_k, H] | 0 \rangle S_{k0} \quad (\text{E.21})$$

The first term in Eq. (E.21) may straightforwardly be expressed as

$$2\langle i | \tilde{H} | j \rangle - 2\delta_{ij}\langle 0 | \tilde{H} | 0 \rangle \quad (\text{E.22})$$

which corresponds to configuration interaction matrix elements of the modified Hamiltonian of Eq. (E.8). The second term in Eq. (E.21) may be written as

$$\begin{aligned} \langle 0 | [V_i, V_j, V_k, H] | 0 \rangle S_{k0} = & \frac{1}{6} \left\{ \langle 0 | [V_i, [V_j, [V_k, H]]] | 0 \rangle \right. \\ & + \langle 0 | [V_i, [V_k, [V_j, H]]] | 0 \rangle \\ & + \langle 0 | [V_j, [V_i, [V_k, H]]] | 0 \rangle \\ & + \langle 0 | [V_j, [V_k, [V_i, H]]] | 0 \rangle \\ & + \langle 0 | [V_k, [V_i, [V_j, H]]] | 0 \rangle \\ & \left. + \langle 0 | [V_k, [V_j, [V_i, H]]] | 0 \rangle \right\} S_{k0} \quad (\text{E.23}) \end{aligned}$$

Expanding the V operators gives

$$\langle 0 | [V_i, V_j, V_k, H] | 0 \rangle S_{k0} = \frac{8}{3} S_{i0} \langle 0 | H | j \rangle + \frac{8}{3} S_{j0} \langle i | H | 0 \rangle + \frac{8}{3} \delta_{ij} S_{k0} \langle k | H | 0 \rangle \quad (\text{E.24})$$

such that the term in Eq. (E.24) is just a sum of standard CI matrix elements. Of course, in the two-step MCSCF procedure, Eq. (E.24) is identically zero. Equation (E.21) can thus be written as a sum of standard and modified CI matrix elements.

E.2. Construction of the Cubic Terms for in the Iterative Cubic Procedure

In the previous section it was shown that the matrix elements of $K_{ijk}\lambda_k$ could be expressed as a sum of three terms. The first term corresponds to a

Hessian with the “Hamiltonian” $[\kappa, H]$, and normal density elements, the second term is a “Hessian” with changed density elements and the hamiltonian H . The third term is a sum of Hessian elements and gradient elements. A straightforward construction of $\mathbf{K}\lambda$ requires thus the construction of two Hessians for each choice of λ . It is, however, possible to construct $\mathbf{G} + \mathbf{K}\lambda$ to the required accuracy (first order in λ) by only one Hessian construction. To do this, introduce

$$\begin{aligned}|0'\rangle &= |0\rangle + |\delta 0\rangle \\|j'\rangle &= |j\rangle + |\delta j\rangle \\H' &= H + [\kappa, H].\end{aligned}\quad (\text{E.25})$$

If second-order terms are neglected, the four types of elements $G_{ij} + K_{ijk}\lambda_k$ separately become

1. $\langle 0' | [Q_i^+ - Q_i, Q_j^+ - Q_j, H'] | 0' \rangle + G'_{ij}$
2. $-2\langle j' | [Q_i^+ - Q_i, H'] | 0' \rangle + G'_{ij}$
3. $-2\langle i' | [Q_j^+ - Q_j, H'] | 0' \rangle + G'_{ij}$
4. $2\langle i' | H' | j' \rangle - 2\delta_{ij} \langle 0' | H | 0' \rangle + G'_{ij}$

where G'_{ij} denotes the sum of Hessian and gradient elements.

In order to construct $(G_{ij} + K_{ijk}\lambda_k)$, the first-order transformed hamiltonian H' must be constructed. An algorithm based directly on Eq. (E.9) requires

$$N'_\mu = \left(\frac{3}{2}r^2 - 2r^4 + r^5\right) N^5 \quad (\text{E.27a})$$

multiplications where r is n/N . The operation count in Eq. (E.27a) is based on the assumption that the only κ elements which are identically zero occur when two indices correspond to unoccupied orbitals.²⁰ Often a very large fraction of the zero κ elements occur when both indices correspond to occupied orbitals. In this case the operation count is about²⁰

$$N''_\mu = \left(\frac{3}{2}r^2 - \frac{3}{2}r^3 - \frac{1}{2}r^4 + \frac{1}{2}r^5\right) N^5 \quad (\text{E.27b})$$

The timings for carrying out this microtransformation are discussed in Appendix F. The microtransformation is not the most time-consuming step in the setting up $(G_{ij} + K_{ijk}\lambda_k)$.

Construction of \mathbf{G}' from \mathbf{G} and \mathbf{F} are not time consuming. The operation count in construction of the orbital-orbital part of \mathbf{G}' is proportional to $r^3 N^5$.

E.3. Calculation of the Cubic Terms in the Perturbative Cubic Procedure

In the perturbative cubic approach [Eq. (137)] only a vector $K_{ijk}^{NR} \lambda_i^{NR} \lambda_k$ is needed. The vector $K_{ijk}^{NR} \lambda_i^{NR} \lambda_k$ can be calculated through a gradient routine, by simple extensions of the ideas developed previously.

APPENDIX F: TWO-ELECTRON INTEGRAL TRANSFORMATIONS IN MCSCF

Transformations of two-electron integrals are the most time-consuming part of an MCSCF iteration when a small or medium sized CI expansion is used. Even when a thousand symmetrized configuration states are included, the transformation occupies a large fraction of an iteration.^{35, 84} The iteration procedures discussed in Sections III, V, VI, VII, and VIII require different transformations. An understanding of each of the transformations' relative complexity is therefore required for combining the methods optimally. In the following we first outline the complexity of usual transformations.^{79, 93–95} We then show how the techniques discussed in Section IX.A simplify the transformations. The number of multiplications directly programmed (the operation count) is used to measure the complexity of a given algorithm. This index is easy to calculate, but it is not a precise indicator of relative CPU times due to neglect of differences at machine language level and the like. Only the terms in the operation count that are of fifth order are included.

F.1. Traditional Two-Electron Integral Transformations

In fixed Hessian methods and update methods only gradients are calculated. This involves integrals (see Appendix C and Section III.D) with three indices confined to occupied (both completely and partially) orbitals. The transformation T_1 required to construct these integrals can be depicted as

$$T_1: (\alpha\beta|\gamma\delta) \rightarrow (ph_3|h_2h_1) \quad (\text{F.1})$$

In Eq. (F.1) and in the following molecular orbitals in general are denoted r, s, t, u and basis orbitals are denoted by $\alpha, \beta, \gamma, \dots$. Fully or partly occupied molecular orbitals are indexed as h_1, h_2, \dots , and completely unoccupied orbitals are indexed as p_1, p_2, \dots . Simplifications can sometimes be obtained if indices corresponding to occupied orbitals are divided into indices corresponding to inactive orbitals and indices corresponding to active orbitals.³⁵ However, for simplicity we will neglect this reduction.

The transformation T_1 can be straightforwardly carried out as indicated in Ref. 20. The operation count N_1 for T_1 is readily obtained

$$N_1 = \left(\frac{r}{2} + \frac{r^2}{4} + r^3 - \frac{3}{8}r^4 \right) N^5 \quad (\text{F.2})$$

In Eq. (F.2) r is n/N , where n is the number of occupied (both fully and partially) orbitals and N is the total number of orbitals. A straightforward T_1 transformation is shown in Fig. 2.

In second-order procedures the Hessian is also required. Explicit formulas for "Hessian"-like elements are given in Appendix C and Section III.D.

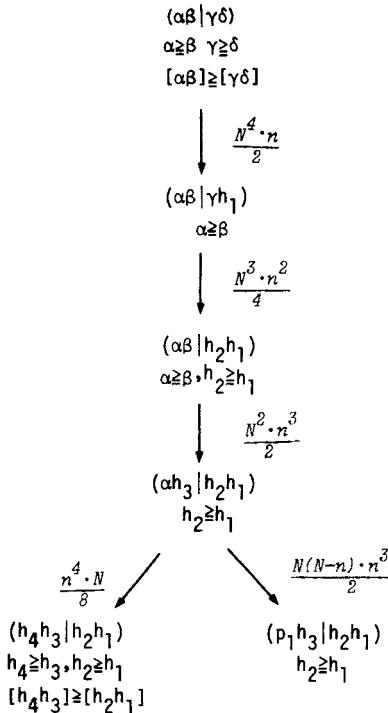


Fig. 2. Outline of a straightforward two-electron integral transformation, T_1 , used in gradient-based methods. Multiplication counts for each step is indicated. Here, N is the total number of orbitals, and n is the number of occupied orbitals. The canonical index $[\alpha\beta]$ equals $\max(\alpha, \beta) \times [\max(\alpha, \beta) + 1]/2 + \min(\alpha, \beta)$.

From Appendix C and Section III.D it may be seen that the following integrals are needed:

$$\begin{aligned} & (rs|h_2 h_1) \\ & (rh_2|sh_1) \end{aligned} \tag{F.3}$$

The transformations that set up the sets of integrals indicated in Eq. (F.3) have been studied by Ruedenberg et al.⁷⁹ They propose two algorithms T_2 and T'_2 with operation counts N_2 and N'_2 :

$$N_2 = \left(r + \frac{3}{2}r^2 - \frac{11}{6}r^3 + \frac{3}{8}r^4 \right) N^5 \tag{F.4}$$

$$N'_2 = \left(\frac{r}{2} + \frac{5}{2}r^2 - \frac{11}{6}r^3 + \frac{3}{8}r^4 \right) N^5 \tag{F.5}$$

Ruedenberg et al.⁷⁹ argue that the smaller operation count in T'_2 is often off-

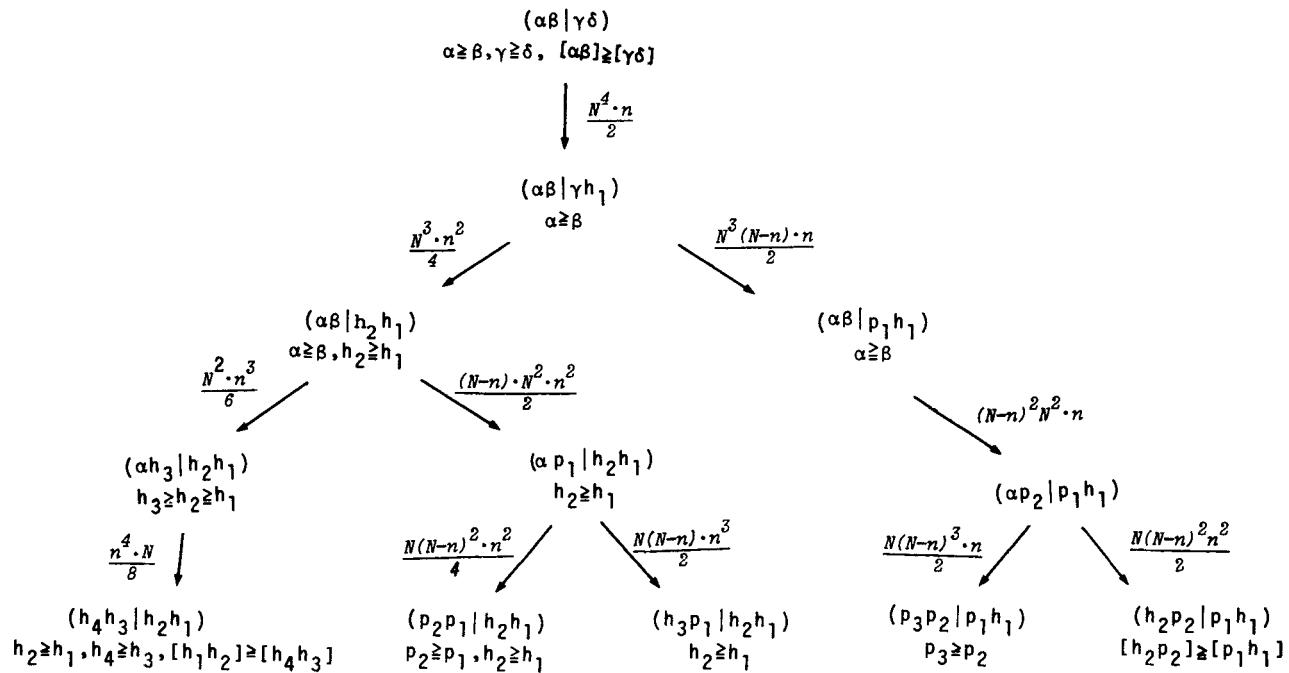


Fig. 3. Outline of a straightforward two-electron integral transformation, T_3 , used in one-point cubic methods. See the legend to Fig. 2 for further explanation.

set by increased I/O processing. As a result of this they implemented T_2 in the ALIS program package.

Two very different transformations are involved in the one-point cubic procedures (iterative and perturbative cubic). Initially a transformation T_3 is required

$$T_3: (\alpha\beta|\gamma\delta) \rightarrow (rs|th_1) \quad (F.6)$$

A straightforward T_3 algorithm is sketched in Fig. 3. The corresponding operation count N_3 is

$$N_3 = \left(\frac{5}{2}r - \frac{5}{2}r^2 + \frac{7}{6}r^3 - \frac{1}{8}r^4 \right) N^5 \quad (F.7)$$

As shown in Appendix E a micro-transformation T_μ is also required in these cubic procedures. The corresponding operation count N_μ has the limits (see Appendix E)

$$\begin{aligned} N'_\mu &= \left(\frac{3}{2}r^2 - 2r^4 + r^5 \right) N^5 \\ N''_\mu &= \left(\frac{3}{2}r^2 - \frac{3}{2}r^3 - \frac{1}{2}r^4 + \frac{1}{2}r^5 \right) N^5 \end{aligned} \quad (F.8)$$

Although none of the procedures discussed requires a complete two-electron transformation, it is of interest to consider such a transformation for purposes of comparison. A very efficient method for the complete transformation, T_4 , is suggested by Elbert.⁹⁵ The operation count is

$$N_4 = \frac{25}{24} N^5 \quad (F.9)$$

Values of N_1 , N_2 , N'_2 , N_3 , N'_μ , and N''_μ are shown in Table F.1 for the typical values $r = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$. The operation counts N_2 , N_3 , N_4 go toward the same limit as $r \rightarrow 1$, while N_1 becomes greater than N_4 in this limit. This can be traced back to the third step of T_1 in Fig. 2. N'_2 also becomes greater than N_4 as $r \rightarrow 1$. Table F.1 indicates that T_3 is somewhat more involved than $T'_2(T_2)$ and

TABLE F.1
Operation Counts for Conventional Transformations^a

r	N_1	N_2	N'_2	N_3	N'_μ	N''_μ
$\frac{1}{2}$	0.41	0.67	0.67	0.76	0.28	0.17
$\frac{1}{3}$	0.23	0.44	0.38	0.60	0.15	0.11
$\frac{1}{4}$	0.15	0.32	0.25	0.49	0.09	0.07
$\frac{1}{5}$	0.12	0.25	0.19	0.41	0.06	0.05

^aIn units of N^5

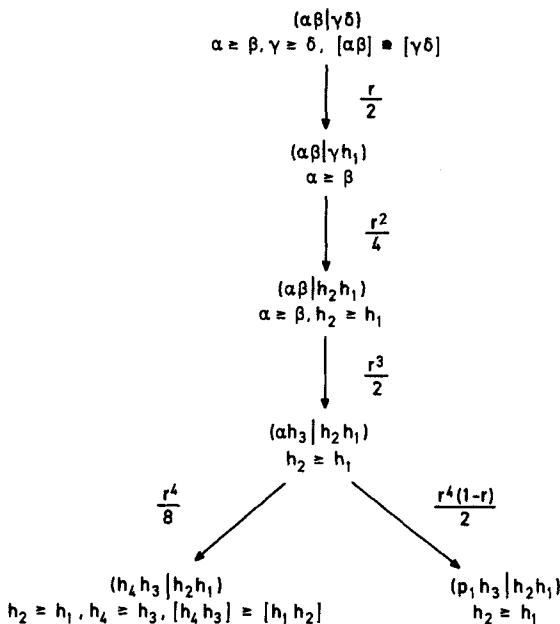


Fig. 4. Outline of the reduced two-electron transformation, T_1 . Multiplication counts ($\times N^5$) corresponding to the simplified expansion matrix [Eq. (252)] are shown; r is n/N . See the legend to Fig. 2 for further explanation.

that $T'_2(T_2)$ is more involved than T_1 . This is due to the increased number of integrals with indices corresponding to unoccupied orbitals.

F.2. Simplified Transformations

In Section IX.A a method to reduce the number of multiplications in the transformations was outlined. The method changed the expansion matrix C so that an unoccupied molecular orbital was spanned by only $n + 1$ (instead of N) basis orbitals. Reductions in the expansion of a completely occupied orbital were also made, but this will not concern us now. We shall discuss here how these simplifications affect the operation counts of various transformations.

We first discuss how the gradient transformation T_1 is affected by the reduction of the matrix C' [Eq. (252)]. The reduced T_1 transformation is summarized in Fig. 4. By summing the operation counts for the individual steps, the total number of multiplications for the reduced T_1 is

$$\tilde{N}_1 = \left(\frac{r}{2} + \frac{r^2}{4} + \frac{r^3}{2} + \frac{5}{8}r^4 - \frac{r^5}{2} \right) N^5 \quad (\text{F.9a})$$

TABLE F.2
Comparison Between Operation Counts
for Normal (N_1) and Reduced (\tilde{N}_1)
Two-Electron Integral Transformation
Required for Gradient Procedures^a

r	N_1	\tilde{N}_1
$\frac{1}{2}$	0.41	0.40
$\frac{1}{3}$	0.23	0.22
$\frac{1}{4}$	0.15	0.15
$\frac{1}{5}$	0.12	0.11

^aIn units of N^5 .

In Table F.2, \tilde{N}_1 is compared to N_1 for our usual choices of r . It is seen that \tilde{N}_1 is not significantly smaller than N_1 . This transformation consists primarily of summing over occupied orbitals, so the small difference between N_1 and \tilde{N}_1 is not surprising.

For second-order procedures T_2 and T'_2 transformations can be used to set up the necessary $(\tilde{rs}|\tilde{tu})$. From Fig. 5 it is seen that an application of T_2 in

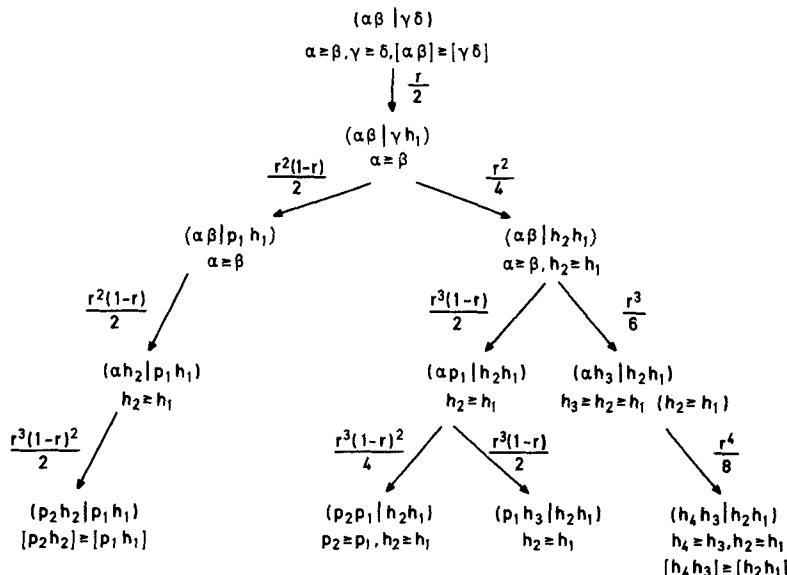


Fig. 5. Outline of the reduced two-electron transformations, \tilde{T}_2 . See the legends to Figs. 2 and 4 for further explanation.

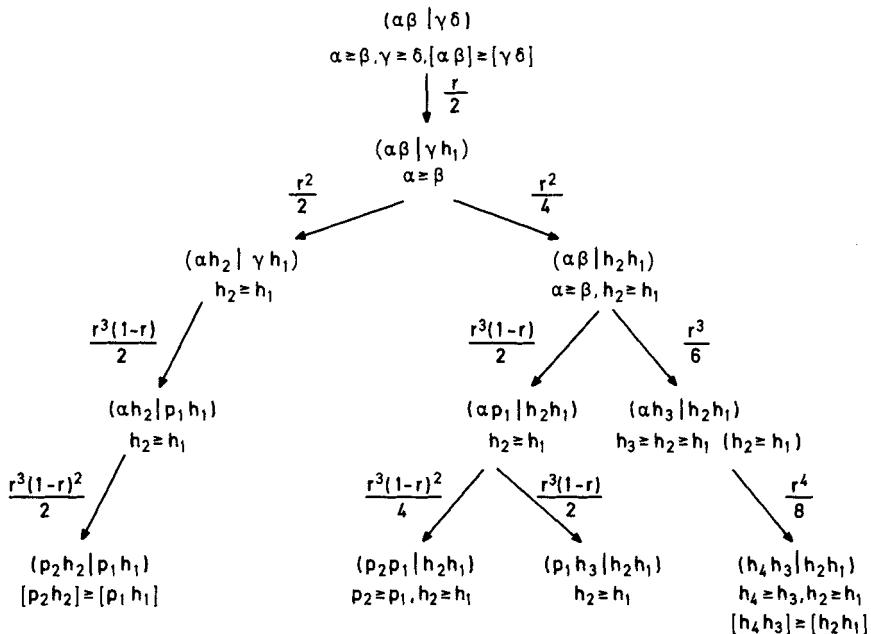


Fig. 6. Outline of the reduced two-electron integral transformations, \tilde{T}'_2 . See the legends to Figs. 2 and 4* for further explanation.

TABLE F.3
 Operation Counts for Reduced (\tilde{N}_2 and \tilde{N}'_2) and Normal (N_2 and N'_2) Two-Electron Integral Transformations Required for Second-Order MCSCF Procedures^a

r	N_2	\tilde{N}_2^b	N'_2	\tilde{N}'_2^b
$\frac{1}{2}$	0.67	0.55	0.67	0.58
$\frac{1}{3}$	0.44	0.31	0.38	0.31
$\frac{1}{4}$	0.32	0.21	0.25	0.20
$\frac{1}{5}$	0.25	0.15	0.19	0.14

^aIn units of N^5 .

^bFrom Eqs. (F.10) and (F.11).

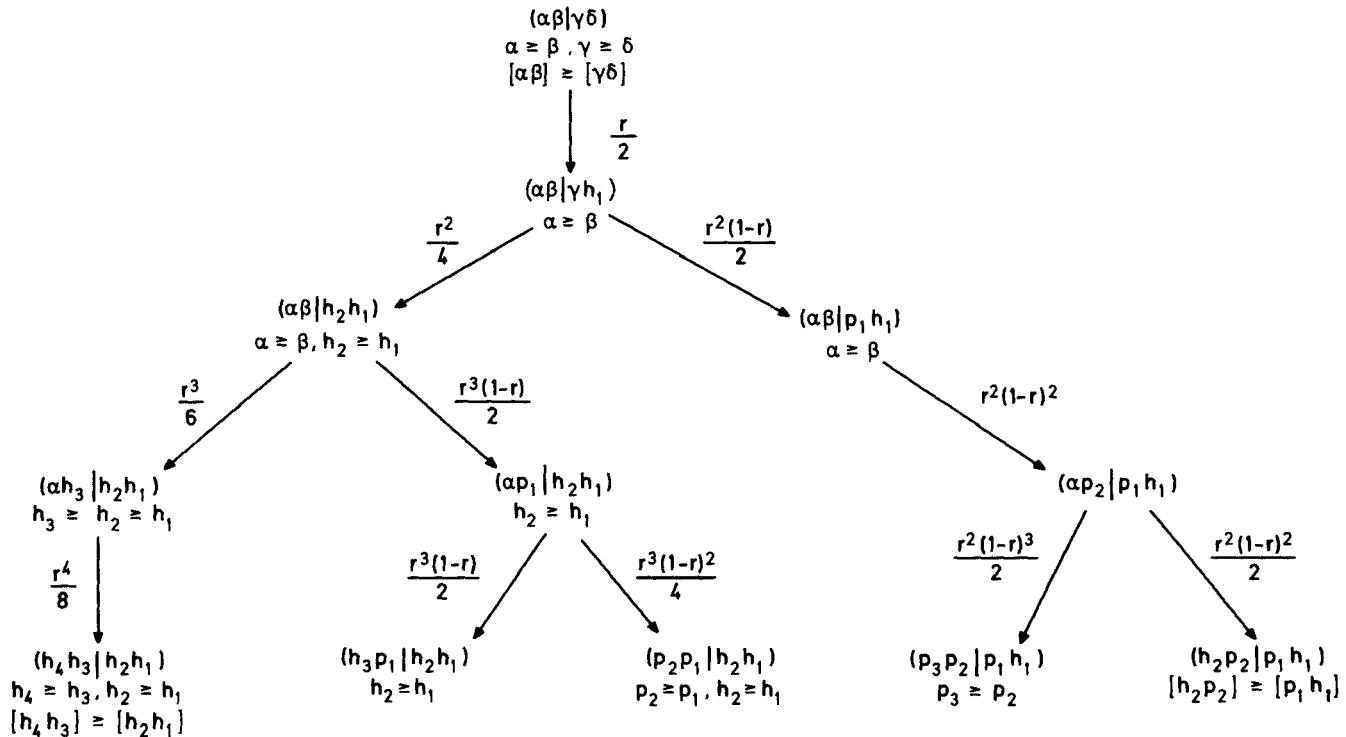


Fig. 7. Outline of the reduced two-electron transformations, \tilde{T}_3 . See the legends to Figs. 2 and 4 for further explanation.

connection with C' gives an operation count

$$\tilde{N}_2 = \left(\frac{r}{2} + \frac{5}{4}r^2 + \frac{11}{12}r^3 - \frac{19}{8}r^4 + \frac{3}{4}r^5 \right) N^5 \quad (\text{F.10})$$

When T'_2 is used the modified operation count is

$$\tilde{N}'_2 = \left(\frac{r}{2} + \frac{3}{4}r^2 + \frac{29}{12}r^3 - \frac{23}{8}r^4 + \frac{3}{4}r^5 \right) N^5 \quad (\text{F.11})$$

The operation counts \tilde{N}_2 and \tilde{N}'_2 are reported in Table F.3. It may be seen that operation counts for T_2 are especially decreased. The greater change in operation counts of T_2 is due to the larger part of the T_2 transformation that consists of transforming to indices corresponding to unoccupied orbitals.

The major part of the original transformation T_3 [Eq. (F.6)] corresponds to summing over expansions of unoccupied orbitals. Thus, the simplified expansion matrix will decrease the operation count very significantly for T_3 . The operation counts for the individual steps of the reduced T_3 transformation are given in Fig. 7. The reduced total operation count is

$$\tilde{N}_3 = \left(\frac{r}{2} + \frac{11}{4}r^2 - \frac{43}{12}r^3 + \frac{13}{8}r^4 - \frac{1}{4}r^5 \right) N^5 \quad (\text{F.12})$$

Values of \tilde{N}_3 are compared to N_3 for different values of r in Table F.4. As expected, very significant reductions are encountered; for example, the reduced operation count \tilde{N}_3 is about 50% of the operation count of N_3 for $r = \frac{1}{4}$ or $r = \frac{1}{5}$. By comparing Table F.4 and Table F.1 it is furthermore seen

TABLE F.4
Operation Counts for Normal (N_3) and
Simplified (\tilde{N}_3) Two-Electron Integral
Transformations Needed for Cubic
MCSCF Procedures^a

r	N_3	\tilde{N}_3
$\frac{1}{2}$	0.76	0.58
$\frac{1}{3}$	0.60	0.36
$\frac{1}{4}$	0.49	0.25
$\frac{1}{5}$	0.41	0.18

^aIn units of N^5 . Here, $r = n/N$, where n is the number of occupied orbitals and N is the total number of orbitals.

that the operation counts for the reduced cubic transformation are below operations counts for the original T_2 transformations.

While the operation counts N_1, N_2, N'_2, N_3 differ significantly for small values of r , the reduced operation counts $\tilde{N}_1, \tilde{N}_2, \tilde{N}'_2, \tilde{N}_3$ are very similar for small values of r . For $r = \frac{1}{3}$ it is seen that \tilde{N}_1 is 0.11, \tilde{N}_2 is 0.15, and \tilde{N}_3 is 0.18. Although it is not advisable to use these numbers to make precise quantitative predictions about relative computer processing times, they do indicate that the reduced T_3 and T_2 transformations can be carried out with computer times of the same magnitude as the reduced T_1 transformation.

Acknowledgments

We would especially like to thank Esper Dalgaard and Jan Linderberg for their many helpful comments and suggestions; Diane Lynch, Jeff Nichols, Preben Albertsen, and Peter Swanstrøm for theoretical and computational assistance; and Birgitte Buus (Aarhus) and Sandy Seligman (Texas A&M) for scientific editing and typing.

We also appreciate receiving copies of many MCSCF articles prior to publication, including preprints from E. Dalgaard, D. G. Hopper, J. Detrich, C. C. J. Roothaan, B. Lengsfeld, R. Shepard, I. Shavitt, J. Simons, H.-J. Werner, W. Meyer, C. Nellin, F. A. Matsen, G. Bacskay, B. Brooks, W. Laidig, H. F. Schaefer, N. Kosugi, H. Kuroda, G. Das, B. Liu, J. van Lenthe, and D. Yarkony.

J.O., P.J. and D.L.Y. acknowledge research support from the National Science Foundation (Grant CHE-8023352) and the Robert A. Welch Foundation (Grant A-770). J.O. also acknowledges support from the Danish Natural Science Research Council, and D.L.Y. has received travel assistance from NATO. (Grant RG 193.80)

References

1. D. R. Hartree, W. Hartree, and B. Swirles, *Phil. Trans. Roy. Soc. (London)* **A238**, 229 (1939); G. Das and A. C. Wahl, *J. Chem. Phys.* **44**, 87 (1966); A. Veillard, *Theoret. Chim. Acta* **4**, 22 (1966); J. Hinze and C. C. J. Roothaan, *Progr. Theoret. Phys. Suppl.* **40**, 37 (1967); see, for example, A. C. Wahl and G. Das, in *Modern Theoretical Chemistry*, H. J. Schaefer III, ed., Plenum Press, New York, 1977, and references therein.
2. See, for example, *Recent Developments and Applications of Multiconfiguration Hartree-Fock Methods*, Proceedings of the National Resource for Computational Chemistry Workshop held at Texas A&M University, July 1980, M. Dupuis, ed.
3. U. Fock, *Z. Physik* **61**, 126 (1930); J. C. Slater, *Phys. Rev.* **35**, 210 (1930).
4. G. G. Hall, *Proc. Roy. Soc. (London)* **A205**, 541 (1951); G. G. Hall, *Proc. Roy. Soc. (London)* **A208**, 328; C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951); R. K. Nesbet, *Proc. Roy. Soc. (London)* **A230**, 312 (1955); C. C. J. Roothaan, *Rev. Mod. Phys.* **32**, 179 (1960); S. Huzinaga, *Phys. Rev.* **120**, 866 (1960); S. Huzinaga, *Phys. Rev.* **122**, 131 (1961); F. W. Birss and S. Fraga, *J. Chem. Phys.* **38**, 2252 (1963); F. W. Birss and S. Fraga, *J. Chem. Phys.* **40**, 3203, 3207, 3212 (1964); C. C. J. Roothaan and P. S. Bagus, *Methods Comp. Phys.* **2**, 47 (1963); W. J. Hunt, T. H. Dunning, and W. A. Goddard, *Chem. Phys. Lett.* **3**, 609 (1969).
5. T. A. Koopmans, *Physica* **1**, 104 (1933).

6. Yale Babylonian Collection Tablet No. 7289; see also Otto Neugebauer and A. Sachs, *Mathematical Cuneiform Texts*, American Oriental Society, New Haven, Conn., 1945.
7. A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, 2nd ed., McGraw-Hill, New York, 1978.
8. J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, N.J., 1964.
9. R. Fletcher, *Practical Methods of Optimization*, Vol. 1, Wiley, New York, 1980.
10. E. Dalgaard and P. Jørgensen, *J. Chem. Phys.* **69**, 3833 (1978).
11. D. L. Yeager and P. Jørgensen, *J. Chem. Phys.* **71**, 755 (1979); D. L. Yeager, 1979 Sanibel Symposia.
12. E. Dalgaard, *Chem. Phys. Lett.* **65**, 559 (1979).
13. D. L. Yeager and P. Jørgensen, *Mol. Phys.* **39**, 587 (1980).
14. P. Jørgensen, P. Albertsen, and D. L. Yeager, *J. Chem. Phys.* **72**, 6466 (1980); **73**, 5408 (1980).
15. D. L. Yeager, P. Albertsen, and P. Jørgensen, *J. Chem. Phys.* **73**, 2811 (1980).
16. P. Jørgensen, J. Olsen, and D. L. Yeager, *J. Chem. Phys.* **75**, 5802 (1981).
17. J. Olsen, P. Jørgensen, and D. L. Yeager, *J. Chem. Phys.* **76**, 527 (1982).
18. D. L. Yeager, D. Lynch, J. Nichols, P. Jørgensen, and J. Olsen, *J. Phys. Chem.* **86**, 2140 (1982).
19. A. Igawa, D. L. Yeager, and H. Fukutome, *J. Chem. Phys.* **76**, 5388 (1982).
20. J. Olsen, P. Jørgensen, and D. L. Yeager, *J. Chem. Phys.* **77**, 356 (1982).
21. J. Olsen and P. Jørgensen, *J. Chem. Phys.* **77**, 6109 (1982).
22. J. Olsen, P. Jørgensen, and D. L. Yeager, "Second and Higher Order Convergence in Linear and Non-linear Multiconfigurational Hartree-Fock Theory," *Int. J. Quantum Chem.* (in press).
23. P. Jørgensen and J. Simons, *Second Quantization-Based Methods in Quantum Chemistry*, Academic Press, New York, 1981.
24. L. G. Yaffe and W. A. Goddard III, *Phys. Rev. A* **13**, 1682 (1976).
25. L. G. J. Douady, J. Ellinger, R. Subra, and B. Levy, *J. Chem. Phys.* **72**, 1452 (1980).
26. C. C. J. Roothaan, J. Detrich, and D. G. Hopper, *Int. J. Quantum Chem.* **S13**, 93 (1979); D. G. Hopper and C. C. J. Roothaan, 1979 Sanibel Symposia.
27. P. Siegbahn, A. Heiberg, B. Roos, and B. Levy, *Phys. Scr.* **21**, 323 (1980).
28. B. Lengsfeld III, *J. Chem. Phys.* **73**, 382 (1980).
29. H.-J. Werner and W. Meyer, *J. Chem. Phys.* **73**, 2342 (1980).
30. H.-J. Werner and W. Meyer, *J. Chem. Phys.* **74**, 5794 (1981).
31. R. Shepard and J. Simons, *Int. J. Quantum Chem.* **S14**, 211 (1980).
32. R. Shepard, I. Shavitt, and J. Simons, *J. Chem. Phys.* **76**, 543 (1982).
33. G. Das, *J. Chem. Phys.* **74**, 5775 (1981).
34. F. A. Matsen and C. J. Nelin, *Int. J. Quantum Chem.* **20**, 861 (1981).
35. P. Siegbahn, J. Almlöf, A. Heiberg, and B. Roos, *J. Chem. Phys.* **74**, 2384 (1981).
36. A. U. Nemukhin, J. Almlöf, and A. Heiberg, *Chem. Phys.* **57**, 197 (1981); J. Almlöf, A. U. Nemukhin, and A. Heiberg, *Int. J. Quantum Chem.* **88**, 655 (1981).

37. C. W. Bauschlicher Jr., P. S. Bagus, D. R. Yarkony, and B. Lengsfield, *J. Chem. Phys.* **74**, 3965 (1981).
38. W. J. Hunt and W. A. Goddard III, *Chem. Phys. Lett.* **3**, 414 (1969).
39. S. Iwata and K. Morokuma, *Theoret. Chim. Acta* **33**, 4972 (1974).
40. S. F. Abdulnur, J. Linderberg, Y. Öhrn, and P. W. Thulstrup, *Phys. Rev. A* **6**, 889 (1972).
41. P. O. Löwdin, in *Perturbation Theory and Its Application in Quantum Mechanics*, C. H. Wilcox, ed., Wiley, New York, 1966.
42. E. A. Hylleraas and B. Undheim, *Z. Phys.* **65**, 759 (1930); J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933).
43. D. L. Yeager and P. Jørgensen, *Chem. Phys. Lett.* **65**, 77 (1979).
44. E. Dalgaard, *J. Chem. Phys.* **72**, 816 (1980).
45. P. Albertsen, P. Jørgensen, and D. L. Yeager, *Mol. Phys.* **41**, 409 (1980).
46. P. Albertsen, P. Jørgensen, and D. L. Yeager, *Int. J. Quantum Chem.* **14S**, 249 (1980).
47. P. Albertsen, P. Jørgensen, and D. L. Yeager, *Chem. Phys. Lett.* **76**, 354 (1980).
48. D. L. Yeager, J. Olsen, and P. Jørgensen, *Int. J. Quantum Chem.* **15S**, 151 (1981).
49. J. Nichols and D. L. Yeager, *Chem. Phys. Lett.* **84**, 77 (1981).
50. D. Lynch, M. F. Herman, and D. L. Yeager, *Chem. Phys.* **64**, 69 (1982).
51. A. Banerjee, J. W. Kenney, and J. Simons, *Int. J. Quantum Chem.* **16**, 1209 (1980).
52. D. J. Thouless, *Nucl. Phys.* **21**, 255.
53. D. J. Rowe, *Nuclear Collective Motion Models and Theory*, Methuen, London, 1970.
54. J. Linderberg and Y. Öhrn, *Propagators in Quantum Chemistry*, Academic Press, New York, 1973.
- 55a. P. Jørgensen, P. Swanstrøm, and D. Yeager, *J. Chem. Phys.* **78**, 347 (1983).
- 55b. J. Golab, D. Yeager, and P. Jørgensen, "Proper Characterization of MCSCF Stationary Points," *Chem. Phys.* (submitted for publication).
56. V. R. Saunders and I. H. Hillier, *Int. J. Quantum Chem.* **7**, 699 (1973).
57. M. F. Guest and V. R. Saunders, *Mol. Phys.* **28**, 819 (1974).
58. J. E. Grabenstetter and F. Grein, *Mol. Phys.* **31**, 1469 (1976).
59. V. A. Kuprievich and O. V. Shamko, *Int. J. Quantum Chem.* **9**, 1009 (1975).
60. H. Tatewaki, K. Tanaka, F. Sasaki, S. Obara, K. Ohno, and M. Yoshimine, *Int. J. Quantum Chem.* **15**, 533 (1979).
61. G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
62. C. E. Fröberg, *Introduction to Numerical Analysis*, 2nd ed., Addison-Wesley, Reading, Mass., 1969.
63. F. E. Dennis and F. F. Moré, *SIAM Rev.* **19**, 46 (1977); E. Spedicato, in *Towards Global Optimization*, L. C. W. Dixon and G. P. Szegő, eds., North-Holland Publs., Amsterdam, 1978.
64. C. G. Broyden, *Math. Comp.* **19**, 577 (1965).
65. M. J. D. Powell, *Math. Progr.* **1**, 26 (1971).
66. H. Kleinmichel, *Num. Math.* **38**, 219 (1981); E. Spedicato, in *Towards Global Optimization*, L. C. W. Dixon and G. P. Szegő, eds., North-Holland Publs., Amsterdam, 1978, 209.
67. J. E. Dennis and J. J. Moré, *Math. Comp.* **28**, 549 (1974).

68. L. C. W. Dixon, *Opt. Theoret. Appl.* **10**, 34 (1972).
69. R. W. H. Sargent and P. I. Sebastian, in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma, ed., Academic Press, New York, 1972; D. M. Himmelblau, *ibid.*
70. C. G. Broyden, *J. Inst. Math. Appl.* **6** 70, 222 (1970); R. Fletcher, *Comp. J.* **13**, 317 (1970); D. Goldfarb, *Math. Comp.* **24**, 23 (1970); D. F. Shanno, *ibid.* **24**, 647 (1970).
71. W. Davidon, AEC Research and Development Rep. ANL 5990, Argonne National Lab., Argonne, Ill., 1959; R. Fletcher and M. Powell; *Comp. J.* **6**, 163 (1963).
72. C. G. Broyden, *Math. Comp.* **21**, 368 (1967); B. A. Murtagh and R. W. H. Sargent, in *Optimization*, R. Fletcher, ed., Academic Press, New York 1969.
73. R. H. A. Eade and M. A. Robb, *Chem. Phys. Lett.* **83**, 362 (1981).
74. J. E. Dennis and R. R. Schnabel, *SIAM Rev.* **21**, 443 (1979), and references therein.
75. D. Simmons, J. Golab, J. Olsen, and D. Yeager, *Chem. Phys. Lett.* (to be submitted).
76. W. C. Davidon, *Math. Progr.* **9**, 1 (1975).
77. S. S. Oren, *Math. Progr.* **7**, 351 (1974).
78. B. Levy and G. Berthier, *Int. J. Quantum Chem.* **2**, 307 (1968).
79. See, for example, K. Ruedenberg, L. M. Cheung, and S. T. Elbert, *Int. J. Quantum Chem.* **16**, 1069 (1979), and references therein.
80. F. Grein and T. C. Chang, *Chem. Phys. Lett.* **1**, 44 (1971); T. C. Chang and W. H. E. Schwarz, *Theoret. Chim. Acta* **44**, 45 (1977); F. Grein and A. Banerjee, *Int. J. Quantum Chem.* **59**, 147 (1975); A. Banerjee and F. Grein, *J. Chem. Phys.* **66**, 1054 (1977).
81. D. J. Rowe, *Nucl. Phys.* **A107**, 99 (1968).
82. B. Levy, *Int. J. Quantum Chem.* **4**, 297 (1970); B. Levy, *Chem. Phys. Lett.* **4**, 17 (1969).
83. R. Pariser and R. G. Parr, *J. Chem. Phys.* **21**, 446 (1953); **21**, 767 (1953); J. A. Pople, *Trans. Faraday Soc.* **42**, 1357 (1953).
84. B. Lengsfeld III and B. Liu, *J. Chem. Phys.* **75**, 478 (1981).
85. G. B. Bacskay, *Chem. Phys.* **61**, 385 (1981); G. B. Bacskay, *Chem. Phys.* **65**, 383 (1982).
86. D. Yarkony, *Chem. Phys. Lett.* **77**, 634 (1981).
87. R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
88. *Numerical Algorithms in Chemistry: Algebraic Methods*, Proceedings of a Workshop of the National Resource for Computational Chemistry held at Lawrence-Berkeley Laboratory (1978).
89. B. Roos, P. Siegbahn, in *Methods of Electronic Structure*, H. F. Schaefer III, ed., Plenum Press, New York, 1977.
90. A. Igawa and H. Fukutome, *Progr. Theoret. Phys.* **54**, 1266 (1975).
91. F. Paldus, in *Theoretical Chemistry: Advances and Perspectives*, H. Eyring and D. Henderson, eds. Vol. 2, Academic Press, New York, 1976.
92. D. Simmons, D. Lynch, and D. Yeager (to be submitted).
93. C. F. Bender, *J. Comp. Phys.* **9**, 547 (1972).
94. I. Shavitt, in *Methods in Electronic Structure Theory*, H. F. Schaefer III, ed., Plenum Press, 1977.
95. S. Elbert, in *Numerical Algorithms in Chemistry: Algebraic Methods*, Proceedings of a Workshop of the National Resource for Computation in Chemistry held at Lawrence-Berkeley Laboratory, 1978.