JOURNAL OF COMPUTATIONAL PHYSICS 74, 265–282 (1988)

Solution of the Large Matrix Equations Which Occur in Response Theory

JEPPE OLSEN, HANS JØRGEN AA. JENSEN, AND POUL JØRGENSEN

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

Received May 9, 1986; revised March 25, 1987

The eigenvalue equation and the sets of linear equations that occur in linear and nonlinear response function calculations have a specific paired structure. We have developed iterative algorithms which utilize this structure to efficiently solve the equations. The algorithms have been designed so they do not require the matrices to be explicitly available, which makes it possible to perform accurate calculations with a dimension of maybe 10^6. Numerical tests show convergence behavior superior to previously suggested algorithms.

I. INTRODUCTION

The simplest response function model is the time-dependent Hartree–Fock (TDHF), also called the random phase approximation (RPA), which is obtained by analyzing the response of a single configuration self-consistent-field (SCF) state to an external perturbation [1]. Transition properties and second and higher order molecular response properties can in principle be calculated very efficiently using a response function approach [1]. Response function models have also been derived for more sophisticated electronic wave functions, such as a multiconfigurational MCSCF state [2–4] and a coupled cluster state [5], and also using a perturbation approach [6]. Common features of these models are (1) a generalized eigenvalue problem with a specific paired structure must be solved to get transition properties and (2) a set of linear equations with the same structured matrices must be solved to get the second and higher order properties. Furthermore the matrices may be very large, approximately 10^6 and in a few years maybe up to 10^7. It is therefore of crucial importance for the response function approach to have efficient methods for solving these two types of equations, preferably methods which do not require explicit matrices. We describe such a new algorithm for each of the two types in Section II, and the efficiency is demonstrated in Section III by numerical examples. We denote the generalized eigenvalue equation, which is encountered in time-dependent response theory, the linear response (LR) eigenvalue equation and the sets of linear equations, which are encountered, for the LR equations.

0021-9991/88 $3.00

For ground state calculations the LR eigenvalue equation may be expressed as two symmetric eigenvalue problems of half the dimension [1]. For small dimensions the solutions are straightforwardly obtained this way, but the dimension of the matrices prohibits this approach even for many RPA calculations of interest. Iterative RPA algorithms have therefore been developed to find the lowest few roots [8, 10] (the lowest excitation energies) which usually are those of chemical interest. These algorithms are similar to the iterative algorithms [11–18] used to find the lowest few eigenvalues of the symmetric eigenvalue problem occurring in configuration interaction (CI) studies of correlation. The developed iterative RPA algorithms solve the RPA eigenvalue equation as a nonsymmetric eigenvalue problem without using the paired structure [8, 10].

In this paper we propose iterative algorithms, which employ the paired structure, for obtaining the interesting roots of the LR eigenvalue equations and for solving the LR equations. The iterative algorithms consist in expanding a basis of orthonormal trial vectors and projecting the exact equations down on this basis using only matrix times vector operations. The paired structure of the LR matrices defines a paired vector for each trial vector, and each time one matrix times vector operation has been carried out we also know the matrix times the paired vector. The process of adding new pairs of trial vectors is continued until the desired root/property is adequately described in the basis of trial vectors. We have a common framework for all of the equations because the same matrices enter the LR eigenvalue equation and the LR equations irrespectively of the type of perturbation. This enables us to reuse information gathered in one such calculation in any other LR calculation for that state vector.

For the eigenvalue problem the algorithm may be viewed as a generalization of the Davidson–Liu algorithm [14, 15] for the symmetric eigenvalue problem. In normal cases, where the second energy derivative matrix (one of the linear response matrices) is positive definite, complex eigenvalues cannot show up in our reduced LR eigenvalue equation, furthermore the eigenvalues in the reduced space will converge monotonically and be upper bounds to the exact eigenvalues [19]. In the previously developed algorithms~~,~~ where nonsymmetric eigenvalue equations are solved, complex eigenvalues can show up in the reduced space, indicating the paired trial vector is missing, and suggesting a slower and potentially unreliable convergence.

Summarizing, we get two trial vectors for the cost of one, the reduced (projected) problems have the same paired structure as the full equations which stabilize the convergence processes, and we may reuse information from one property for other properties.

In Section II.a we present our new algorithm for the LR eigenvalue problem and in Section II.b we describe a very similar algorithm for the LR property equations based on a generalization of the conjugate gradient method [20]. In Section III we demonstrate by numerical examples the convergence characteristics of the algorithms. Section IV contains some concluding remarks.

II. ALGORITHMS

II.a. *The Linear Response Eeigenvalue Equation*

The linear response (LR) eigenvalue equation for a real wave function is in its most general form [4]

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega \begin{pmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{pmatrix}\right]\begin{pmatrix} {}^1\mathbf{X} \\ {}^2\mathbf{X} \end{pmatrix} = 0, \tag{1}$$

where the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{\Sigma}$ are symmetric and $\mathbf{\Delta}$ is antisymmetric. For MCSCF wave functions $\mathbf{A}$, $\mathbf{B}$, $\mathbf{\Sigma}$, and $\mathbf{\Delta}$ are defined in Eqs. (5.62), (5.63), (5.65), and (5.66), respectively, of Ref. [4]. For a closed shell Hartree–Fock wave function (the random phase approximation) $\mathbf{\Delta}$ is the zero matrix and $\mathbf{\Sigma}$ is the unit matrix. A simple rearrangement of Eq. (1) gives

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} + \omega \begin{pmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{pmatrix}\right]\begin{pmatrix} {}^2\mathbf{X} \\ {}^1\mathbf{X} \end{pmatrix} = 0 \tag{2}$$

which shows that $-\omega$ is an eigenvalue corresponding to the eigenvector $\left(\begin{smallmatrix} {}^2\mathbf{X} \\ {}^1\mathbf{X} \end{smallmatrix}\right)$. If we assume that the eigenvector $\left(\begin{smallmatrix} {}^1\mathbf{X} \\ {}^2\mathbf{X} \end{smallmatrix}\right)$ can be normalized to one over the metric $\left(\begin{smallmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{smallmatrix}\right)$,

$$\begin{pmatrix} {}^1\tilde{\mathbf{X}} & {}^2\tilde{\mathbf{X}} \end{pmatrix}\begin{pmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{pmatrix}\begin{pmatrix} {}^1\mathbf{X} \\ {}^2\mathbf{X} \end{pmatrix} = 1 \tag{3}$$

then the eigenvector $\left(\begin{smallmatrix} {}^2\mathbf{X} \\ {}^1\mathbf{X} \end{smallmatrix}\right)$ is normalized to minus one:

$$\begin{pmatrix} {}^2\tilde{\mathbf{X}} & {}^1\tilde{\mathbf{X}} \end{pmatrix}\begin{pmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{pmatrix}\begin{pmatrix} {}^2\mathbf{X} \\ {}^1\mathbf{X} \end{pmatrix} = -1. \tag{4}$$

The eigenvalues corresponding to the eigenvectors that can be normalized to one approximate the excitation energies of the molecular system [1].

For large dimensions of the LR matrices it is necessary to find the solution vectors using direct methods, that is, with algorithms that just require the two linear transformations

$$\begin{pmatrix} {}^1\mathbf{u} \\ {}^2\mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}\begin{pmatrix} {}^1\mathbf{b} \\ {}^2\mathbf{b} \end{pmatrix} \tag{5}$$

$$\begin{pmatrix} {}^1\mathbf{m} \\ {}^2\mathbf{m} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{\Sigma} \end{pmatrix}\begin{pmatrix} {}^1\mathbf{b} \\ {}^2\mathbf{b} \end{pmatrix} \tag{6}$$

to be carried out. In Section VI.B of Ref. [4] we showed how the linear transformations may be implemented for MCSCF wave functions without explicitly setting up the linear response matrices. The topic of this section is to describe an algorithm which finds the lowest roots of Eq. (1) using a sequence of the linear transformations in Eqs. (5) and (6).

In the algorithm we describe for determining the few lowest roots of Eq. (1) we assume that in the nth iteration we have a set of k trial vectors

$$\left\{\begin{pmatrix}{}^1\underline{b}\\ {}^2\underline{b}\end{pmatrix}_1, \begin{pmatrix}{}^1\underline{b}\\ {}^2\underline{b}\end{pmatrix}_2, \begin{pmatrix}{}^1\underline{b}\\ {}^2\underline{b}\end{pmatrix}_2, \ldots, \begin{pmatrix}{}^1\underline{b}\\ {}^2\underline{b}\end{pmatrix}_k\right\} \tag{7}$$

and have carried out the linear transformations in Eqs. (5) and (6)

$$\left\{\begin{pmatrix}{}^1\underline{u}\\ {}^2\underline{u}\end{pmatrix}_1, \begin{pmatrix}{}^1\underline{u}\\ {}^2\underline{u}\end{pmatrix}_2, \begin{pmatrix}{}^1\underline{u}\\ {}^2\underline{u}\end{pmatrix}_k\right\} \tag{8}$$

$$\left\{\begin{pmatrix}{}^1\underline{m}\\ {}^2\underline{m}\end{pmatrix}_1, \begin{pmatrix}{}^1\underline{m}\\ {}^2\underline{m}\end{pmatrix}_2, \begin{pmatrix}{}^1\underline{m}\\ {}^2\underline{m}\end{pmatrix}_k\right\}. \tag{9}$$

Because of the structure of the two LR matrices the two linear transformations are also known for the vectors

$$\left\{\begin{pmatrix}{}^2\underline{b}\\ {}^1\underline{b}\end{pmatrix}_1, \begin{pmatrix}{}^2\underline{b}\\ {}^1\underline{b}\end{pmatrix}_2, \ldots, \begin{pmatrix}{}^2\underline{b}\\ {}^1\underline{b}\end{pmatrix}_k\right\}. \tag{10}$$

The linear transformation in Eq. (5) gives

$$\left\{\begin{pmatrix}{}^2\underline{u}\\ {}^1\underline{u}\end{pmatrix}_1, \begin{pmatrix}{}^2\underline{u}\\ {}^1\underline{u}\end{pmatrix}_2, \ldots, \begin{pmatrix}{}^2\underline{u}\\ {}^1\underline{u}\end{pmatrix}_k\right\} \tag{11}$$

and the linear transformation in Eq. (6) gives

$$\left\{-\begin{pmatrix}{}^2\underline{m}\\ {}^1\underline{m}\end{pmatrix}_2, -\begin{pmatrix}{}^2\underline{m}\\ {}^1\underline{m}\end{pmatrix}_2, \ldots, -\begin{pmatrix}{}^2\underline{m}\\ {}^1\underline{m}\end{pmatrix}_k\right\}. \tag{12}$$

We assume that the combined set of paired trial vectors in Eqs. (8) and (11) have been orthonormalized. The optimal linear combination of the basis vectors in Eqs. (8) and (11) is determined from the reduced linear response eigenvalue equation

$$\left[\begin{pmatrix}\underset{\sim}{A}^R & \underset{\sim}{B}^R\\ \underset{\sim}{B}^R & \underset{\sim}{A}^R\end{pmatrix} - \omega^R\begin{pmatrix}\underset{\sim}{\Sigma}^R & \underset{\sim}{\Delta}^R\\ -\underset{\sim}{\Delta}^R & -\underset{\sim}{\Sigma}^R\end{pmatrix}\right]\begin{pmatrix}{}^1\underline{X}^R\\ {}^2\underline{X}^R\end{pmatrix} = \underline{0} \tag{13}$$

where the matrices $\underset{\sim}{A}^R$, $\underset{\sim}{B}^R$, and $\underset{\sim}{\Sigma}^R$ are symmetric and $\underset{\sim}{\Delta}^R$ is antisymmetric and defined as

$$A^R_{lm} = ({}^1\tilde{\underline{b}} \;\; {}^2\tilde{\underline{b}})_l \begin{pmatrix}{}^1\underline{u}\\ {}^2\underline{u}\end{pmatrix}_m = ({}^1\tilde{\underline{b}} \;\; {}^2\tilde{\underline{b}})_l \begin{pmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B} & \underset{\sim}{A}\end{pmatrix}\begin{pmatrix}{}^1\underline{b}\\ {}^2\underline{b}\end{pmatrix}_m \tag{14}$$

$$B^R_{lm} = ({}^1\tilde{\underline{b}} \;\; {}^2\tilde{\underline{b}})_l \begin{pmatrix}{}^2\underline{u}\\ {}^1\underline{u}\end{pmatrix}_m = ({}^2\tilde{\underline{b}} \;\; {}^1\tilde{\underline{b}})_l \begin{pmatrix}{}^1\underline{u}\\ {}^2\underline{u}\end{pmatrix}_m \tag{15}$$

$$\Sigma^R_{lm} = ({}^1\tilde{\underline{b}} \quad {}^2\tilde{\underline{b}})_l \begin{pmatrix} {}^1\underline{m} \\ {}^2\underline{m} \end{pmatrix}_m \tag{16}$$

$$\Delta^R_{lm} = ({}^1\tilde{\underline{b}} \quad {}^2\tilde{\underline{b}})_l \begin{pmatrix} {}^2\underline{m} \\ {}^1\underline{m} \end{pmatrix}_m = -({}^2\tilde{\underline{b}} \quad {}^1\tilde{\underline{b}})_l \begin{pmatrix} {}^1\underline{m} \\ {}^2\underline{m} \end{pmatrix}_m . \tag{17}$$

The reduced $2k$-dimensional eigenvalue equation for the basis vectors in Eqs. (7) and (10) can thus be set up based on knowledge of the vectors in Eqs. (7)–(9).

If we want to determine the jth excitation energy then the jth root ω^R_j of the reduced equation in Eq. (14) which satisfies positive normalization condition is an approximation to that eigenvalue and $\binom{{}^1\underline{X}^R_j}{{}^2\underline{X}^R_j}$ is the best approximation to the eigenvector within the basis described by Eqs. (7) and (10).

Convergence of the jth root is obtained when the norm of the residual

$$\underline{R} = \left[\begin{pmatrix} \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{B} & \underset{\sim}{A} \end{pmatrix} - \omega^R_j \begin{pmatrix} \underset{\sim}{\Sigma} & \underset{\sim}{\Delta} \\ -\underset{\sim}{\Delta} & -\underset{\sim}{\Sigma} \end{pmatrix} \right] \begin{pmatrix} {}^1\underline{X}^R_j \\ {}^2\underline{X}^R_j \end{pmatrix} \tag{18}$$

is smaller than a given tolerance.

A straightforward extension of the Davidson algorithm [14] to the more general eigenvalue equation in linear response gives the trial vector for the $(n+1)$th iteration as

$$\begin{pmatrix} {}^1\underline{b} \\ {}^2\underline{b} \end{pmatrix}_{k+1} = \begin{pmatrix} {}^1\underset{\sim}{d}_j & \underset{\sim}{0} \\ \underset{\sim}{0} & {}^2\underset{\sim}{d}_j \end{pmatrix}^{-1} \underline{R}, \tag{19}$$

where ${}^1\underset{\sim}{d}_j$ and ${}^2\underset{\sim}{d}_j$ are diagonal matrices containing the elements

$${}^1d_{j,ii} = A_{ii} - \omega^R_j \Sigma_{ii} \tag{20}$$

$${}^2d_{j,ii} = A_{ii} + \omega^R_j \Sigma_{ii}. \tag{21}$$

In Eq. (19) $\binom{{}^1\underline{X}^R_j}{{}^2\underline{X}^R_j}$ is considered to be a vector of the dimension of the LR eigenvalue equation in Eq. (7); that is, it is expanded in the original basis.

For convenience the vector $\binom{{}^1\underline{b}}{{}^2\underline{b}}_{k+1}$ is orthogonalized against the $2k$ previous trial vectors in Eqs. (7) and (10). The pair of new trial vectors, $\binom{{}^1\underline{b}}{{}^2\underline{b}}_{k+1}$ and $\binom{{}^2\underline{b}}{{}^1\underline{b}}_{k+1}$, are also orthogonalized against each other using an orthonormalization scheme which preserves the paired structure. We use symmetric orthonormalization. All trial vectors in any iteration are then orthonormal with respect to the usual Euclidean norm. For the LR eigenvalue equation we could have used vectors orthonormal with respect to the LR metric (Eq. (3)), but then we would only be able to reuse those trial vectors for the LR linear equations in Section II.b (e.g., for frequency dependent polarizability) with great difficulty—in fact, we would then have to reorthonormalize the trial vectors with respect to the Euclidean norm. The optimal choice for efficiency is thus to use the Euclidean norm in the first place.

Several roots can be determined simultaneously by applying Eq. (19) for as many roots as desired (as described by Liu for the symmetric eigenvalue equation [15]). The Σ_{ii} values are one for all operators in the CSF space. Small Σ_{ii} values may appear in the orbital space for operators which are mainly introduced to describe correlation effects.

In Eq. (1) where $\left(\begin{smallmatrix}\underset{\sim}{\Sigma} & \underset{\sim}{\Delta}\\ -\underset{\sim}{\Delta} & -\underset{\sim}{\Sigma}\end{smallmatrix}\right)$ is considered as metric, the negative and the smallest positive eigenvalues which satisfy a positive normalization condition (see Eq. (3)) are usually those of interest as they describe the lowest excitation energies. The LR Hessian matrix $\left(\begin{smallmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B} & \underset{\sim}{A}\end{smallmatrix}\right)$ may also be considered as metric [19] and the LR eigenvalue equation then reads

$$\left(\begin{pmatrix}\underset{\sim}{\Sigma} & \underset{\sim}{\Delta}\\ -\underset{\sim}{\Delta} & -\underset{\sim}{\Sigma}\end{pmatrix}-\frac{1}{\omega}\begin{pmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B} & \underset{\sim}{A}\end{pmatrix}\right)\begin{pmatrix}{}^{1}\underset{\sim}{X}\\ {}^{2}\underset{\sim}{X}\end{pmatrix}=0. \tag{22}$$

The eigenvalues of interest are now those of largest numerical value [19]. If we had based the algorithm on this equation instead of Eq. (1) we would have obtained the same sequence of trial vectors. We can thus use Eq. (22) to make statements about the convergence. For ground state calculations $\left(\begin{smallmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B} & \underset{\sim}{A}\end{smallmatrix}\right)$ is positive definite and the square of the eigenvalues may be obtained by solving a Hermitian eigenvalue problem (see, for example, Eq. (6.109) of Ref. [1]). McDonald's theorem then gives that the roots (all excitation energies are positive) of the reduced eigenvalue equation in Eq. (14) separate the roots of the eigenvalue equation in Eq. (7). A monotonic convergence towards the eigenvalues of Eq. (7) is therefore observed in each iteration when the dimension of the reduced space is increased.

In order to elucidate the implementation of this algorithm a step-by-step description is given in the Appendix.

II.b. *Linear Response Equations*

The set of linear equations which has to be solved to determine response properties for an external perturbation is

$$\left(\begin{pmatrix}\underset{\sim}{A} & \underset{\sim}{B}\\ \underset{\sim}{B} & \underset{\sim}{A}\end{pmatrix}-\omega\begin{pmatrix}\underset{\sim}{\Sigma} & \underset{\sim}{\Delta}\\ -\underset{\sim}{\Delta} & -\underset{\sim}{\Sigma}\end{pmatrix}\right)\begin{pmatrix}{}^{1}\underset{\sim}{Z}\\ {}^{2}\underset{\sim}{Z}\end{pmatrix}=\begin{pmatrix}\underset{\sim}{C}\\ \pm\underset{\sim}{C}\end{pmatrix}. \tag{23}$$

The matrices $\underset{\sim}{A}$, $\underset{\sim}{B}$, $\underset{\sim}{\Sigma}$, and $\underset{\sim}{\Delta}$ are the same as those appearing in the linear response eigenvalue equation in Eq. (1) and ω is the frequency of the external perturbation. The elements of the vector $\underset{\sim}{C}$ describe the nonoptimality (the gradient) of the wave function after the perturbation has been switched on. For an MCSCF wave function and a real perturbation $\underset{\sim}{C}$ is equal to the vector $\left(\begin{smallmatrix}\underset{\sim}{F}\\ \underset{\sim}{G}\end{smallmatrix}\right)$ in Eqs. (5.11)–(5.12) of Ref. [1]. The + and − signs in the second component are used when the perturbation operator is imaginary and real, respectively. The vector $\left(\begin{smallmatrix}{}^{1}\underset{\sim}{Z}\\ {}^{2}\underset{\sim}{Z}\end{smallmatrix}\right)$ describes the first-order correction to the wave function due to the external perturbation [4] and second-order molecular properties are determined as $(\tilde{\underset{\sim}{C}} \pm \tilde{\underset{\sim}{C}})\left(\begin{smallmatrix}{}^{1}\underset{\sim}{Z}\\ {}^{2}\underset{\sim}{Z}\end{smallmatrix}\right)$.

We now describe an iterative algorithm to solve Eq. (23) which finds the solution vector from a sequence of the linear transformations in Eqs. (5) and (6). We assume that in the nth iteration of the algorithm we have generated the set of k orthonormal vectors in Eq. (7) and that the linear transformations in Eqs. (5) and (6) have been carried out giving the vectors in Eqs. (8) and (9), respectively. As in the eigenvalue case the linear transformations corresponding to the paired vectors in Eq. (10) are also known. The trial vectors in Eqs. (7) and (10) must be orthonormal. The optimal linear combination of the basis vectors in Eqs. (7) and (10) is then determined from the reduced set of linear equations

$$\left[\begin{pmatrix} \utilde{A}^R & \utilde{B}^R \\ \utilde{B}^R & \utilde{A}^R \end{pmatrix} - \omega \begin{pmatrix} \utilde{\varSigma}^R & \utilde{\varDelta}^R \\ -\utilde{\varDelta}^R & -\utilde{\varSigma}^R \end{pmatrix}\right]\begin{pmatrix} {}^1\underline{Z}^R \\ {}^2\underline{Z}^R \end{pmatrix} = \begin{pmatrix} \underline{C}^R \\ \pm\underline{C}^R \end{pmatrix}, \tag{24}$$

where $\utilde{A}^R$, $\utilde{B}^R$, $\utilde{\varSigma}^R$, and $\utilde{\varDelta}^R$ are defined in Eqs. (14)–(17) and

$$C_l^R = ({}^1\tilde{\underline{b}} \quad {}^2\tilde{\underline{b}})_l \begin{pmatrix} \underline{C} \\ \pm\underline{C} \end{pmatrix}. \tag{25}$$

The vector $\binom{{}^1\underline{Z}^R}{{}^2\underline{Z}^R}$ is the optimal solution within the set of basis vectors in Eqs. (7) and (10).

The accuracy of a solution vector may be measured in terms of the norm of the residual vector

$$\underline{R} = \left[\begin{pmatrix} \utilde{A} & \utilde{B} \\ \utilde{B} & \utilde{A} \end{pmatrix} - \omega \begin{pmatrix} \utilde{\varSigma} & \utilde{\varDelta} \\ -\utilde{\varDelta} & -\utilde{\varSigma} \end{pmatrix}\right]\begin{pmatrix} {}^1\underline{Z}^R \\ {}^2\underline{Z}^R \end{pmatrix} - \begin{pmatrix} \underline{C} \\ \pm\underline{C} \end{pmatrix}. \tag{26}$$

Convergence is obtained when the norm of the residual is smaller than a specified tolerance. Trial vectors for the $(n+1)$th iteration may be generated using a generalization of the conjugate gradient algorithm [20]

$$\begin{pmatrix} {}^1\underline{b} \\ {}^2\underline{b} \end{pmatrix}_{k+1} = \begin{pmatrix} \utilde{f}^1 & \utilde{0} \\ \utilde{0} & \utilde{f}^2 \end{pmatrix}^{-1} \underline{R}, \tag{27}$$

where $\utilde{f}^1$ and $\utilde{f}^2$ are diagonal matrices with elements

$$f^1_{kk} = A_{kk} - \omega \varSigma_{kk} \tag{28}$$

$$f^2_{kk} = A_{kk} + \omega \varSigma_{kk}. \tag{29}$$

The vector $\binom{{}^1\underline{b}}{{}^2\underline{b}}_{k+1}$ is orthogonalized against the $2k$ trial vectors of the nth iteration in Eqs. (7) and (10) and the new pair of trial vectors of the $(n+1)$th iteration $\binom{{}^1\underline{b}}{{}^2\underline{b}}_{k+1}$ and $\binom{{}^2\underline{b}}{{}^1\underline{b}}_{k+1}$ are orthogonalized against each other using symmetric orthonormalization. The iterative procedure is continued until the norm of the residual is smaller than a specified tolerance. As starting vector we may use Eq. (27) with $\underline{R}$ equal to $\binom{\underline{C}}{\pm\underline{C}}$. When ω in Eq. (23) is close to a resonance it is advantageously first to locate the resonance from the linear response eigenvalue equation. The vectors which determine the resonance may then be used as starting vectors to solve the linear response equations. The resonance contribution to the solution is then contained in the initial basis of trial vectors.

The solution vector to Eq. (23) can be used to determine second-order molecular properties as $({}^1\underline{\tilde{Z}}\ {}^2\underline{\tilde{Z}})\begin{pmatrix}\underline{C}\\ \pm\underline{C}\end{pmatrix}$. In Ref. [21] it is shown that the accuracy in the second-order property is squared in the residual. If several components are calculated of a second-order property (for example, the different components of the polarizability), it is also shown in Ref. [21] that the out-of-diagonal components of second-order properties have an accuracy squared in the residual if the same reduced space is used to span all the solution vectors.

III. RESULTS

To illustrate the efficiency of the proposed algorithms we compare it to the convergence characteristics of a previously reported RPA calculation on ethylene by Bouman *et al.* [7], and we report a new MCLR calculation on the Neon atom.

The RPA calculation on ethylene was used by Bouman *et al.* [7] to illustrate the convergence characteristics of their algorithm. Details of geometry and basis set can be found there. In Table I we report the convergence characteristics of our new algorithm for finding the lowest 6 singlet states, and in Table II we report the corresponding characteristics for the algorithm of Bouman *et al.* derived from the results reported in Ref. [7]. The convergence in each iteration is measured as the

TABLE I

The Difference between the Eigenvalues of the Reduced RPA Equation and the Converged RPA Eigenvalues in a Sequence of Iterations for the Lowest 6 Singlet Excitations in Ethylene. Convergence to a Residual Tolerance 10^{-3}.

Root number[a] / Iteration number[b]	1	2	3	4	5	6
1(12)	0.0039027	0.0064854	0.0168965	0.0137942	0.0142670	0.0025675
2(18)	0.0000884	0.0014301	0.0029791	0.0009608	0.0003791	0.0000500
3(24)	0.0000008	0.0001017	0.0000006	0.0000265	0.0000070	0.0000005
4(30)	0.0000000	0.0000013	0.0000000	0.0000023	0.0000002	0.0000000
5(32)		0.0000001		0.0000002		

[a] The fully converged excitation energies for the lowest 7 states are in a.u. 0.2626717, 0.2830822, 0.2844209, 0.2898198, 0.3009313, 0.3224419, and 0.3233291, respectively.

[b] The numbers in parentheses denote the number of linear transformations which is equal to half the dimension of the reduced space RPA matrices.

TABLE II

The Difference Calculated by Bouman *et al.* [7] between the Eigenvalues of the Reduced RPA Equation and the Converged RPA Eigenvalues in a Sequence of Iterations for the Lowest 6 Singlet Excitations in Ethylene

Root number / Iteration number[a]	1	2	3	4	5	6
1(12)	0.0052563	0.0083572	0.0182450	0.0139583	0.0092640	0.0015487
2(18)	0.0004751	0.0015831	0.0057538	0.0016486	0.0008344	0.0000712
3(24)	0.0000114	0.00012557	0.0000526	0.0001223	-0.0000219	0.0000039
4(30)	0.0000053	0.0000768	0.0000187	0.0000104	0.0000042	$<10^{-6}$
5(35)	$<10^{-6}$	0.0001431	0.0000098	0.0000021	0.0000025	
6(39)		-0.0000949	-0.0000002	$<10^{-6}$	$<10^{-6}$	
7(41)		0.0000035	$<10^{-6}$			
8(42)		$<10^{-6}$				

[a] The numbers in parentheses denote the number of trial vectors which have been used.

difference between the current approximate excitation energy (from the reduced space) and the converged excitation energy. We converge to a residual tolerance 10^{-3}, which leads to an accuracy of approximately 10^{-6} in the excitation energies (as in the case of a symmetric eigenvalue problem as standard configuration interaction (CI)). Bouman *et al.* converge to a residual tolerance of 10^{-4}; however, their residual is defined differently from ours and their tolerance also leads to approximately 10^{-6} error in excitation energies. We note that our calculation converged using a total of 32 trial vectors while Bouman *et* al. used 42 trial vectors. More important for this rather small example, we observe from Table I that we obtain a monotonic decrease in the excitation energies towards the totally converged results, while in the calculation by Bouman *et al.* in Table II such a monotonic decrease is not always obtained. For example, the error in root 2 increases from 0.000,076,8 in iteration 4 to 0.000,143,1 in iteration 5, and in iteration 6 the approximate excitation energy is 0.000,094,9 *below* the totally converged excitation energy.

A RPA calculation is usually much simpler than a MCLR calculation because in the RPA calculation both the matrices $\begin{pmatrix} \underline{A} & \underline{B} \\ \underline{B} & \underline{A} \end{pmatrix}$ and $\begin{pmatrix} \underline{\Sigma} & \underline{\Delta} \\ -\underline{\Delta} & -\underline{\Sigma} \end{pmatrix}$ are diagonally dominant. In an MCLR calculation parts of these matrices have no diagonal dominance. In Tables III and IV we report convergence charististics of a MCLR calculation on the NE atom with 65 GTO's [22]. All single double, triple, and quadruple excitations out of the $2s$ and $2p$ orbitals into the $3s$ and $3p$ orbitals are included in the configuration space. In Table III the convergence characteristic is given for a calculation of the lowest excitations of $X^1S \rightarrow {}^1S$ and $X^1S \rightarrow {}^1D$ symmetry. The calculations are carried out in the D_{2h} subgroup and two components are therefore

TABLE III

The Difference between the Eigenvalues of the Reduced MCLR Equation and the Exact MCLR Eigenvalues in a Sequence of Iterations for the Two Excitations of $X^1S \rightarrow {}^1S$ and $X^1S \rightarrow {}^1D$ Symmetry in the Ne Atom. Convergence to a Residual Tolerance 10^{-3}

Iteration[b] number \ Root number[a]	1	2	3	4	5	6
1(13)	0.1329207(2)[c]	0.1329207(3)	0.1319480(1)	0.1859297(5)	0.1859297(6)	0.1817246(4)
2(19)	0.0194656(1)	0.0194661(2)	0.0160613(3)	0.0404303(4)	0.0404309(5)	0.0549739(6)
3(25)	0.0098700(1)	0.0098714(2)	0.0110460(3)	0.0240211(4)	0.0240221(5)	0.0413698(6)
4(31)	0.0001223(1)	0.0001225(2)	0.0070920(3)	0.0007897(4)	0.0007907(5)	0.0233305(6)
5(37)	0.0000111(1)	0.0000111(2)	0.0001210(3)	0.0000272(4)	0.0000272(5)	0.0004029(6)
6(43)	0.0000070(1)	0.0000070(2)	0.0000458(3)	0.0000030(4)	0.0000030(5)	0.0001180(6)
7(47)			0.0000261(3)	0.0000019(4)	0.0000019(5)	0.0000323(6)
8(49)			0.0000218(3)			0.0000082(6)
9(50)						0.0000065(6)

[a] (1, 2) and (4, 5) are two components of the lowest two $X^1S \rightarrow {}^1D$ excitations with excitation energy 0.6961506 a.u. and 0.755463 a.u., respectively; 3 and 6 are the lowest two $X^1S \rightarrow {}^{1S}S$ excitations with excitation energy 0.7087031 a.u. and 0.7593817 a.u., respectively.

[b] The numbers in parentheses denote the dimension of the reduced space MCLR matrices.

[c] The numbers in parentheses denote the order of the eigenvalues in the reduced space.

included for each ${}^1D$ state. The dimension of the matrices $(\underset{\sim}{A}, \underset{\sim}{B}, \underset{\sim}{\Sigma}, \underset{\sim}{\Delta})$ is 291 of which 105 is of orbital type. The convergence is, as expected, slower than in the ethylene RPA calculation. Convergence to a residual tolerance of 10^{-3} is obtained in 6–9 iterations. In the first iteration in Table III the order of the lowest ${}^1D$ and ${}^1S$ excitation energies is switched compared to the order in the converged calculation. The convergence in Table III is typically linear convergence and, as expected, monotonic towards the totally converged roots. In Table IV the convergence characteristic is given for a calculation on the lowest excitations of $X^1S \rightarrow {}^1P$, $X^1S \rightarrow {}^1F$, and $X^1S \rightarrow {}^1H$ symmetry. The dimension of the MCLR matrices is 225 of which 79 is of orbital type. In the first three iterations in Table IV an incorrect order is obtained in the excitation energies. After iteration 3 the excitations of ${}^1F$ and ${}^1H$ symmetry have converged mainly due to small dimensions in these spaces and in the final six iterations all new trial vectors are of ${}^1P$ symmetry. At iteration 4 a hitherto missing root shows up (the error in the excitation energy drops from 0.11811145 to 0.0056951) and typical linear convergence is obtained from that iteration onwards.

To demonstrate the efficiency of the iterative algorithm for solving the MC linear response equations we report in Table V the convergence characteristics of a frequency independent $(\omega = 0)$ polarizability calculation on Ne. The initial trial vector was the property vector $\underset{\sim}{C}$ divided by the diagonal LR matrix elements as described

TABLE IV

The Difference between the Eigenvalues of the Reduced MCLR Equation and the Exact MCLR Eigenvalues in a Sequence of Iterations for the Lowest Excitations of $X^1S \rightarrow {}^1P$ and $X^1S \rightarrow F$ Symmetry in the Ne Atom. Convergence to a Residual Tolerance 10^{-3}

Root number[a] Iteration number[b]	1	2	3	4	5	6	7
1(12)	0.1793625(1)[c]	0.1976572(2)	0.2163368(6)	0.0938972(3)	0.0938972(4)	0.7612798(7)	0.0935171(5)
2(18)	0.0476997(1)	0.0422119(2)	0.1188079(5)	0.0000953(3)	0.0000955(4)	0.4805246(7)	0.0000711(6)
3(24)	0.0308931(1)	0.0272284(2)	0.1181145(5)	0.0000032(3)	0.0000032(4)	0.0686924(7)	0.0000010(6)
4(27)	0.0042317(1)	0.0113135(2)	0.0056951(3)	(4)	(5)	0.0000311(6)	(7)
5(31)	0.0001901(1)	0.0007622(2)	0.0002447(3)	(4)	(5)	0.0000087(6)	(7)
6(35)	0.0000241(1)	0.0000448(2)	0.0000187(3)	(4)	(5)	0.0000012(6)	(7)
7(39)	0.0000158(1)	0.0000039(2)	0.0000052(3)	(4)	(5)	0.0000003(6)	(7)
8(40)	(1)	0.0000036(2)	(3)	(4)	(5)	(6)	(7)

[a] 1, 2, 3, and 6 are the four excitation $X^1S \rightarrow {}^1P^0$ with excitation energies 0.6316345 a.u., 0.7374811 a.u., 0.7759858 a.u., and 0.8939950 a.u., respectively; 4, 5 are two components of the lowest $X^1S \rightarrow {}^1F^0$ excitation with excitation energy 0.8888564 a.u., and 7 is the lowest $X^1S \rightarrow {}^1H^0$ excitation with excitation energy 0.8947578 a.u.

[b] The number in parentheses denotes the dimension of the reduced space MCLR matrices.

[c] The number in parentheses denotes the order of the eigenvalues in the reduced space.

TABLE V

Convergence Characteristics when Solving the Linear Equations for the Ne Atom at the Frequencies $\omega = 0.0$ a.u. and $\omega = 0.1$ a.u.

Iteration number	$\omega = 0.0$		$\omega = 0.1$		$\omega = 0.1$[a]	
	Polarizability	Residual	Polarizability	Residual	Polarizability	Residual
1	2.1660116	1.2551	2.1796743	1.2591	2.6447380	0.0614
2	2.4394803	0.8155	2.4583579	0.8314	2.6457370	0.0231
3	2.5663182	0.4149	2.5898047	0.4226	2.6457862	0.0032
4	2.5898329	0.2403	2.6140995	0.2443	2.6457878	0.0009
5	2.6089461	0.2371	2.6347751	0.2447		
6	2.6169385	0.0732	2.6435753	0.0801		
7	2.6184280	0.0191	2.6456083	0.0248		
8	2.6185188	0.0061	2.6457734	0.0074		
9	2.6185263	0.0024	2.6457853	0.0029		
10	2.6185278	0.0014	2.6457874	0.0013		
11	2.6185280	0.0004	2.6457878	0.0004		

[a] The 11 trial vectors from the $\omega = 0.0$ a.u. are used as initial trial vectors.

in Eqs. (27)–(29). The calculation converged to a residual tolerance 10^{-3} in 11 iterations. The convergence characteristics of a calculation at frequency $\omega = 0.1$ a.u. using the same initial trial vector as in the $\omega = 0$ calculation is also given in Table V. The convergence characteristics of the $\omega = 0$ and $\omega = 0.1$ calculation are very similar and convergence is obtained in the same number of iterations. In Table V we also report the convergence characteristics of a $\omega = 0.1$ calculation where we have included the 11 trial vectors which were determined in the $\omega = 0$ calculation, as initial trial vectors. Only four additional iterations are then required to obtain convergence to the threshold 10^{-3}. It is thus, as expected, important with respect to computational efficiency to reuse the trial vectors when solving the linear response equations at a new close-lying frequency.

When the frequency ω in Eq. (23) is close to an excitation energy (a resonance) the matrix $\begin{pmatrix} \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{B} & \underset{\sim}{A} \end{pmatrix} - \omega \begin{pmatrix} \underset{\sim}{\Sigma} & \underset{\sim}{\Delta} \\ -\underset{\sim}{\Delta} & \underset{\sim}{\Sigma} \end{pmatrix}$ is nearly singular and straightforward application of the generalized conjugate gradient algorithm will give slow convergence. To illustrate this point we report in Table VI the convergence characteristics obtained in a calculation with $\omega = 0.6320$ a.u. (the excitation energy is 0.6316). The calculation converges in 36 iterations which are significantly more iterations than the 11 iterations required in the calculations at the frequencies 0.0 a.u. and in 0.1 a.u. (compare to Table V). In Table VI we also report the convergence characteristics which are obtained if we include the 12 vectors which determine the excitation energy to a residual tolerance 10^{-3} as initial trial vector. Table VI shows that the use of the trial vectors from an excitation energy calculation reduces significantly the number of iterations which is required to obtain convergence. The excitation energy trial vectors eliminate the near singularity contribution in the solution vector. In Table VI we also report a calculation where we, in addition to the excitation energy trial vectors, have included the trial vectors of the frequency independent calculation. These last trial vectors are seen to have basically no influence on the convergence in this case.

IV. DISCUSSION

A major advantage of our proposed eigenvalue algorithm is that although the LR eigenvalue problem is formally a generalized eigenvalue problem we can use the specific paired structure to obtain an algorithm which is just as stable and efficient as the Davidson–Liu [14, 15] algorithm for the normal symmetric eigenvalue problem. The algorithm and the analogous iterative algorithm for solving LR sets of linear equations are both solely based on the two linear transformations in Eqs. (5) and (6). As we envision LR matrices which may have a dimension of 10^5–10^6 and larger this is an important part of the algorithms. For MCSCF wave functions we have shown how the two linear transformations can be calculated directly without explicit construction of the two matrices [4] and this is what makes it feasible to perform large-scale, accurate LR calculations. The linear trans-

TABLE VI

Convergence Characteristics when Solving the Linear Equations for the Ne Atom at the Frequency $\omega = 0.6320$ a.u.

Iteration number	Polarizability	Residual	Polarizability[a]	Residual[a]	Polarizability[b]	Residual[b]
1	3.032813	1.5539	−372.657935	1.5335	−372.188383	1.5784
2	4.020402	2.1092	−371.513264	0.5436	−371.464683	0.5213
3	4.986168	1.3434	−371.393860	0.3075	−371.381807	0.2846
4	5.857768	1.6223	−371.366185	0.2080	−371.366475	0.2770
5	8.302796	2.6672	−371.342024	0.2294	−371.309785	0.4882
6	10.610454	1.6946	−371.281485	0.6740	−370.756720	4.3490
7	12.383247	2.0933	−370.702647	1.3975	−365.817129	4.1263
8	21.809483	7.1883	−364.751608	6.1279	−356.669646	7.0414
9	76.294185	14.1546	−356.970923	2.0847	−355.792362	0.6765
10	400.450290	48.8363	−355.827357	0.9557	−355.761008	0.2362
11	−542.921660	31.5497	−355.763852	0.2132	−355.759675	0.0737
12	−388.254881	6.3980	−355.761050	0.0355	−355.759595	0.0170
13	−373.090253	1.8456	−355.760911	0.0136	−355.759452	0.0334
14	−371.486757	0.7968	−355.760864	0.0137	−355.749894	0.5955
15	−371.313014	0.2628	−355.760277	0.1079	−355.702040	0.2911
16	−371.277449	0.1989	−355.751667	0.1870	−355.692364	0.1016
17	−371.254621	0.2256	−355.699394	0.5915	−355.689809	0.1856

18	−371.162879	0.5684	−355.688409	0.1061	−355.685195	0.0798
19	−370.598225	1.3445	−355.685121	0.0396	−355.684856	0.0108
20	−367.397972	3.0772	−355.684862	0.0117	−355.684845	0.0027
21	−360.966370	2.3886	−355.684845	0.0032	−355.684844	0.0005
22	−357.193464	1.8356	−355.684843	0.0008		
23	−355.838456	0.5578				
24	−355.765930	0.1423				
25	−355.762272	0.0311				
26	−355.762131	0.0094				
27	−355.762080	0.0189				
28	−355.761425	0.0505				
29	−355.752155	0.2829				
30	−355.702275	0.2692				
31	−355.690311	0.0766				
32	−355.686472	0.0565				
33	−355.684922	0.0208				
34	−355.684850	0.0042				
35	−355.684844	0.0021				
36	−355.684843	0.0004				

[a] The 12 trial vectors required to converge the lowest MCLR eigenvalue of $X^1S \rightarrow {}^1P$ symmetry to a residual tolerance 10^{-3} are used as initial trial vectors.

[b] As in [a] but with the additional 6 vectors that were required to get the polarizability at $\omega = 0.0$ a.u.

formations will, by, far be the most CPU and IO demanding part of the calculation, and the primary emphasis in the algorithm design has therefore been to minimize the number of iterations (the number of linear transformations) rather than CPU and IO in the eigenvalue algorithm. Therefore, if possible within the external constraints (memory, disk space), we prefer to keep all trial vectors in the reduced space. In cases where vectors with no or very little importance can be isolated these can be discarded without effecting the total convergence. In less fortunate cases the reduction of the number of trialvectors can result in extra iterations.

The stable and efficient convergence has been obtained by always adding pairs of trial vectors to the reduced space such that the reduced space matrices maintain the paired structure of the full matrices. This is important as the roots of the reduced eigenvalue equation will then monotonically converge to the roots of the linear response eigenvalue equation. Recall that it does not require additional linear transformations to include the paired trial vectors. The paired structure also guarantees that complex roots cannot be obtained in the reduced space. In previously used iterative algorithms complex roots have occasionally shown up.

When solving the linear response equations for frequency dependent properties we have demonstrated the economy of reusing trial vectors from one frequency at new frequencies. We have also shown that for frequencies near resonances the near singularity contribution can be removed by including the trial vectors of the resonance eigenvalue among the initial trial vectors. We have thus developed a very efficient and numerically stable algorithm for finding the solution vectors of linear response eigenvalue and linear equations directly, without explicit construction of the large matrices. We believe this algorithm will be very important for the future development of large scale linear and nonlinear response calculations.

APPENDIX: THE LR EIGENVALUE ALGORITHM

To elucidate implementation this Appendix contains a detailed step-by-step description of the linear response eigenvalue algorithm which was verbally described and discussed in Section III.a. We use here "#" to denote the paired vector

$$\text{if } \underline{X}_j = \begin{pmatrix} {}^1\underline{X}_j \\ {}^2\underline{X}_j \end{pmatrix} \qquad \text{then } \underline{X}_j^{\#} = \begin{pmatrix} {}^2\underline{X}_j \\ {}^1\underline{X}_j \end{pmatrix}.$$

The other matrices and vectors used below are defined in Section II.a. A step-by-step description of the LR property algorithm in Section II.b is obtained by reading "set of linear equations" instead of "eigenvalue equation" below (K may be greater than one if several properties are solved for).

PROBLEM. Determine the K first excitation energies to a residual tolerance of t (e.g., $t = 10^{-3}$).

Initialization.

(a) Select $k \geqslant K$.

(b) Select a set $U_k = \{\underline{b}_1, \underline{b}_1^{\#}, ..., \underline{b}_k, \underline{b}_k^{\#}\}$ of initial trial vectors.

(c) Set $p = 0$.

Iteration $(n = 0, 1, 2, ...)$.

(a) Calculate $\{\underline{u}_i = \begin{pmatrix} \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{B} & \underset{\sim}{A} \end{pmatrix} \underline{b}_i, \underline{m}_i = \begin{pmatrix} \underset{\sim}{\Sigma} & \underset{\sim}{\Delta} \\ -\underset{\sim}{\Delta} & -\underset{\sim}{\Sigma} \end{pmatrix} \underline{b}_i; i = p+1, p+k\}$.

(b) Save $\{\underline{b}_i, \underline{u}_i, \underline{m}_i; i = p+1, p+k\}$ in memory or on external storage (for steps (c) and (g)).

(c) Extend the reduced eigenvalue equation with $\{\underline{b}_i, \underline{b}_i^{\#}; i = p+1, p+k\}$.

(d) Set $p = p + k$.

(e) Solve reduced eigenvalue equation of dimension $2p$.

(f) Select $k \geqslant K$.

(g) Calculate the residual vectors $\underline{R}_j$, $j = 1, k$ for nonconverged roots.

(h) Test for convergence: $\|\underline{R}_j\| \leqslant t$. Decrement k with one for each converged root. If the K lowest roots are converged then exit, otherwise continue to (i) with the k $\underline{R}_j$-vectors associated with nonconverged roots.

(i) Calculate $\underline{b}_{p+j} = \underline{D}_j^{-1} \underline{R}_j$, $j = 1, k$.

(j) Gram–Schmidt orthonormalize these new trial vector pairs to previous trial vectors U_p and symmetrically orthonormalize them to each other. Omit any linear dependent $\{\underline{b}_j, \underline{b}_j^{\#}\}$ pairs and decrement k by one for each linear dependent pair.

(k) If $k = 0$ then error exit; otherwise go back to (a) for next iteration with $n = n + 1$.

If desired because space requirements dictate so, or because the I/O involved becomes significant, it is possible to restart as in the Davidson algorithm [14] or to only keep a small number of vectors in the reduced basis as in Ref. [18].

ACKNOWLEDGMENT

H. J. Aa. Jensen acknowledges support under Grant 11-5420 from the Danish Natural Science Research Council.

REFERENCES

1. P. JØRGENSEN AND J. SIMONS, *Second Quantization Based Methods in Quantum Chemistry* (Academic Press, New York, 1981).
2. D. L. YEAGER AND P. JØRGENSEN, *Chem. Phys. Lett.* **65**, 77 (1979).
3. E. DALGAARD, *J. Chem. Phys.* **72**, 816 (1980).

4. J. Olsen and P. Jørgensen, *J. Chem. Phys.* **82**, 3225 (1985).
5. H. J. Monkhorst, *Int. J. Quantum Chem.* **S11**, 421 (1977).
6. J. Oddershede, P. Jørgensen, and D. L. Yeager, *Comput. Phys. Rep.* **2**, 35 (1984).
7. T. Bouman, Aa. E. Hansen, B. Voigt, and S. Rettrup, *Int. J. Quantum Chem.* **23**, 595 (1983).
8. S. Rettrup, *J. Comput. Phys.* **45**, 100 (1982).
9. K. Hirao and H. Nakatsuji, *J. Comput. Phys.* **45**, 246 (1982).
10. J. P. Flament and H. G. Gervais, *Int. J. Quantum Chem.* **16**, 1347 (1979).
11. R. K. Nesbet, *J. Chem. Phys.* **43**, 311 (1965).
12. I. Shavitt, *J. Comput. Phys.* **6**, 124 (1970); I. Shavitt, C. F. Bender, A. Pepano, and R. P. Hesteney, *J. Comput. Phys.* **11**, 90 (1973).
13. C. Lanczos, *J. Res. Nat. Bur. Stand.* **45**, 255 (1950).
14. E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975); E. R. Davidson, *J. Phys. A* 13, L179 (1980).
15. B. Liu, "The Simultaneous Expansion Method," Proceedings, Numerical Algorithms in Chemistry: Algebraic Methods, Workshop of the National Resource for Computation in Chemistry, Berkeley, 1978, p. 49.
16. R. C. Raffenetti, *J. Comput. Phys.* **32**, 403 (1979).
17. G. A. Gallup, *J. Comput. Chem.* **3**, 127 (1982).
18. N. Kosugi, *J. Comput. Phys.* **55**, 426 (1984).
19. J. Linderberg and Y. Ohrn, *Int. J. Quantum Chem.* **12**, 161 (1977).
20. M. Hestenes, *Conjugate Direction Methods in Optimization* (Springer, Berlin, 1980).
21. T. U. Helgaker, J. Almlöf, H. J. Aa. Jensen, and P. Jørgensen, *J. Chem. Phys.* **84**, 6266 (1986).
22. The 14s 9p 4d GTO basis was obtained from the 11s 6p 3d basis of Werner and Meyer (*Mol. Phys.* **31**, 8551 (1976)) by adding three *s*-functions (exponents 0.05, 0.015, 0.004), three *p*-sets (exponents 0.05, 0.015, 0.004) and a single *d*-set (exponent 0.11). The *s* components of the *d*-sets were excluded.