An Improvement of Davidson's Iteration Method: Applications to MRCI and MRCEPA Calculations

H. J. J. VAN DAM* and J. H. VAN LENTHE

Theoretical Chemistry Group, Debye Institute, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands

G. L. G. SLEIJPEN and H. A. VAN DER VORST

Department of Mathematics, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands

ABSTRACT

Davidson's method is widely used for finding the lowest eigenvalues of large matrices. Recently, mathematicians have shown that Davidson's derivation could be improved. They have corrected the derivation yielding a new iteration method. In this article this new method is adapted for realistic MRCI and MRCEPA calculations. Results show that the new method converges significantly faster in H_2O and O_2 with moderately elongated bonds than Davidson's original method. The new method offers new insights into the rate of convergence of Davidson's original method. © 1996 by John Wiley & Sons, Inc.

Introduction

Configuration interaction (CI) methods are well established and are widely applied. In CI calculations one aims at the few lowest eigenvalues and the corresponding eigenvectors of CI Hamilton matrices. Because these matrices are large and sparse, iterative methods are the methods of choice. These methods require a matrix–vector multiplication with the CI matrix in every iteration. Because the CI matrices are so large, these matrix–vector multiplications determine the cost of the algorithm. Therefore, almost every modification that reduces the total number of these matrix–vector multiplications can be applied to reduce the total cost.

The eigenvalue problem solver which is most often used in CI calculations was proposed by Davidson.[1-3] In this method the eigenvectors are expanded in a subspace and the subspace is extended with an update vector in every iteration. Recently, mathematicians have studied Davidson's method. Saad[4] pointed out that Davidson's deduc-

*Author to whom all correspondence should be addressed.

tion of the update vector did not lead to an optimal choice. Sleijpen et al.[5] have corrected the deduction. Their results lead to extended insight in the convergence properties of Davidson's method and suggest improvements.

In this article we summarize the discussion that resulted in the corrected derivation. In addition, we describe the application of the improved method in multireference CI (MRCI) calculations. The results are compared with the results obtained with Davidson's original method. A related method has been proposed in J. Olsen, P. Jørgensen, and J. Simons, *Chem. Phys. Lett.*, **169**, 463 (1990).

Davidson's Update Vector

In his original article, Davidson[1] started from the observation that the lowest eigenvalue of a matrix is at the minimum of the corresponding Rayleigh quotient:

$$R(|\mathbf{x}\rangle) = \frac{\langle \mathbf{x}|\mathbf{A}|\mathbf{x}\rangle}{\langle \mathbf{x}|\mathbf{x}\rangle} \quad (1)$$

Therefore, if the desired eigenvector is approximated by a vector $|\mathbf{u}\rangle$, one is interested in an update vector $|\boldsymbol{\delta}\rangle$ that minimizes $R(|\mathbf{u}\rangle + |\boldsymbol{\delta}\rangle)$. To this end, Davidson expanded the Rayleigh quotient to second order in the update vector and minimized the expression yielding the linear system

$$(\mathbf{A} - \lambda\mathbf{I})|\boldsymbol{\delta}\rangle = -|\mathbf{r}\rangle \quad (2)$$

where $|\mathbf{r}\rangle$ is the residual vector

$$|\mathbf{r}\rangle = (\mathbf{A} - \lambda\mathbf{I})|\mathbf{u}\rangle$$

and $\lambda = R(|\mathbf{u}\rangle)$. If the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is diagonally dominant, that is[6]

$$|(\mathbf{A}_{jj} - \lambda)| > \sum_{k \neq j} |\mathbf{A}_{jk}| \quad (3)$$

then eq. (2) may be approximated by

$$(\mathbf{D} - \lambda\mathbf{I})|\boldsymbol{\delta}\rangle = -|\mathbf{r}\rangle \quad (4)$$

Equation (4) is the famous Davidson preconditioning. The fact that this method results from a second-order energy expression should explain the good convergence properties compared to gradient methods.

However, Saad[4] pointed out that Davidson's derivation is not complete. Following Davidson's line of reasoning, the best results would be ob-

tained if one solves eq. (2) for $|\boldsymbol{\delta}\rangle$. This would result in $|\boldsymbol{\delta}\rangle = -|\mathbf{u}\rangle$. However, after orthogonalization of $|\boldsymbol{\delta}\rangle$ on the subspace spanned by update vectors from earlier iterations, the null vector remains. Clearly, this vector cannot serve as an effective extension of the subspace. Consequently, the update vector $|\boldsymbol{\delta}\rangle$ should be linearly independent to the subspace whereas $|\mathbf{u}\rangle$ is a linear combination in that subspace. This suggests that Davidson's argument is not complete. In a recent article, Sleijpen et al.[5] pointed out that while deriving eq. (2), Davidson assumed $|\boldsymbol{\delta}\rangle$ orthogonal to $|\mathbf{u}\rangle$. However, to assure that $|\boldsymbol{\delta}\rangle$ is an effective update vector, one should determine $|\boldsymbol{\delta}\rangle$ under the condition that $|\boldsymbol{\delta}\rangle$ is orthogonal to $|\mathbf{u}\rangle$. This may be effected by computing the update vector from the projection of $\mathbf{A}$ onto the subspace orthogonal to $|\mathbf{u}\rangle$. Sleijpen et al. showed that this is equivalent to substituting the projected matrix

$$\mathbf{B} = (\mathbf{I} - |\mathbf{u}\rangle\langle \mathbf{u}|)\mathbf{A}(\mathbf{I} - |\mathbf{u}\rangle\langle \mathbf{u}|) \quad (5)$$

for $\mathbf{A}$ in eq. (2), yielding

$$(\mathbf{B} - \lambda\mathbf{I})|\boldsymbol{\delta}\rangle = -|\mathbf{r}\rangle \quad (6)$$

If the matrix $(\mathbf{B} - \lambda\mathbf{I})$ is diagonally dominant, eq. (6) may be approximated to give

$$(\mathbf{D_B} - \lambda\mathbf{I})|\boldsymbol{\delta}\rangle = -|\mathbf{r}\rangle \quad (7)$$

where $\mathbf{D_B}$ is the diagonal of $\mathbf{B}$. This equation is analogous to eq. (4). Moreover, if the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is strongly diagonally dominant, then a unit vector is a good approximation to the eigenvector and to $|\mathbf{u}\rangle$. In that case, eq. (7) will only differ from eq. (4) in the first element. The difference in the resulting update vectors nearly vanishes upon orthogonalizing the update vector on the subspace. Therefore, both eqs. (4) and (7) may then be considered equivalent. However, eqs. (2) and (6) are certainly not equivalent. Due to Saad, we know that a calculation based on eq. (2) will not converge. To investigate the convergence of calculations with eq. (6), a calculation on H_2O was performed. The results are shown in Table I (computational details are provided later in this article). Note that the convergence of the energy is at least second order. That is, the number of converged digits is doubled in every it eration. This is in accordance with the fact that eq. (6) resulted from a second-order energy expression.

For practical applications, we will concentrate on eq. (6). Approaches based on this equation will be referred to by the name generalized Jacobi-Davidson[5] (GJD). When Davidson's method con-

TABLE I.
The Convergence of the Generalized Jacobi-Davidson Method in Extreme for H_2O at $2.0 \cdot R_e$.

Iteration Number	Energy (Hartree)
1	−75.839984920556
2	−75.950820825635
3	−75.951032652229
4	−75.951032652292

verges slowly due to large off-diagonal elements, eq. (6) may allow approaches with better convergence properties. However, in its current formulation, the equation is rather impractical because of the size of the matrix $(B - \lambda I)$. Therefore, some suitable approximations to this matrix must be found.

Application of the Generalized Jacobi-Davidson Method to MRCI Problems

In MRCI we distinguish two categories of molecular orbitals (MOs): internal orbitals and external orbitals.[7] The set of internal orbitals contains the MOs which are of primary importance in the construction of a qualitatively correct wave function. From the MOs, two types of n-electron states may be constructed: vacuum states and external states. The vacuum states have no electrons occupying external orbitals. All other states are external states. Based on these definitions, CI vectors can be separated into a vacuum part and an external part. We will denote a vacuum part with a subscript v and an external part with a subscript e.

In MRCI calculations with a suitably chosen vacuum space, the vacuum coefficients are much more important than the external coefficients. This suggests that if the projector P is defined as

$$|u\rangle\langle u| = P = \begin{bmatrix} P_{vv} & P_{ev} \\ P_{ve} & P_{ee} \end{bmatrix} \tag{8}$$

this may be approximated by

$$P \approx \begin{bmatrix} P_{vv} & 0 \\ 0 & 0 \end{bmatrix}. \tag{9}$$

Introducing this approximation into the matrix $(B - \lambda I)$ from eq. (6) yields

$$(B - \lambda I)$$
$$= \begin{bmatrix} (I - P)_{vv} A_{vv} (I - P)_{vv} & (I - P)_{vv} A_{ve} \\ A_{ev} (I - P)_{vv} & A_{ee} \end{bmatrix}$$
$$- \lambda \begin{bmatrix} I_{vv} & 0 \\ 0 & I_{ee} \end{bmatrix}$$

Furthermore, we assume that $(B - \lambda I)$ is diagonally dominant in the external space. This means we assume that we may approximate the matrix by

$$(B - \lambda I)$$
$$\approx \begin{bmatrix} (I - P)_{vv} A_{vv} (I - P)_{vv} - \lambda I_{vv} & 0 \\ 0 & D_{ee} - \lambda I_{ee} \end{bmatrix} \tag{10}$$

where D_{ee} is the diagonal of the external−external block. Substituting eq. (10) into eq. (6) results in

$$(B - \lambda I)_{vv} |\delta\rangle_v = -|r\rangle_v \tag{11}$$
$$(D - \lambda I)_{ee} |\delta\rangle_e = -|r\rangle_e \tag{12}$$

Because the linear system in the vacuum space is relatively small, we expect that these equations may be cost effectively applied to realistic MRCI problems. Note that in Table CI[8] calculations, similar approximations can be used based on a selection of the most important configurations. Furthermore, eqs. (11) and (12) may also be applied to MRCEPA[9] calculations because MRCEPA may be expressed in a form analogous to MRCI. That is, the MRCEPA matrix is equivalent to a MRCI matrix with shifted diagonal elements.

The linear system in eq. (11) can be treated with Gaussian elimination if the matrix is small. However, if the vacuum space is large (but small compared to the total CI space), an iterative method such as conjugate gradient[10, 11] may be more efficient. In the conjugate gradient method, a matrix−vector multiplication with the vacuum part of the matrix $(B - \lambda I)$ is performed in every iteration. These multiplications may be computed efficiently as follows. Assume the vacuum matrix $(B - \lambda I)$ is to be multiplied with a vector $|q\rangle$. Then, using eq. (5), we have

$$(B - \lambda I)|q\rangle = A|q\rangle - |u\rangle\langle u|A|q\rangle - A|u\rangle\langle u|q\rangle$$
$$+ |u\rangle\langle u|A|u\rangle\langle u|q\rangle - |q\rangle\lambda \tag{13}$$

Writing

$$\lambda = \langle u|A|u\rangle$$

defining

$$|z\rangle = A|u\rangle$$

and using that A is hermitian, eq. (13) becomes

$$(B - \lambda I)|q\rangle = A|q\rangle + |u\rangle(\langle u|q\rangle\lambda - \langle z|q\rangle)$$
$$- |z\rangle\langle u|q\rangle - |q\rangle\lambda \quad (14)$$

Equation (14) involves a matrix–vector product with the original vacuum matrix A and some simple vector–vector operations. The vector $|z\rangle$ can be obtained without additional cost because it is already calculated in the Davidson algorithm in the computation of the residual vector.

Symmetric Dissociation of H_2O

The first test case we consider involves the cleavage of both OH bonds at a constant angle in the water molecule. The geometries and the basis sets of Bauschlicher and Taylor[12] are used. This means that the energy of the molecule was calculated at the equilibrium OH bond length R_e, at $1.5 \cdot R_e$ and at $2.0 \cdot R_e$, where R_e was 1.889 726 Bohr. Additionally, we also calculate the energy at 4.0, 6.0, and 8.0 times R_E. The HOH angle is fixed at 104.5°. The $1s$ orbital is frozen on the self-consistent field (SCF) level. The active space in the Multi Configuration SCF (MCSCF) calculation also defines the reference space in the MRSDCI and MRCEPA calculations. The results in Tables II and III are obtained with the 55 Configuration State Function (CSF) Complete Active Space (CAS) from Ruttink et al.[9] The initial vector is computed by solving the reference part of the CI matrix for the lowest eigenvalue. The calculations are performed with the ATMOL program package.[13]

As a first result, we found that the GJD iterations are almost as efficient as the Davidson iterations in terms of central processing unit (CPU) time. This is in accordance with the assumption that solving a linear system in the vacuum space is efficient compared with a matrix–vector multiplication with the complete CI matrix. Furthermore, the results in Table II show that Davidson's method and the GJD method have comparable convergence rates at the extreme geometries. To explain this, we have made a comparison of the diagonal dominance of the matrices at the different geometries. This comparison is a tentative one because a robust measure for diagonal dominance has not been defined in the literature. We found that Davidson's method converges fast at the extreme geometries because the matrix $(B - \lambda I)$ is relatively diagonally dominant. However, this diagonal dominance has different causes in the equilibrium geometry and near the dissociation limit. On the one hand, at the equilibrium geometry the matrix $(B - \lambda I)$ is diagonally dominant due to the structure of the CI matrix. This is obvious from the resulting normalized CI vector, which has one element that is larger than 0.96. Therefore, the projections contained in B have no significant effect on the diagonal dominance of $(B - \lambda I)$. On the other hand, near the dissociation limit the CI matrix has large off-diagonal elements. This is reflected in the resulting CI vector, which has five components exceeding 0.42. However, the matrix $(B - \lambda I)$ is diagonally dominant because the off-diagonal elements are reduced due to the projections contained in B. The GJD method converges significantly faster with moderately elongated bonds than Davidson's method. This improved convergence is due to a more accurately calculated reference part of the update vector.

TABLE II.

Comparison of the Convergence of Davidson and the Generalized Jacobi-Davidson for H_2O at Various Geometries in MRSDCI.

OH Bond Length (R_e)	Energy (Hartree)	No. of Iterations	
		Davidson	Generalized Jacobi-Davidson
1.0	−76.2559104	9	8
1.5	−76.0706700	13	11
2.0	−75.9510327	14	10
4.0	−75.9129160	11	7
6.0	−75.9128205	12	9
8.0	−75.9128177	12	9

The reference space is a CAS containing the $3a_1$, $4a_1$, $1b_1$, $2b_1$, $1b_2$, and $2b_2$ orbitals. The convergence threshold is $\tilde{E} - E < 0.1E - 6$, where $\tilde{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

TABLE III.

Comparison of the Convergence of Davidson and the Generalized Jacobi-Davidson for H_2O at Various Geometries in MRCEPA.

OH Bond Length (R_e)	Energy (Hartree)	No. of Iterations Davidson	No. of Iterations Generalized Jacobi-Davidson
1.0	-76.2590888	8	8
1.5	-76.0731227	10	9
2.0	-75.9531400	12	8
4.0	-75.9148751	12	8
6.0	-75.9147778	12	8
8.0	-75.9147748	12	8

The reference space is a CAS containing the $3a_1$, $4a_1$, $1b_1$, $2b_1$, $2b_1$, $1b_2$, and $2b_2$ orbitals. The convergence threshold is $\tilde{E} - E < 0.1E - 6$, where $\tilde{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

The convergence behavior of both methods is shown Figure 1. The data for this figure are taken from a calculation at $2.0 \cdot R_e$. In this calculation the energy is converged to 14 digits. The figure shows that Davidson's method and the GJD method initially converge at almost equal rates. However, Davidson's method slows down after a few iterations whereas the GJD method maintains an almost constant convergence rate.

Dissociation of O_2

We consider the dissociation of O_2 as a second test case. The calculations were performed using a $(9s, 5p) \Rightarrow [4s, 2p]$ double zeta basis.[14] The energy of the ground state was computed at three geometries: at the equilibrium bond length R_e, at $1.5 \cdot R_e$, and at $2.0 \cdot R_e$, where R_e was 2.28 Bohr. The CAS contained the $3a_g$, $1b_{2g}$, $1b_{3g}$, $3b_{1u}$, $1b_{2u}$, and $1b_{3u}$ orbitals. The initial CI vector is computed as described for the H_2O test case. The MRCI and MRCEPA results are shown in Tables IV and V, respectively.

Again, Davidson's method and the GJD method converge at a comparative rate near the equilibrium geometry in the MRCI calculations. However, at moderately elongated bonds the GJD method converges significantly faster in MRCI as well as in MRCEPA calculations.
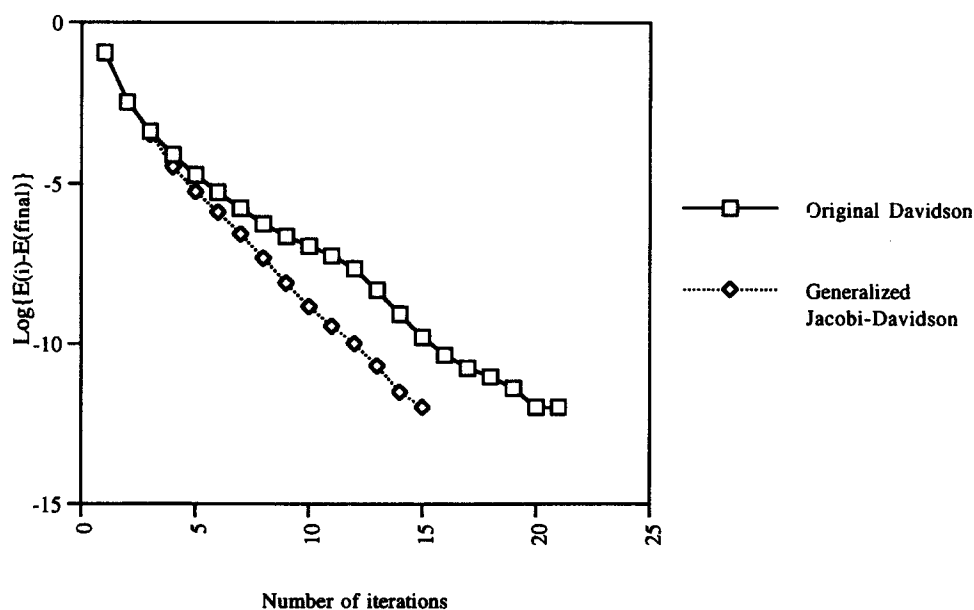
FIGURE 1. Comparison of the convergence of Davidson and Generalized Jacobi-Davidson.

TABLE IV.
Comparison of the Convergence of Davidson and the Generalized Jacobi-Davidson for O_2 at Various Geometries in MRSDCI.

O_2 Bond Length (R_e)	Energy (Hartree)	No. of Iterations	
		Davidson	Generalized Jacobi-Davidson
1.0	−149.8404582	9	9
1.5	−149.7729223	13	12
2.0	−149.7366070	16	12

The reference space is a CAS containing the $3a_g$, $1b_{2g}$, $1b_{3g}$, $3b_{1u}$, $1b_{2u}$, and $1b_{3u}$ orbitals. The convergence threshold is $\bar{E} - E < 0.1E - 6$, where $\bar{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

TABLE V.
Comparison of the Convergence of Davidson and the Generalized Jacobi-Davidson for O_2 at Various Geometries in MRCEPA.

O_2 Bond Length (R_e)	Energy (Hartree)	No. of Iterations	
		Davidson	Generalized Jacobi-Davidson
1.0	−149.8468435	10	10
1.5	−149.7802435	15	12
2.0	−149.7425726	21	13

The reference space is a CAS containing the $3a_g$, $1b_{2g}$, $1b_{3g}$, $3b_{1u}$, $1b_{2u}$, and $1b_{3u}$ orbitals. The convergence threshold is $\bar{E} - E < 0.1E - 6$, where $\bar{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

Conclusions

We derived a formulation of the modified Davidson method due to Sleijpen et al.[5] that is applicable to realistic MRCI and MRCEPA calculations. In calculations on H_2O and O_2, the new method converges significantly faster than the traditional Davidson method for geometries with moderately elongated bonds. The improved convergence properties are due to the more accurate calculation of the vacuum part of the update vector. Consequently, it is expected that this method is effective if the vacuum space is well chosen.

References

1. E. R. Davidson, *J. Comp. Phys.*, **17**, 87 (1975).

2. E. R. Davidson, *Com. Phys. Comm.*, **53**, 49 (1989).

3. C. W. Murray, S. C. Racine, and E. R. Davidson, *J. Comp. Phys.*, **103**, 382 (1992).

4. Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, 1992.

5. G. L. G. Sleijpen and H. A. van der Vorst, *A Generalized Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems*, Preprint 856, Dept. Math., Utrecht University, 1994.

6. E. Kreyszig, *Advanced Engineering Mathematics*, Wiley, Singapore, 1993.

7. V. R. Saunders and J. H. van Lenthe, *Mol. Phys.*, **48**, 923 (1983).

8. R. J. Buenker, In *Studies in Physical and Theoretical Chemistry*, R. Carbó, Ed., Vol. 21, Elsevier Scientific Publishing Company, Amsterdam, 1982, p. 17.

9. P. J. A. Ruttink, J. H. v. Lenthe, R. Zwaans, and G. C. Groenenboom, *J. Chem. Phys.*, **94**, 7212 (1991).

10. P. E. S. Wormer, F. Visser, and J. Paldus, *J. Comp. Phys.*, **48**, 23 (1982).

11. G. H. Golub and C. F. van Loan, *Matrix Computations*, North Oxford Academic, Oxford, 1986, p. 363.

12. C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.*, **85**, 2779 (1986).

13. V. R. Saunders and M. F. Guest, *ATMOL*, 1974.

14. F. B. v. Duijneveldt, *Gaussian Basis Sets for the Atoms H—Ne for use in Molecular Calculations*, RJ 945, IBM Research San José, 1971.