Chapter 6
An improvement of Davidsons
iteration method

Applications to MRCI and MRCEPA(0) calculations

H. J. J. van Dam and J. H. van Lenthe
Debye Institute, Utrecht University

G. L. G. Sleijpen and H. A. van der Vorst
Department of Mathematics
Utrecht University


Abstract

Davidsons method is widely used for finding the lowest eigenvalues of large matrices. Recently, mathematicians have shown that Davidsons 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 H₂O and O₂ with moderately elongated bonds than Davidsons original method. The new method offers new insights into the rate of convergence of Davidsons original method.
6.1 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 a perturbation vector in every iteration. Recently, mathematicians have studied Davidsons method. Saad [4] pointed out that Davidsons deduction of the perturbation 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 Davidsons method and suggest improvements.

In this paper we summarise the discussion that resulted in the corrected derivation. Also, we describe the application of the improved method in multi-reference CI (MRCI) calculations. The results are compared with the results obtained with Davidsons original method.

6.2 Davidsons perturbation vector

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

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

Therefore, if the desired eigenvector is approximated by a vector $|u\rangle$ one is interested in a perturbation vector $|\delta\rangle$ that minimises $R(|u\rangle + |\delta\rangle)$. To this end
Davidson expanded the Rayleigh quotient to second order in the perturbation vector and minimised the expression essentially yielding the linear system (see also Davidson [1] Eq (8) and Davidson [2] section 5).

\[(A - \lambda I)|\delta\rangle = -|r\rangle \]  \hspace{1cm} \text{(2)}

where \(|r\rangle\) is the residual vector

\[|r\rangle = (A - \lambda I)|u\rangle \]

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

\[|A_{jj} - \lambda| > \sum_{k \neq j} |A_{jk}| \]  \hspace{1cm} \text{(3)}

then equation (2) may be approximated by

\[(D - \lambda I)|\delta\rangle = -|r\rangle. \]  \hspace{1cm} \text{(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. The vector \(|\delta\rangle\) may be added to \(|u\rangle\) to obtain an updated approximation of the eigenvector. This updated vector may then be used to calculate a new vector \(|\delta\rangle\) and the procedure may be repeated until convergence. This approach has poor convergence characteristics. Davidson [1] pointed out that much better convergence characteristics can be obtained by expanding the eigenvector in a basis of the successive perturbation vectors \(|\delta\rangle\). The new approximation to the eigenvector is calculated by minimising the Rayleigh quotient in this basis. This is equivalent to solving a small projected eigenvalue problem. The resulting eigenvector is then used to calculate a new perturbation vector \(|\delta\rangle\) that is added to the basis. Throughout this article we assume that the perturbation vectors are used in the latter iteration scheme.

However, Saad [4] pointed out that Davidsons derivation of equation (4) is not complete. Following Davidsons line of reasoning the best results would be obtained if one solves equation (2) for \(|\delta\rangle\). This would result in \(|\delta\rangle = -|u\rangle\). However, after orthogonalisation of \(|\delta\rangle\) on the subspace spanned by perturbation vectors from earlier iterations the null vector remains. Clearly this vector can not serve as an effective extension of the subspace. Consequently, the
perturbation vector $|\delta\rangle$ should be linearly independent to the subspace whereas $|\mathbf{u}\rangle$ is a linear combination in that subspace. This suggests that Davidsons argument is not complete. In a recent paper Sleijpen et al. [5] pointed out that while deriving equation (2) Davidson assumed $|\delta\rangle$ orthogonal to $|\mathbf{u}\rangle$. However, to assure that $|\delta\rangle$ is an effective perturbation vector one should determine $|\delta\rangle$ under the condition that it is orthogonal to $|\mathbf{u}\rangle$. This may be effected by computing the perturbation vector from the projection of $A$ onto the subspace orthogonal to $|\mathbf{u}\rangle$. Sleijpen et al. showed that this is equivalent to substituting the projected matrix

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

(5)

for $A$ in equation (2) yielding

$$(B - \lambda I)|\delta\rangle = -|r\rangle.$$  

(6)

If the matrix $(B - \lambda I)$ is diagonally dominant equation (6) may be approximated to give

$$(D_B - \lambda I)|\delta\rangle = -|r\rangle$$

(7)

where $D_B$ is the diagonal of $B$. This equation is analogous to equation (4). Moreover, if the matrix $(A - \lambda I)$ is strongly diagonally dominant then a unit vector is a good approximation to the eigenvector and to $|\mathbf{u}\rangle$. In that case equation (7) will only differ from equation (4) in the first element. The difference in the resulting perturbation vectors nearly vanishes upon orthogonalising the perturbation vector on the subspace. Therefore both equations (4) and (7) may then be considered equivalent. However, equations (2) and (6) are certainly not equivalent. Due to Saad we know that a calculation based on equation (2) will not converge. To investigate the convergence of calculations with equation (6) a calculation on H$_2$O was performed. The results are shown in table I (for computational details see the Section 6.4). Note that the convergence of the energy is at least second order. That is, the number of converged digits is doubled in every iteration. This is in accordance with the fact that equation (6) resulted from a second order energy expression. The results shown in table I were obtained computing the approximations of the eigenvector from the subspace projected eigenvalue equation. We found that in this case updating the eigenvector by simply adding the perturbation vector and renormalising results in similar convergence.
Table I: The convergence of the generalised Jacobi-Davidson method in extreme for H₂O at 2.0 \( R_e \).

<table>
<thead>
<tr>
<th>Iteration Number</th>
<th>Energy (Hartree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-75.839984920556</td>
</tr>
<tr>
<td>2</td>
<td>-75.950820825635</td>
</tr>
<tr>
<td>3</td>
<td>-75.951032652229</td>
</tr>
<tr>
<td>4</td>
<td>-75.951032652292</td>
</tr>
</tbody>
</table>

For practical applications we will concentrate on equation (6). Approaches based on this equation will be referred to by the name “Generalised Jacobi-Davidson” [5] (GJD). In cases where Davidson's method converges slowly due to large off-diagonal elements equation (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 have to be found.

6.3 Application of the generalised Jacobi-Davidson method to MRCI problems

In MRCI we distinguish two categories of molecular orbitals (MO), i.e. 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 wavefunction. From the MOs two types of \( n \)-electron states may be constructed, i.e. 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
\[ |u\rangle\langle u| = P = \begin{bmatrix} P_{vv} & P_{ev} \\ P_{ve} & P_{ee} \end{bmatrix} \]

this may be approximated by

\[ P \approx \begin{bmatrix} P_{vv} & 0 \\ 0 & 0 \end{bmatrix}. \] (9)

Introducing this approximation into the matrix \((B - \lambda I)\) from equation (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

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

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

\[ (B - \lambda I)_{vv} |\delta\rangle_v = -|r\rangle_v \] (11)

\[ (D - \lambda I)_{ee} |\delta\rangle_e = -|r\rangle_e \] (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, equations (11) and (12) may also be applied to MRCEPA(0) [9] calculations because MRCEPA(0) may be expressed in a form analogous to MRCI. That is, the MRCEPA(0) matrix is equivalent to a MRCI matrix with shifted diagonal elements.

The linear system in equation (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 efficiently computed as follows. Assume the vacuum matrix \((B - \lambda I)\) is to be multiplied with a vector \(|q\rangle\). Then, using equation (5) we have

\[
(B - \lambda I)|q\rangle = A|q\rangle - |u\rangle\langle u|A|q\rangle - A|u\rangle\langle u|q\rangle
\]

\[+
|u\rangle\langle u|A|u\rangle\langle u|q\rangle - |q\rangle\lambda
\]

Writing

\[
\lambda = \langle u|A|u\rangle,
\]
deriving

\[
|z\rangle = A|u\rangle,
\]
and using that \(A\) is hermitian equation (13) becomes

\[
(B - \lambda I)|q\rangle = A|q\rangle + |u\rangle\left(\langle u|q\rangle\lambda - \langle z|q\rangle - \langle z\rangle\langle u|q\rangle - |q\rangle\lambda\right).
\]

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.

### 6.4 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. Recapitulating this means that the energy of the molecule was calculated at the equilibrium OH bond length \(R_e\), at 1.5 \(R_e\) and at 2.0 \(R_e\), 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 SCF level. The active space in the MCSCF calculation also defines the reference space in the MRSDCI and MRCEPA(0) calculations. The results in table II and III are obtained with the 55 CSF CAS from Rutting et al. [9]. The initial vector is computed solving the reference part of the CI matrix for the lowest eigenvalue. The calculations are performed with the ATMOL program package [13].
Table II: Comparison of the convergence of Davidson and the generalized Jacobi-Davidson for H$_2$O at various geometries in MRSDCI.

<table>
<thead>
<tr>
<th>OH Bond Length ($R_e$)</th>
<th>Energy (Hartree)</th>
<th>Number of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Davidson</td>
</tr>
<tr>
<td>1.0</td>
<td>-76.2559104</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>-76.0706700</td>
<td>13</td>
</tr>
<tr>
<td>2.0</td>
<td>-75.9510327</td>
<td>14</td>
</tr>
<tr>
<td>4.0</td>
<td>-75.9129160</td>
<td>11</td>
</tr>
<tr>
<td>6.0</td>
<td>-75.9128205</td>
<td>12</td>
</tr>
<tr>
<td>8.0</td>
<td>-75.9128177</td>
<td>12</td>
</tr>
</tbody>
</table>

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$_2$O at various geometries in MRCEPA(0).

<table>
<thead>
<tr>
<th>OH Bond Length ($R_e$)</th>
<th>Energy (Hartree)</th>
<th>Number of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Davidson</td>
</tr>
<tr>
<td>1.0</td>
<td>-76.2590888</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>-76.0731227</td>
<td>10</td>
</tr>
<tr>
<td>2.0</td>
<td>-75.9531400</td>
<td>12</td>
</tr>
<tr>
<td>4.0</td>
<td>-75.9148751</td>
<td>12</td>
</tr>
<tr>
<td>6.0</td>
<td>-75.9147778</td>
<td>12</td>
</tr>
<tr>
<td>8.0</td>
<td>-75.9147748</td>
<td>12</td>
</tr>
</tbody>
</table>

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.

As a first result we found that the GJD iterations are almost as efficient as the Davidson iterations in terms of 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 Davidsons method and the GJD method have comparable convergence rates at the extreme geometries. To explain this we have made a comparison of the diagonally dominance of the matrices at the different geometries. This comparison is a tentative one because a robust measure for diagonally dominance has not been defined in the literature. We found that Davidsons method converges fast at the extreme geometries because the matrix \((B - \lambda I)\) is relatively diagonally dominant. However, this diagonally 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 normalised CI-vector which has one element that is larger than 0.96. Therefore, the projections contained in \(B\) have no significant effect on the diagonally dominance of \((B - \lambda I)\). On the other hand, near the dissociation limit the CI-matrix has large off-diagonal elements.

**Figure I:** Comparison of the convergence of Davidson and generalized Jacobi-Davidson

![Graph showing convergence comparison](image)

Vertically the deviation of the energy in hartree is plotted on a logarithmic scale. The deviation is taken to be energy in the \(i\)-th iteration minus the fully converged energy. Horizontally the iteration number is plotted.
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 Davidsons method. This improved convergence is due to a more accurately calculated reference part of the perturbation vector.

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

6.5 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, i.e. at the equilibrium bond length \(R_e\), at 1.5 \(R_e\) and at 2.0 \(R_e\), where \(R_e\) was 2.28 bohr. The CAS contains the 3ag, 1b2g, 1b3g, 3b1u, 1b2u, and 1b3u orbitals. The initial CI vector is computed as described for the H\(_2\)O test case. The MRCI and MRCEPA(0) results are shown in table IV and V respectively.

Again Davidsons method and the GJD method converge at a similar 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(0) calculations.

6.6 Conclusions

We derived a formulation of the modified Davidson method due to Sleijpen et al. that is applicable to realistic MRCI and MRCEPA(0) calculations. In calculations on H\(_2\)O 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 perturbation vector. Consequently, it is expected that this method is effective if the vacuum space is well chosen.

Table IV: Comparison of the convergence of Davidson and the generalized Jacobi-Davidson for O₂ at various geometries in MRSDCI.

<table>
<thead>
<tr>
<th>O₂ Bond Length (Rₑ)</th>
<th>Energy (Hartree)</th>
<th>Number of Iterations</th>
<th>Davidson</th>
<th>Generalized Jacobi-Davidson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-149.8404582</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>-149.7729223</td>
<td>13</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>2.0</td>
<td>-149.7366070</td>
<td>16</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

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 $\frac{\Delta E}{E} < 0.1 \times 10^{-6}$, where $\Delta E$ 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 O₂ at various geometries in MRCEPA(0).

<table>
<thead>
<tr>
<th>O₂ Bond Length (Rₑ)</th>
<th>Energy (Hartree)</th>
<th>Number of Iterations</th>
<th>Davidson</th>
<th>Generalized Jacobi-Davidson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-149.8468435</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>-149.7802435</td>
<td>15</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>-149.7425726</td>
<td>21</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

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 $\frac{\Delta E}{E} < 0.1 \times 10^{-6}$, where $\Delta E$ is the current approximation to the eigenvalue and $E$ is the fully converged eigenvalue.
References