Convergence behaviour of multireference perturbation theory: an indicator

By HUUB J. J. VAN DAM and JOOP H. VAN LENTHE

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

(Received 10 September 1996; revised version accepted 4 November 1996)

An approach to investigate the convergence behaviour of a (multireference) Möller–Plesset perturbation series is proposed in which the series is evaluated up to arbitrary order without reference to excitations beyond doubles. It is shown that this approach qualitatively reproduces the convergence behaviour of the complete perturbation series for stretched H₂ molecules and H₂O molecules. The MRMP approach does show improved convergence characteristics compared with single reference MP for stretched H₂ molecules. Applications of the approach to O₂ at 4·0 a₀ and N₂ at 5·0 a₀ show that the MRMP series is divergent for these cases. Furthermore, in the latter case it was found that the multireference series shows much stronger divergence than the single reference series.

1. Introduction

Recent developments in multireference perturbation theory (MRPT), and complete active space perturbation theory (CASPT) [1–4] in particular, have led to programs which accomplish correlating calculations on relatively large molecular systems. As a consequence, MRPT has become popular. However, to limit the computational cost most programs cannot evaluate the perturbation series beyond second order, which is only the first term in the correlation energy. A few programs based on a multireference singles–doubles configuration interaction (MRSDCI) approach can evaluate at best the third-order correction. Thus, these results are not converged or, worse, they do not have to be part of a converging series. If the latter is the case it even is debatable whether to consider the truncated series as an approximation to some full series.

We constructed an MRPT program [5] and calculated some potential energy surfaces [6]. Already at the Möller–Plesset MP3 level some results suggested that the perturbation series may diverge. Therefore, we developed an indicator that should make it possible to investigate the convergence behaviour of the series. The indicator is based on the assumption that the singularities leading to divergence belong to the space spanned by the reference function and all single and double excitations.

2. Theory

The program [5] is based on the theory presented by Pulay, Wolinski, and Sellers [7, 8] and uses an approach similar to those in Molcas [1–3] and Molpro [4]. Basically the program uses the ‘first diagonalize then perturb’ approach, thus applying Möller–Plesset perturbation theory to a general contracted multiconfigurational reference function. The singly and doubly excited states are obtained by (repeated) action of single excitation operators on the reference function and subsequent orthonormalization, thus generating a contracted CI space.

The zero-order Hamiltonian was chosen to be

\[ H_0 = P_0 F P_0 + P_1 F P_1 + P_2 F P_2 + \cdots, \]

where \( F \) is taken to be the closed shell Fock operator, \( P_0 \) is the projector onto the reference function, \( P_1 \) is the projector onto the space of singly excited states, \( P_2 \) projects onto the doubles, etc. Note that there is no interaction between states of different excitation levels through \( H_0 \). This \( H_0 \) is slightly different from that proposed by Andersson et al. [2] where there is a separate projector onto the internally excited states. We assume that this does not make an essential difference, because if there are near degenerate internally excited states in our \( H_0 \) they will be low lying excitations in Andersson’s \( H_0 \) too. This has been found to be true in a test on the O₂ molecule. The perturbation series [9, 10] is generated by

\[ (H_0 - E^{(0)}) \psi^{(k)} = \sum_{l=1}^{k} E^{(l)} \psi^{(k-l)} - (H - H_0) \psi^{(k-1)}, \]

where

\[ E^{(0)} = \langle \psi^{(0)} | H_0 | \psi^{(0)} \rangle, \]

\[ E^{(k)} = \langle \psi^{(0)} | H - H_0 | \psi^{(k-1)} \rangle, \quad k > 0, \]
where $H$ is the Hamiltonian matrix in the bases of contracted states, and $E^{(k)}$ and $\psi^{(k)}$ are the $k$th order correction to the energy and the wavefunction, respectively. Using the above equations the perturbation series can be evaluated to any order provided that the complete $H_0$ and $H$ are available. In that case the total energy and the total wavefunction upon convergence would be the full CI energy and wavefunction. In practice the complete matrices are too large to deal with, and may not be needed to investigate the convergence behaviour of the series. The latter is due to the fact that the convergence or divergence of the series depends mostly on the properties of $(H_0 - E^{(0)})^{-1}$.

Due to the choice for the projection operators $H_0$ is block diagonal, and as a consequence the inverse,

\begin{equation}
(H_0 - E^{(0)})^{-1} = (P_0 FP_0 - E^{(0)})^{-1} + (P_1 FP_1 - E^{(0)})^{-1} + \cdots,
\end{equation}

is block diagonal too. Note that the first term is undefined because $(P_0 FP_0 - E^{(0)})$ is singular, but this term operates only on $\psi^{(0)}$ which is orthogonal to any of the other $\psi^{(k)}$. Therefore the first term is allowed to take any value, or it may even be removed. Considering the other diagonal blocks in the inverse matrix it is extremely unlikely to have singularities due to near degeneracies in blocks corresponding to triply and higher excited states without having singularities in the blocks corresponding to the singly and doubly excited states. Thus, assuming that if the matrix contains singularities then this implies there will be singularities in the space spanned by the singles and doubles, we may drop the terms in the spaces of higher excitations and still expect to capture all contributions that determine the convergence characteristics. In essence this approach is equivalent to performing contracted MRSDCI using Møller–Plesset perturbation theory as eigenvalue solver. However, our goal is to study the convergence characteristics, not to do contracted CI.

To obtain the limiting energies to which the perturbation series should converge, we implemented a contracted CI option also. For this we used the contracted matrix–vector product $H \psi$ from the perturbation theory code and plugged it into a Davidson diagonalization [11, 12] routine.

### 3. Tests

To validate the approach in practice we choose non-interacting H$_2$ molecules as a first test case. Because one H$_2$ molecule is a 2 electron system the space spanned by all singles and doubles is equivalent to the full CI space. This means we can evaluate the perturbation series without approximations to any order, and use these results as a reference. If we have $N$ non-interacting H$_2$ molecules this system is physically equivalent to $N$ systems each containing one H$_2$ molecule. Therefore, an appropriate model should behave similarly for $N$ non-interacting H$_2$ molecules as for $N$ times one H$_2$ molecule. Here this means that the perturbation approach restricted to singly and doubly excited states should show similar convergence behaviour for different numbers of non-interacting H$_2$ molecules. Furthermore the orbital energies in the molecule can be varied through changing the bond length, thus introducing near degeneracies at will.

In the calculations we use H$_2$ molecules stretched to twice the equilibrium bond length. The molecules are lined up parallel to each other $50 \sigma_0$ apart ($a_0 = \text{bohr} \approx 5.29177 \times 10^{-11} \text{ m}$). The DZ basis set by Dunning [13] is used with an additional p function (exponent 0.80). Both single reference and multi-reference calculations are performed. The single reference calculations involve 1, 3 and 5 H$_2$ molecules. In the multireference calculations 1, 2 and 3 molecules are involved, using a CAS reference containing the 1s and 2s* orbitals of each H$_2$ molecule. In figures 1 and 2 the deviation of the $N$th order corrected perturbation energy from the (contracted) CI energy per molecule is plotted on a logarithmic scale. Figure 1 shows that although convergence is not smooth because of the stretched bonds, the single reference perturbation series converges in a similar manner for 1, 3 and 5 H$_2$ molecules. Figure 2 shows that the multireference approach removes the degeneracy and the series converges smoothly. Also, for the multireference series the convergence is similar for 1, 2 and 3 H$_2$ molecules.

The H$_2$ test may seem an academic one. We now turn to a more realistic test involving the H$_2$O molecule. We use the results of Handy et al. [10] in a study on the convergence of single reference MP for the groundstate

![Figure 1. MPPT calculations on non-interacting H$_2$ molecules; $r_{\text{H}-\text{H}} = 2.776 \sigma_0$, and $r_{\text{H}_2-\text{H}_2} \geq 50.0 \sigma_0$.](image-url)
of the water molecule. We use their geometries at \( r_e \) and 2\( r_e \) (\( r_e = 0.967 \) Å and \( \theta = 107.6^\circ \)) and basis set (6-21G) [14–18] to calculate the MP series in the singles–doubles space only. The results are shown in figures 3 and 4 in which the absolute value of the deviation of the MPPT energy from the CI energy is plotted logarithmically. At the equilibrium geometry the perturbation series converges reasonably well in both the SDCI space and the full CI space. In the full CI space convergence is somewhat slower because the 9-fold and 10-fold excited states do not come into play before the ninth order, whereas in the SDCI space all configurations are involved already at second order. Apart from this the curves are similar. At twice the equilibrium bond length convergence is much slower, but also in this case both curves are qualitatively the same. Unfortunately, Handy et al. did not provide results for even larger OH distances, where the perturbation series may be just divergent. With our approach we find extremely slow convergence leading to real divergence setting in at 23rd order at 2.5\( r_e \).

Olsen and coworkers have shown that the convergence behaviour of the Møller–Plesset perturbation series is basis set dependent [19, 20]. For example, the single reference MP series in the full CI space for the neon atom diverges if a basis set with diffuse functions is used. The approach we suggest in this article failed to reproduce these results both in single and multireference calculations. The reason for this is that the highly excited and highly diffuse states that are responsible for those divergences are not present in our expansion because we allow for at most two electrons in the external orbitals.

Recapitulating, the proposed convergence test qualitatively reproduces the convergence behaviour in the full CI space for \( \text{H}_2 \) molecules both in single reference and multireference calculations, as well as in single reference calculations on \( \text{H}_2\text{O} \). From the tests performed we conclude that the approach chosen seems to be satisfactory for studying the convergence behaviour. Using this result we apply the test to the nitrogen molecule at the equilibrium geometry (\( R = 2.1 \) \( a_0 \)) and at a stretched geometry (\( R = 5.0 \) \( a_0 \)), comparing the single reference and the multireference series. Also, the test is applied to the oxygen molecule at \( R = 4.0 \) \( a_0 \).

4. Application to \( \text{N}_2 \)

In our report of our implementation of multireference Møller–Plesset perturbation theory [6] an application to the nitrogen molecule suggested that the multireference series is divergent at \( R = 5.0 \) \( a_0 \). Therefore, a closer inspection seems appropriate. Also, we want to comment on the 'classical' idea that perturbation theory is allowed when the perturbation (\( H - H_0 \)) is small enough, whereas we emphasized in the theory section that non-singularity of \( (H_0 - E^{(0)}) \) in the orthogonal
complement space of $\psi^{(0)}$ is essential. For this the perturbation series is calculated at the equilibrium geometry ($R = 2\cdot1 \ a_0$) and the stretched geometry ($R = 5\cdot0 \ a_0$) using the 6-31G* [21] basis set. Both multi-reference calculations using a CAS in the $5\sigma$, $6\sigma^*$, $1\pi_x$, $2\pi_y^*$, $1\pi_y$ and $2\pi_x^*$ orbitals as a reference and single reference calculations are performed. The IS orbitals are frozen at the MCSCF level and not correlated. The results are shown in figures 5 and 6.

From the calculations at the equilibrium geometry it is clear that the multireference series converges a little faster than the single reference series. However, at the stretched geometry both series diverge but there is a significant difference in the behaviour of the approaches. The single reference series diverges slowly and the divergence starts immediately (second order). In the multireference series the first term (MRMP2) improves the energy and the divergence sets at third order, after which the series diverges strongly.

We measure the size of the perturbation by the size of the first-order correction to the energy. These values together with the zero-order energies are given in the table. At both geometries the perturbation is smaller for the single reference approach than for the multi-reference approach. Also, the zero-order energy is lower for the single reference calculations than for the multireference calculations, suggesting that the single reference zero-order approximation is the better one. This suggestion is odd because the single reference zero-order wavefunction is lacking an essential physical quality, i.e., proper dissociation. Furthermore, at the equilibrium geometry the multireference series converges faster than the single reference, even though the perturbation is larger. At the stretched geometry the multireference approach diverges faster than the single reference approach. Therefore, it is not possible to deduce information on the convergence from the size of $E^{(0)}$ or $E^{(1)}$.

5. Application to $\text{O}_2$

As well as the $\text{N}_2$ molecule we have reported also the potential energy surface of $\text{O}_2$ [6]. Because the $\text{N}_2$ series diverges at $R = 5\cdot0 \ a_0$ we revisit the $\text{O}_2$ molecule here to validate our original results. The triplet groundstate $\text{O}_2$ molecule dissociates into two triplet groundstate O atoms. In various calculations it was found that the potential energy surface for this dissociation has a maximum [22–26]. Recently, Byrman [22] has studied this maximum using valence bond calculations in which the wavefunction was written as a superposition of structures describing the triplet $\text{O}_2$ molecule (molecular part) and structures describing two triplet O atoms (atomic part). These two parts differ in the spin paths of their structures. It was found that the energy of the molecular part rises sharply for $\text{O}–\text{O}$ distances larger than 3.0 $a_0$, whereas the energy of the atomic part comes down. Thus the maximum was found to be due to the ‘crossing’ of these two curves, which means a spin recoupling. The spin recoupling suggests that the dissociation of the $\text{O}_2$ molecule is a serious test for a method aimed at calculating potential energy surfaces. Therefore, we calculate the perturbation series within the singles–doubles space up to 7th order. The calculations are performed within $D_{2h}$ symmetry $B_{1u}$ using a DZ basis [13] with additional spherical harmonic d functions (exponent 1.2). We use a 6-orbital 8-electron CAS reference function, but the IS orbitals are frozen at the MCSCF level and not correlated. In figure 7 the deviation of the perturbation energy is plotted and in figure 8 the 2-norm of the $i$th order correction to the wavefunction is plotted. Both plots are on a logarithmic scale, and from them it is clear that the perturbation series is divergent in this case. The norm of the correction vector
Zero-order and first-order energies for single and multireference calculations on the nitrogen molecule.

<table>
<thead>
<tr>
<th></th>
<th>$E^{(0)}/E_h$</th>
<th>$E^{(1)}/E_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR 2.1</td>
<td>-70.91866361</td>
<td>-38.04781784</td>
</tr>
<tr>
<td>MR 2.1</td>
<td>-70.38351456</td>
<td>-38.72366820</td>
</tr>
<tr>
<td>SR 5.0</td>
<td>-69.62628236</td>
<td>-38.46756589</td>
</tr>
<tr>
<td>MR 5.0</td>
<td>-67.38477070</td>
<td>-41.40496354</td>
</tr>
</tbody>
</table>

$E_h$ = hartree $\approx 4.35975 \times 10^{-18}$ J.

Increases monotonically. Although the series is divergent our MRMP2/3 estimates of the potential surface [6] show no sign of any problems. This suggests that it is rather dangerous to judge a perturbation series by its first few terms only.

6. Conclusion

We propose an approach to study the convergence behaviour of MRPT series without reference to excitations beyond doubles. Using this approach we qualitatively reproduce the convergence behaviour of the full series for the stretched $H_2$ molecule, both single reference and multireference, and the single reference $H_2$ molecule at the equilibrium geometry and a stretched geometry. Applying the same approach to the oxygen molecule stretched to a bond length of 4.0 $a_0$ shows that the MRPT series is divergent at that geometry. Also we confirm that the MRPT series is divergent in the nitrogen molecule stretched to a bond length of 5.0 $a_0$, as we suspected from calculations published recently. While we failed to show the diffuse basis set effects of Olsen and coworkers on the neon atom, we do show that divergences may appear even with small basis sets in typical multireference situations. Because the MRMP2/3 energies do not always indicate any possible trouble although the series is divergent we believe that MRPT results should be used with caution.

The authors would like to thank Jeppe Olsen for sending copies of his manuscripts and for helpful discussions. The investigations were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO). Some basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 1.0, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory, which is part of the Pacific Northwest Laboratory, PO Box 999, Richland, Washington 99352, USA, and funded by the US Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the US Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller, Karen Schuchardt, or Don Jones for further information.

References


