Calculations on correlation effects in molecules

Convergence and size-consistency of multi-reference methods

Berekeningen aan correlatie effecten in moleculen

Convergentie en size-consistency van multi-reference methoden

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht, op gezag van de Rector Magnificus, Prof dr. H. O. Voorma, ingevolge het besluit van het college van decanen in het openbaar te verdedigen op maandag 13 oktober 1997 des middags te 12.45 uur

door

Hubertus Johannes Jacobus van Dam

geboren op 20 juli 1968, te Grubbenvorst promotor: Prof. dr. F. B. van Duijneveldt

co-promotor: Dr. J. H. van Lenthe

beiden verbonden aan de faculteit Scheikunde
van de Universiteit Utrecht

Het in dit proefschrift beschreven onderzoek werd verricht onder auspiciën van de Stichting Scheikundig Onderzoek Nederland (SON) met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

ISBN 90-393-2031-4

Calculations on correlation effects in molecules

Convergence and size-consistency of multi-reference methods

Huub van Dam

Contents

Chapter 1

Introduction

This thesis is about methods for electronic structure calculations on molecular systems. The ultimate goal is to construct methods that yield potential energy surfaces of sufficient accuracy to allow a qualitatively correct description of the chemistry of these systems; i.e. heat of formation, isomerisation barriers, equilibrium geometries, and vibrational spectra. As a step towards this goal most of this text deals with multi-reference Møller-Plesset [1] perturbation theory (MRMP). Additionally, multi-reference configuration interaction (MRCI) [2,3] and multi-reference coupled electron pair approximations (MRCEPA(0)) [4] feature regularly as competing alternatives.

In quantum chemistry potential energy surfaces are calculated from the time independent Schrödinger equation [5,6] within the Born-Oppenheimer approximation [7], symbolically denoted as

$$\hat{H}\Psi = E\Psi . \tag{1}$$

This equation defines the total energy E of the molecule at a chosen fixed geometry. Actually, E is a function of the nuclear coordinates called the potential energy surface of the molecule. The name potential energy surface stems from the fact that this function is the potential energy in the Schrödinger equation for the nuclear motion. The potential energy surface contains information of chemical importance, e.g. equilibrium geometries are at the minima, transition states are at saddle points, force constants are determined by the curvature near the minima. So the potential energy surface holds the key information needed in thermochemistry, reaction kinetics and infrared spectroscopy. Thus it is worthwhile to try and calculate it.

1.1 Approaches to solving the Schrödinger equation

1.1.1 Discretisation

The standard approach to solving differential equations such as the Schrödinger equation is to discretise it using some basis. The solution is then represented as a function of some parameters that remain to be calculated. In quantum chemistry this is typically a multi level approach executed as follows.

First functions φ_i of the coordinates of one electron ("1") are defined. These functions are called orbitals referring to the orbits of particles in classical dynamics. In most approaches today the orbitals are expressed as a linear combination of Gaussian functions $\chi^A(1)$

$$\phi_i = \sum_j c_{ij} \chi_j^A(1) \tag{2}$$

centered at the atomic positions A. These Gaussian functions are termed atomic orbitals and the linear combinations are called molecular orbitals.

In this thesis it is assumed that these expansions are chosen such that the molecular orbitals are orthonormal. To incorporate the intrinsic magnetic moment of the electrons the molecular orbitals are multiplied by a spin function, α or β, depending on the spin state of the electron, giving spin-orbitals. Throughout this thesis a spin-restricted approach is assumed, that is, the same expansion coefficients are used for both α and β spin-orbitals. The advantage of the spin restricted formalism is that methods may be constructed that are guaranteed to yield solutions that are spin-eigenfunctions. This often is convenient in applications in the field of spectroscopy.

From these spin-orbitals *N*-electron functions are constructed in the form of Slater determinants [8,9] to enforce the Pauli exclusion principle [10]

$$\Theta = |\phi_1 \phi_2 \ldots \phi_n|. \tag{3}$$

Different determinants may be constructed by replacing one or more orbitals by an other orbital. Subsequently the total wavefunction can be approximated by a linear combination of Slater determinants

$$\Psi = \sum_i C_i \Theta_i. \tag{4}$$

The wavefunction presented in this form is completely determined if the coefficients in equations (2) and (4) are known. The aim is to calculate these coefficients to approximate a solution of the Schrödinger equation. Various recipes to determine these coefficients will be considered.

1.1.2 Levels of approximation: Hartree-Fock, limited CI, full CI

The simplest recipe is the one in which the wavefunction is approximated by a single determinant. This approach is called the Hartree-Fock method [11,12] and it amounts to separating the Schrödinger equation into differential equations for single electrons. To decouple the differential equation for an electron from the equations of the other electrons the electron-electron interaction is integrated over the coordinate space of the other electrons. As a result an effective one-particle equation is obtained that accounts for an average interaction with the other electrons.

The average electron interaction that is obtained in the Hartree-Fock method may be too poor an approximation in some applications. To describe the actual interaction between the electrons as completely as possible all linearly independent determinants have to be taken into account variationally. Such an approach is called full Configuration Interaction (full CI) and it gives the exact answer within the chosen atomic orbital basis. The energy difference between the Hartree-Fock energy and the full CI energy is called the correlation energy [13].

Hartree-Fock and full CI are at the extrema of the spectrum of methods that may be applied. In principle the exact solution within the basis may be obtained using full CI. However, in practice such calculations are feasible only for small molecules. The problem is that the number of determinants grows factorially with the number of electrons and basis functions. Therefore in practice often only a subset of all determinants is selected. It can be shown that due to the orthogonality of the orbitals and the absence of more than two-particle interactions in the Hamilton operator only determinants with at most double substitutions from a reference determinant can have direct interaction with this reference. Based on this notion the approaches most often used today describe the wavefunction in terms of a (few) reference determinant(s) and all single and double substitutions. The most straightforward approach in this vein is the singles-doubles configuration interaction approach (CISD) where the determinantal coefficients C_i are optimised variationally.

The limited many-configuration approaches like CISD yield energies that depend on the shape of the orbitals. Therefore the orbital optimisation becomes crucial and a qualitatively correct starting point is essential. However the

Hartree-Fock approach as described up to now may not yield a suitable starting point for an accurate many-configuration treatment. One reason is that the Hartree-Fock method fails to describe the bond breaking in molecules when an electron pair must be dissociated.

As an example consider the potential energy surface of ethylene for the rotation of one methylene group relative to the other [14]. In the planar geometry (D_{2h}) the singlet ground state is well described by a single configuration

$$1a_g^2 1b_{1u}^2 2a_g^2 2b_{1u}^2 1b_{2u}^2 3a_g^2 1b_{3g}^2 1b_{3u}^2. \tag{5}$$

However if the C-C torsion angle approaches 90˚ the π–bond is broken and two singlet configurations with exactly the same energy appear (D_2)

$$1a_1^2 1b_1^2 2a_1^2 2b_1^2 1b_2^2 1b_3^2 3a_1^2 2b_2^2 \tag{6}$$

and

$$1a_1^2 1b_1^2 2a_1^2 2b_1^2 1b_2^2 1b_3^2 3a_1^2 2b_3^2. \tag{7}$$

Such a situation is a typical case of near degeneracy. Clearly it is completely arbitrary to select either of these configurations as the Hartree-Fock singlet ground state. Also, in a CI treatment both configurations will have an equally large coefficient. Therefore, the only reasonable alternative is to select both configurations to describe the potential energy surface of the singlet A state.

To deal with these situations a Hartree-Fock type approach called multi-configurational Hartree-Fock (MCHF) may be used that can take a few determinants into account. For historical reasons this method is often referred to as multi-configuration Self Consistent Field (MCSCF). The latter name refers to the optimisation strategy that was initially used for such wavefunctions.

In a MCHF calculation three classes of orbitals may be distinguished, those that are doubly occupied in all determinants, those that are unoccupied in all determinants, and orbitals that differ in occupation for different determinants. These orbitals are called inactive, virtual and active orbitals respectively. The term "virtual orbitals" refers to the fact that these orbitals do not contribute to the MCHF wavefunction and to the physical properties.

The energy difference between MCHF and Hartree-Fock is called non-dynamical correlation [15,16] because MCHF does not really take the actual motion of the electrons into account. MCHF merely corrects the most severe artefacts of the Hartree-Fock method by including near-degenerate configurations. In contrast to this the energy difference between MCHF and Full CI is called dynamical correlation [17]. The terms dynamic and non-dynamic correlation are, however, not well defined. For example MCHF will include some dynamical correlation also. Therefore these terms may only be used in a qualitative sense.

1.1.3 Multi-reference approaches

In the case that a MCHF approach is used for the first step in a calculation it is natural to define many-configuration methods that include all single and double substitutions from all the determinants used in the MCHF calculation. There are three main approaches to do this.

The conceptually most simple approach is to just generate all single and double excitations from all reference determinants and use this determinantal basis as such. This approach may be called an uncontracted approach. It has been studied by Siegbahn [2] and it is used for example in the MRCI program by Saunders *et al.* [3]. To reduce the computational cost the most significant singles and doubles may be selected to participate in the CI calculation. This has been implemented by Buenker *et al.* [18].

Alternatively all single and double excitations from each determinant may be generated yielding a many-determinantal function for each reference determinant. Per reference determinant the singles-doubles wavefunction is calculated and subsequently kept fixed. The total wavefunction is expressed as a linear combination of these many-determinantal wavefunctions. Such an approach is called externally contracted. It has been investigated for example by Siegbahn [19,20] and it is implied in a number of effective Hamiltonian approaches in the literature [21-23].

Finally, a (fixed) linear combination of the reference determinants may be selected as the reference function, e.g. the ground-state MCHF wavefunction for the system under consideration. Subsequently the total wavefunction may be expanded using all single and double excitations from this multi-determinantal

reference function. This approach is said to be internally contracted. It has been investigated by Siegbahn [24], and by Werner [25].
In this thesis only the uncontracted and internally contracted approaches will be discussed.

1.2 Properties of potential energy surfaces

1.2.1 Size extensivity and size consistency

The total energy of a system depends on the number of particles involved, i.e. the size of the molecule. In general this dependence is not a simple relation. However, there are some limiting cases to be discussed below, in which a simple relation can be given. An acceptable method should be consistent with this relation, i.e. be size consistent. In practice size consistency is not easily obtained and a significant part of this thesis deals with this problem.

To study the scaling of the energy with the number of particles in a molecule consider the example of a large monoatomic crystal. In this case the total energy of the crystal may be calculated with high precision from an average energy per atom in the crystal. So for a large crystal the total energy may be considered linear in the number of composing particles. More precisely, in the limit when the number of particles approaches infinity at constant density the average total energy per composing particle approaches a constant. In thermodynamics a property that satisfies this requirement is said to be extensive. Inspired by this Bartlett [26] has suggested to call approaches that maintain this quality size extensive.

For approximate approaches to solving the Schrödinger equation it is rather difficult to prove the size extensivity. This has been done only for Hartree-Fock, single-reference Møller-Plesset perturbation theory, single-reference coupled cluster, and full CI. To show by calculations that a method is size extensive is cumbersome because the average energy per particle converges slowly with the number of particles. This means that a numerical test is out of the question because the number of particles needed is too large for practical calculations.

However in the limiting case where the interaction between the subsystems approaches zero the total energy becomes linear for any number of subsystems.

An approach that maintains this quality is said to be size consistent [27]. Such a situation is obtained for instance when considering two atoms at a large separation. In this case the total energy clearly should equal the sum of the atomic energies, i.e. for identical atoms this is equivalent to stating that $1+1$ should be 2. In cases like these it may be possible to test the size consistency both theoretically and numerically. Therefore in this thesis size consistency is considered rather than size extensivity.

It may seem a bit overdone to stress the proper scaling of the energy with the number of particles so much. However there are historical reasons for this. The often used singles-doubles configuration interaction method (CISD) yields dynamical correlation energies that scale like the square root of the number of non-interacting electron pairs (i.e. $1+1 \approx \sqrt{2}$). This scaling violates the required linear scaling so severely that it will cause a major error in the calculated energies even for relatively small numbers of electrons.

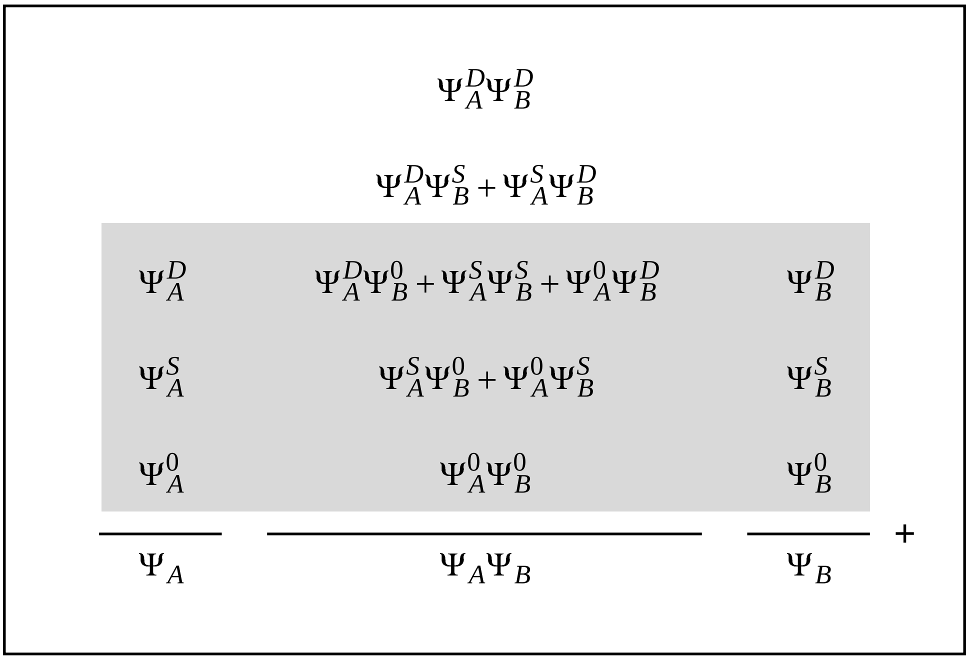


Figure I: The CISD wavefunctions for the monomers and the direct product wavefunction of the dimer in relation to the CISD wavefunction for the dimer. The CISD terms are in the grey area.

This poor scaling may be explained by considering a system of two identical atoms without interaction. Because there is no interaction the wavefunction of the two atoms together should be the direct product of the wavefunctions of the separate atoms. This means if a CISD calculation is performed on the separate atoms then the direct product wavefunction for the dimer should contain at least some four-fold excitations, see figure **I**. Clearly a CISD calculation on the total system cannot reproduce this wavefunction because the triple and quadruple excitations are missing. CISD is clearly unsuitable for obtaining size consistent results. However size concistency is essential in comparing results for systems of different size.

1.2.2 Invariance of the energy for orbital rotations.

As mentioned before full CI gives the exact energy within the atomic orbital basis. This means that the full CI energy does not depend on the particular shape of the orbitals. This is no longer true for an approach that uses a limited set of determinants. In such a case the orbitals are optimised using a Hartree-Fock type approach. However, the orbitals obtained from such an approach are not uniquely defined. This results from the fact that the wavefunction obtained is in general invariant for unitary transformations among the inactive orbitals, among the virtual orbitals and among degenerate active orbitals. Therefore there is no physical criterion to restrict these rotations. This means that any method to calculate the dynamic correlation energy using these orbitals must be invariant for these rotations also, otherwise a discontinuous potential energy surface may be obtained. Such a potential energy surface would be unphysical.

1.3 Requirements for an appropriate *ab initio* method

Having described the most important aspects in *ab initio* theory with respect to this thesis a list of required properties for an acceptable method may be compiled. An early attempt to generate such a list was published by Pople *et al.* [27] when trying to formulate "theoretical model chemistry". This list contained four points

1) A method should provide well-defined results for the energies of electronic states for any arrangement of fixed nuclei, leading to a set of continuous potential energy surfaces.
2) A method should be such that the amount of computation does not increase too rapidly with the size of the system.

3) A method should be size consistent.
4) A method should yield upper bounds to the exact solutions, i.e. be variational.

This list followed an earlier version presented at the Boulder conference in 1972 [28]. A remarkable difference is that at the conference Pople felt that interpretability of a model was important, while the upper bound property was not mentioned then. In the list reproduced above Pople introduced the upper bound property at the expense of the interpretability but he did not comment on this change, nor did he explain why the upper bound property is important.

Half a decade later another list was published by Bartlett [26]. This list stated that a method should be

1) size extensive.
2) generally applicable to a wide class of problems within one framework, i.e. not dependent on specific choices of configurations
3) invariant to unitary transformations among degenerate orbitals
4) efficient and cost effective
5) applicable to excited states and open shells
6) able to dissociate a molecule correctly into its fragments.

The second, third and fifth item on this list can together be considered as a more explicit statement of Pople's first requirement. New on this list is the realisation that even a method that yields continuous potential energy surfaces may yield unphysical results if it can not describe dissociations properly. This realisation is at the heart of the development of MCHF and multi-reference methods although these methods conflict with the second requirement. Multi-configurational approaches for proper dissociation explicitly depend on the right choice of configurations. Furthermore, the size extensivity of a method has been raised to the prime position whereas the upper bound property has been dropped.

The reason the upper bound property was found to be of little importance is that this bound is of little value without a lower bound. There have been some attempts to calculate lower bounds (e.g. [29]) but none of the proposed methods has been very successful. So we are left in a situation where the lower bound is unknown. In this case the upper bound gives no real information about the accuracy of the calculated property. Also, many quantities of chemical interest are calculated from energy differences. Therefore, even if the energies are upper bounds this does not imply any bound on the calculated difference. So in practice the upper bound property is of little value.

Before turning to the list of requirements aimed at in this thesis one should distinguish between two approaches to obtain good descriptions of chemical phenomena.

The first approach is to calculate the energy of chemical systems as accurately as possible. From these very accurate energies relevant properties may be calculated with high precision.

The second approach is to devise methods that capture all the necessary physical qualities but are not necessarily extremely accurate. However because all the relevant physics is included it may be hoped that the calculated properties are of chemical significance and allow for qualitatively correct chemical predictions.

The first approach may be inevitable in some areas of research such as in the studies on van der Waals interactions or floppy molecules. The second approach may be sufficient to study chemical reactions and electronic spectra. This thesis follows the second approach.

In order to obtain results that are sufficiently accurate for studying potential energy surfaces it is at least required that a method should be

1) well defined at a wide range of geometries
2) able to dissociate a molecule properly
3) invariant for different orbital representations of the (reference) wavefunction
4) size consistent
5) able to calculate energies of similar quality at a wide range of geometries
6) efficient.

The first item underlines that a method should yield a continuous potential energy surface at all reasonable geometries. The fifth point in practice means that the dynamic correlation energy should be accounted for to some extent.

1.4 Outline of this thesis

This thesis will address methods to calculate potential energy surfaces. The performance of these methods will be checked against the requirements listed in the previous section. Initially, it was clear that in order to dissociate molecules properly a multi-configurational starting point is essential. This means that all methods that will be discussed are based on a multi-reference wavefunction where the reference function is optimised using MCHF.

Beyond the MCHF method there are various methods to account for the correlation energy of a molecule. From studies on single-reference treatments it was known that the perturbation theory formulated in 1934 by Møller and Plesset [1] (MP) is an exactly size consistent and efficient method for such calculations. It was known also that the single-reference MP method diverges in cases where two states are close in energy. The basic idea at the start of the work described here was to avoid these divergences by including all nearly degenerate states in the reference function thus generalising the method to the multi-reference case. If this would be possible while retaining the size consistency a very efficient and highly accurate multi-reference Møller-Plesset (MRMP) method would be obtained.

The questions concerning this approach were: "Is it possible to avoid the divergences using a multi-reference treatment? Does the perturbation theory retain its size consistency if it is generalised?" The literature at the time [30,31] suggested that MRMP naturally would lead to an improvement of the results calculated. However, as far as the size consistency was concerned it was uncertain if it could be retained. Some had come to believe that size consistent multi-reference treatments are impossible in principle [32,33].

The major part of this thesis addresses the points raised in the previous paragraph. In chapter 2 the implementation of MRMP in the ATMOL [34] program package is described. Chapter 3 addresses the convergence behaviour of the perturbation series and a method to test it is suggested. Chapters 4 and 5 discuss the size consistency of MRMP both mathematically and based on applications.

Perturbation theory is not the only method that yields size consistent results. Already in the sixties it was known that some electron pair approximations give size consistent correlation energies also. In the single-reference case it was shown that coupled electron-pair approximations (CEPA) could be used as a size consistent approximation to coupled cluster in the singles-doubles configuration space. In practice CEPA may be implemented such that estimates of the energy contributions lacking in the CISD method are added to the diagonal of the Hamiltonian matrix. These estimates depend on the shape of the wavefunction obtained so far. This way the size consistency is incorporated in an iterative manner. Ruttink *et al.* [4] have generalised this approach to the multi-reference case (MRCEPA(0)). Although this approach is not exactly size consistent it is the best alternative to MRMP we have available. For this reason the MRMP results in this thesis are often compared to results obtained with MRCEPA(0).

At the heart of the MRCEPA(0) is the Davidson diagonalisation method that is used to iteratively solve the eigenvalue equations. The efficiency of the MRCEPA(0) depends primairily on the rapid convergence of the Davidson method. Essentially, the Davidson method calculates the best approximation to the wavefunction from a given set of vectors. Through extending this set by one vector (the update vector) in every iteration convergence is guaranteed. The speed of convergence depends on how appropriate the update vectors are. However, Sleijpen and van der Vorst [35] realised that if the method was applied exactly as suggested by Davidson it would never converge. A detailed analysis allowed for improvements of the method and the application of these improvements in quantum chemistry is discussed in chapter 6.

References

[1] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[2] P. E. M. Siegbahn, J. Chem. Phys. **72**, 1647 (1980).

[3] V. R. Saunders and J. H. v. Lenthe, Mol. Phys. **48**, 923 (1983).

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

[5] E. Schrödinger, Ann. Phys. **79**, 361 (1926).

[6] E. Schrödinger, Ann. Phys. **79**, 734 (1926).

[7] M. Born and R. J. Oppenheimer, Ann. Phys. **84**, 457 (1927).

[8] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[9] J. C. Slater, Phys. Rev. **35**, 210 (1930).

[10] W. Pauli Jr., Z. Physik **31**, 765 (1925).

[11] V. Fock, Z. Physik **61**, 126 (1930).

[12] D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

[13] I. Shavitt, in: "Methods of Electronic Structure Theory" (Plenum Press, New York, 1977).

[14] G. C. Groenenboom, "Novel Approaches to the Calculation of the Electronic Structure and Dynamics of Excited States", thesis, Technische Universiteit Eindhoven (1991).

[15] H. J. Silverstone and O. Sinanoglu, J. Chem. Phys. **44**, 1899 (1966).

[16] C. Hollister and O. Sinanoglu, J. Am. Chem. Soc. **88**, 13 (1966).

[17] O. Sinanoglu, Proc. Nat. Acad. Sci. USA **47**, 1217 (1961).

[18] R. J. Buenker, S. D. Peyerimhoff and W. Butscher, Mol. Phys. **35**, 771 (1978).

[19] P. E. M. Siegbahn, "The externally contracted CI method" in: "Semin. Comput. Methods Quantum Chem., 5th", 291 (Garching, Fed. Rep. Ger., 1981).

[20] P. E. M. Siegbahn, Stud. Phys. Theor. Chem. **21**, 65 (1982).

[21] J. P. Finley, R. K. Chaudhuri and K. F. Freed, J. Chem. Phys. **103**, 4990 (1995).

[22] G. Hose, "Multi-reference-state Rayleigh-Schrödinger perturbation theory" in: "Many-body methods in quantum chemistry", Lecture Notes in Chemistry 52, ed. U. Kaldor (Springer-Verlag, Berlin, 1989).

[23] U. Kaldor, J. Chem. Phys. **81**, 2406 (1984).

[24] P. E. M. Siegbahn, Int. J. Quant. Chem. **18**, 1229 (1980).

[25] J. Werner and E. A. Reinsch, J. Chem. Phys. **76**, 3144 (1982).

[26] R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).

[27] J. A. Pople, J. S. Binkley and R. Seeger, Int. J. Quant. Chem. Symp. **10**, 1 (1976).

[28] J. A. Pople, "Theoretical models for chemistry" in: "Energy, Structure and Reactivity", 51 (Boulder, 1972).

[29] J. H. Choi and D. W. Smith, J. Chem. Phys. **45**, 4425 (1966).

[30] I. Shavitt and L. T. Redmon, J. Chem. Phys. **73**, 5711 (1980).

[31] K. Wolinski and P. Pulay, J. Chem. Phys **90**, 3647 (1989).

[32] B. Huron, J. P. Malrieu and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[33] E. R. Davidson, Chem. Phys. Lett. **241**, 432 (1995).

[34] V. R. Saunders and M. F. Guest, ATMOL (Utrecht University, Daresbury, 1974).

[35] G. L. G. Sleijpen and H. A. van der Vorst, Preprint: 856, Dept. Math., Utrecht University (1994).

Chapter 2

Multi-reference Møller-Plesset perturbation theory

Implementation to third order and applications

Abstract

The generalised Møller-Plesset perturbation theory for open-shell and multi-configurational reference functions as formulated by Wolinski, Sellers and Pulay is implemented in a multi-reference CI program that allows general reference functions. Extension to third order is straightforward at the cost of one CI iteration. The method is applied to several small model problems, i.e. the singlet-triplet splitting in methylene, the dispersion energy of helium dimer, and the dissociation energy of N_2 and O_2. Also the method was applied to a larger system, i.e. the transition energy $1A_g \rightarrow 1B_u$ in 1,1'-bicyclohexylidene was calculated. The results depend little on the projection operators chosen in the definition of the zeroth-order Hamiltonian. The dissociation energies obtained are fairly good. In most calculations, MRMP3 is closer to higher order methods [Multi-Reference CI (MRCI), MRCI with Davidsons correction, and MRCEPA(0)] than MRMP2. However, in the calculation of the dissociation energy of dioxygen the MRMP2 results are closer to the higher order methods than those of the MRMP3 results. The method is not free of problems, however, as we have found divergences on the potential energy surface of N_2 at the MRMP3 level. For the calculation of dispersion energies, which depend on the reordering among the reference states, the MRMP2/3 method brings no improvement over the single-reference approach, although this may be cured at higher orders in perturbation theory. The method was also applied to a larger system, i.e. the $1A_g \rightarrow 1B_u$ transition energy in 1,1'-bicyclohexylidene was calculated. Size consistency corrected MRCI, MRCEPA(0) and MRMP3 all agree closely on the transition energy, which is, however, still too high compared to the experimental value.

2.1 Introduction

The calculation of potential energy surfaces is perhaps the most important application of quantum chemistry. Since during a reaction bonds are broken and reformed, a single-determinant description is not of consistent quality. The commonly used Restricted Hartree Fock (RHF) method breaks down for radical bond breaking while the Unrestricted Hartree Fock (UHF) method leads to spin contamination. The Multi-Configurational SCF (MCSCF) approach remedies these shortcomings by including all configurations that are important in various regions of the potential energy, creating a qualitatively correct zeroth-order wavefunction. Though the use of a multi-reference function is a useful prerequisite for obtaining a consistent description and it ensures a proper description of the non-dynamic correlation energy, it is not sufficient since in general the dynamic correlation energy, which is essentially impossible to include to a significant extent in MCSCF, changes with the molecular geometry [1-3].

The principal ways to describe the dynamic correlation energy are the variational Configuration Interaction (CI) technique, the Coupled Cluster (CC) approach, and the use of Perturbation Theory (PT). Whereas the former methods offer a more complete description of the correlation, low orders of PT are much more efficient. The Møller-Plesset form of PT shares with the more expensive CC techniques the desirable property of size consistency [4,5].

Perturbation theory may be generalized to the multi-reference case in different ways, for example using effective Hamiltonians, using different types of H_0 using, different expansions for the wavefunction, etc. [5-22]. In this paper we explore the performance of the generalisation of Møller-Plesset Perturbation Theory (MPPT) [23] as formulated by Wolinski *et al.* [24,25]. Different approaches to multi-reference MPPT are the CASPT approach implemented in MOLCAS [26-28] and the multi-configurational MPPT approach in MOLPRO [29] . We discuss some aspects of the implementation of the theory. In addition we consider the implications of various choices possible in the zeroth order Hamiltonian. We study a set of problems ranging from the breaking of double and triple bonds for closed and open shell systems to the dispersion energy in Van der Waals systems, and to the calculation of transition energies in large organic molecules. A number of these cases have been studied by Werner [29] also. However for the diatomic molecules we concentrated on calculating the potential energy surface instead of the spectroscopic constants.

2.2 Theory

The theory corresponds closely to the closed shell Møller-Plesset theory and will indeed reduce to this approach in the single-reference closed shell case. We will therefore recapitulate the single-reference closed shell approach first. As usual in perturbation theory, the total Hamiltonian is split into a zeroth order part and a perturbation:

$$H = H_0 + V \tag{1}$$

H_0 in Møller-Plesset Perturbation Theory is the closed-shell Fock operator

$$H_0 = F = \sum_i f_i \tag{2}$$

where f is the usual one-electron closed-shell Fock operator, the elements of which are obtained from the one-electron density matrix P and the one- and two-electron integrals as

$$f_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}\left[(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\lambda|\sigma\nu)\right] \tag{3}$$

The zeroth order energy is the sum of orbital energies

$$E_0 = E^{(0)} = \langle \Psi_0 | H_0 | \Psi_0 \rangle = \sum_i \varepsilon_i \tag{4}$$

and the first, second and third order energies are given by

$$E_1 = E^{(0)} + E^{(1)} = \langle \Psi_0 | H | \Psi_0 \rangle = E_{ref} \tag{5}$$

were E_{ref} is the energy of the Hartree-Fock determinant and

$$\begin{aligned} E_2 &= E_1 + \langle \Psi_0 | V | \Psi_1 \rangle \\ &= E_1 + \langle \Psi_0 | H | \Psi_1 \rangle \end{aligned} \tag{6}$$

$$\begin{aligned} E_3 &= \langle \Psi_0 + \Psi_1 | H - E_1 | \Psi_0 + \Psi_1 \rangle + E_1 \\ &= \langle \Psi_0 + \Psi_1 | H - E_{ref} | \Psi_0 + \Psi_1 \rangle \end{aligned} \tag{7}$$

The first order wavefunction is formally obtained by

$$\Psi_1 = -(H_0 - E_0)^{-1} QV\Psi_0 \tag{8}$$

where $Q = 1 - | \psi_0 \rangle \langle \psi_0 |$ to guarantee that Ψ_1 will have no component in Ψ_0. Ψ_1 is computed by expanding it formally in the space of all determinants obtained from the reference determinant by substituting virtual orbitals for occupied ones. If canonical molecular orbitals are used, the zeroth order Hamiltonian (the Fock operator) is diagonal in the configuration basis. Only doubly substituted determinants can contribute to Ψ_1 and the matrix inversion implied in Eq. (8) becomes trivial. Equation (7) shows that calculating the MP3 energy requires a matrix vector product with the Hamilton matrix, which costs as much as a single iteration in a CI calculation.

In the general case of a multi-configurational or spin restricted open shell RHF wavefunction, one may still define a Fock operator through the first order density matrix as in Eq. (2) [30]. By comparing this (spin-averaged) zeroth-order Hamiltonian with the unrestricted Fock matrix, it is evident that H_0 defined by Eqs. (2,3) has some deficiency for high-spin open-shell systems, and there has been a number of attempts to improve it for these cases [9,12,18,31]. We will not comment on perturbation theories aimed at restricted Hartree-Fock reference only [32-34], as in our opinion an open-shell PT should be the limiting case of a multi-reference theory. The most promising of the modifications suggested for H_0 appears to be the one by Andersson [31], because it allows to remove the often low lying singly and doubly excited states in the orthogonal complement space to the reference function from the MP2 and MP3 equations.

A difficulty in all open-shell and multi-reference perturbation theories is that the zeroth-order wavefunction Ψ_0 is not necessarily an eigenfunction of H_0 defined by Fock operators. To restore the eigenfunction property we can define a projector on Ψ_0 as

$$P_0 = | \Psi_0 \rangle \langle \Psi_0 | \tag{9}$$

Using this projector, a plausible choice for H_0 is

$$H_0 = P_0 F P_0 + (1 - P_0) F (1 - P_0) \qquad (10)$$

Unfortunately, this zeroth-order Hamiltonian does not lead to a practically useful general PT since it is not block-diagonal in the excitation levels, and thus the first-order wavefunction with this H_0 contains not only doubly but also triply, quadruply, etc. substituted configurations. To escape this problem H_0 may be defined as [25]

$$H_0 = P_0 F P_0 + P_1 F P_1 + P_2 F P_2 + \cdots \qquad (11)$$

where P_1 is a projector on the singly substituted states, P_2 on the doubly substituted ones, etc. We denote this as model 1. It is not trivial to define substitution levels for a multi-configurational reference state. Wolinski and Pulay [24] defined them by applying repeatedly the single substitution operator manifold (E_{ij}-E_{ji}) to Ψ_0, and orthogonalising the resulting set to any lower order substituted subspace. This method allows the consistent definition of arbitrary high orders of PT, although the practical usefulness of high orders is doubtful. This formulation leads naturally to the internally contracted [35] form of CI theory, with the advantage that the configuration list is dramatically shortened, although at the expense of a more complicated structure of the matrix elements.

In most early CI programs, single and double substitutions were treated together. In this vein Anderson *et al.* [26] have defined their zeroth-order Hamiltonian as

$$H_0 = P_0 F P_0 + P_{1+2} F P_{1+2} \qquad (12)$$

where the part of the Hamiltonian that is needed only in fourth and higher orders (i.e. the orthogonal complement of the reference function, and three-fold and higher excitations) is left out. For historical reasons we call this model 3; its projection operators are defined as model 1, i.e. using a contracted formalism. Although MRMP is not strictly size consistent in various implementations, test calculations on Be-dimer ($R_{\text{Be-Be}}$ 200 bohr, basis 6-311g*, CAS reference in 3σ, $4\sigma^*$, 5σ, $6\sigma^*$, $1\pi_x$, $2\pi_x^*$, $1\pi_y$, $2\pi_y^*$ orbitals) yielded a size consistency error of less than 0.1 μH for model 1, whereas model 3 was found to be not strictly size consistent. A number of applications have been reported by this method at the second order level by Roos and co-workers [36-45].

An alternative definition of the projectors to the space of substituted configurations is to consider the substitution operators applied to each reference configuration separately, not to the full reference wavefunction. This uncontracted approach leads to much larger substituted subspaces, the ratio to the contracted approach being roughly the number of configurations in the reference space. This is the principal disadvantage of this method, since it is not applicable to long reference expansions, say 10^5 or 10^6 configurations. Using an H_0 that is formally identical with Eq. (12), except that the projector P_{1+2} projects to the space spanned by applying the substitution operators to each reference configuration, Murphy and Messmer [46] implemented this form of multi-configurational PT first. We call this model 0. As model 1 and model 3 differ only in some small coupling terms between singles and doubles and are numerically almost equivalent, we will concentrate on models 0 and 1.

Finally, Werner [29] defined a compromise method that uses uncontracted functions for the internal and singly external configuration spaces, and internally contracted states for the rest of the space. We have not studied this possibility.

Unlike in the closed-shell case, it is not possible in general to define canonical orbitals that are eigenfunctions of the Fock operator, Eq. (3), particularly if a limited CI expansion is used (not a Complete Active Space [47-49]). This means that, unless the Fock operator is truncated to a diagonal form, the first-order wavefunction must be determined using the full formula Eq. (8). Although this looks formidable at first, the first order wavefunction can be determined by an efficient iterative procedure, since the matrix H_0 is very sparse, just as in the case of closed-shell Møller-Plesset theory with non-canonical (localised) orbitals [50]. The use of a non-diagonal H_0 is essential for the formulation of the theory in an orbital-invariant manner, and it usually also leads to better results than a zeroth-order Hamiltonian truncated to diagonal form [51].

2.3 Implementation

We implemented the multi-reference Møller-Plesset formalism as an option in the Direct CI program of Saunders and van Lenthe [52]. This program uses a basis of spin-adapted Configuration State Functions (CSF). Both the projected formalism using the CSFs as a basis, and the contracted formalism using the contracted reference state and the contracted excited states as a basis were implemented.

The singly and doubly excited states needed in both approaches are obtained by applying the substitution operators to the MCSCF reference function [24], and subsequent orthonormalisation of the resulting states. This is a deviation from the internally contracted CI methods [53], where possible non-orthogonalities can be handled by diagonalisation. The orthogonalisations are only required within very restricted subspaces, i.e. configuration spaces having the same number of holes in the doubly occupied and the same number of particles in the virtual spaces of the multi-configurational reference function. In the case of model 0 no orthogonalisation is required except for excitations completely within the active orbitals of the MCSCF wavefunction [46]. The orthogonalisation may be performed through diagonalising the overlap matrix as is done by Andersson [26] or through a Löwdin orthogonalisation [54] as done by Werner [29], but we choose the Modified Gramm-Schmidt (MGS) method. A slight problem with MGS is that it is numerically less accurate [55]. Therefore, if a vector has been orthogonalised to a set of orthogonal vectors using MGS the orthogonality should be checked. If needed, a second iteration on the vector may be performed to achieve orthogonality within the desired accuracy.

In both the projected and contracted form Eq. (8) is solved using the Jacobi method [56,57] which involves evaluating

$$C_1^{n+1} = C_1^n - D^{-1}\left[\left(H - E_{ref}\right)C_0 + \left(H_0 - E_0\right)C_1^n\right] \tag{13}$$

each iteration. Here D is the diagonal of the Fock operator H_0 possibly adjusted with a level shifter. The part between square brackets is the residue, the norm of which may be used to monitor the convergence of the iteration process. The HC_0 term in Eq. (13) is obtained from the first iteration of the Direct CI program using the MCSCF vector C_0 as CI vector. Convergence is accelerated using a standard DIIS procedure [58,59]. In this method only a matrix-vector multiplication with the zeroth order Hamiltonian is required in each iteration. The efficiency of this multiplication is important to the overall efficiency and depends on the formalism (projected or contracted) used. Therefore the details involved in the multiplication which are important to the efficiency will be discussed in the next two paragraphs for the projected and contracted formalism respectively.

In the projected form the matrix-vector multiplication uses the CSF basis of the direct CI program [52] (for notation we refer to [52] also). The standard symbolic matrix elements of this program are used to evaluate the matrix elements of H_0. Many of the routines originally written for handling the CI Hamiltonian may be reused as such or in a slightly modified form. Basically, these routines use internal matrix elements that are precomputed, the external part is generated on the fly. Only the original precomputed elements for H need to be replaced by the elements for H_0. To generate these elements the projection operators have to be incorporated in the Fock matrix elements. In CSF basis the projector on the reference function may be written as

$$\left|\Psi_0\right\rangle\left\langle\Psi_0\right| = \sum_{ST\sigma\tau}\left|\Psi^{S\sigma}\right\rangle C^{S\sigma}C^{T\tau}\left\langle\Psi^{T\tau}\right| \tag{14}$$

where S and T label the internal occupation patterns, and σ and τ label the spin patterns in the internal MOs. A projector on an excited state (here two electrons are excited from the internal space to the external space) becomes

$$\left|E_{ij}^{ab}\Psi_0\right\rangle\left\langle E_{ij}^{ab}\Psi_0\right| = \sum_{ST\sigma\tau}\left|E_{ij}^{ab}\Psi^{S\sigma}\right\rangle C^{S\sigma}C^{T\tau}\left\langle E_{ij}^{ab}\Psi^{T\tau}\right| \tag{15}$$

Taking into account that the Fock operator is a one-electron operator the internal occupation pattern may differ in one place to allow a CSF $E_{ij}^{ab}\Psi^{V\upsilon}$ to contribute to a CSF $E_{ij}^{ab}\Psi^{S\sigma}$. This leads to

$$\begin{aligned}&\left|E_{ij}^{ab}\Psi_0\right\rangle\left\langle E_{ij}^{ab}\Psi_0\right|F\left|E_{ij}^{ab}\Psi_0\right\rangle\left\langle E_{ij}^{ab}\Psi_0\right| =\\&\sum_{\substack{ST\sigma\tau\\UV\rho\upsilon}}\left|E_{ij}^{ab}\Psi^{S\sigma}\right\rangle C^{S\sigma}C^{T\tau}\left\langle E_{ij}^{ab}\Psi^{T\tau}\right|F\left|E_{ij}^{ab}\Psi^{U\rho}\right\rangle C^{U\rho}C^{V\upsilon}\left\langle E_{ij}^{ab}\Psi^{V\upsilon}\right|\end{aligned} \tag{16}$$

which results in a contribution to the matrix-vector product with H_0 of the form

$$Z_{ab}^{S\sigma} = \sum_{\upsilon} B_{ab}^{S\sigma V\upsilon} FC_{ab}^{V\upsilon} + \sum_{\overline{\upsilon}} B_{ab}^{S\sigma\overline{V}\overline{\upsilon}} FC_{ab}^{\overline{V}\overline{\upsilon}} \tag{17}$$

where B represents coupling coefficients (the bar over V and υ indicates triplet coupling of the electrons in the external MOs).

On the other hand the excitation operator may differ for one-electron to allow a CSF $E_{ij}^{ac}\Psi^{U\rho}$ to contribute to a CSF $E_{ij}^{ab}\Psi^{S\sigma}$. This leads to

$$\begin{gathered}\left|E_{ij}^{ab}\Psi_0\right\rangle\left\langle E_{ij}^{ab}\Psi_0\right|F\left|E_{ij}^{ac}\Psi_0\right\rangle\left\langle E_{ij}^{ac}\Psi_0\right| = \\ \sum_{STU\sigma\tau\rho}\left|E_{ij}^{ab}\Psi^{S\sigma}\right\rangle C^{S\sigma}C^{T\tau}\left\langle E_{ij}^{ab}\Psi^{T\tau}\right|F\left|E_{ij}^{ac}\Psi^{T\tau}\right\rangle C^{T\tau}C^{U\rho}\left\langle E_{ij}^{ac}\Psi^{U\rho}\right|\end{gathered} \tag{18}$$

which results in a contribution to the matrix-vector product with H_0 of the form

$$Z_{ab}^{S\sigma} = \sum_{\rho c} B_{ac}^{S\sigma U\rho}F_{bc}C_{ac}^{U\rho} + \sum_{\bar{\rho} c} B_{ac}^{S\sigma\bar{U}\bar{\rho}}F_{bc}C_{ac}^{\bar{U}\bar{\rho}} \tag{19}$$

This means that the Fock matrix transformed to CSF basis behaves as a one-electron operator among the external orbitals in the sense that it only couples CSFs differing one orbital in the external space. However, it couples all internal occupation patterns and in that sense it behaves similar to a multiple electron operator. In practice this means that the integrals needed in the matrix-vector multiplication may be classified as shown in table **I**

Table I: Classification of zeroth-order hamiltonian interactions

CSF interactions	**Integral type**
Vacuum-vacuum Doublet-Doublet Singlet-Singlet Triplet-Triplet	[ij|kl] and [ij]
Doublet-Vacuum Singlet-Doublet Triplet-Doublet	[ia]
Doublet-Doublet Singlet-Singlet Triplet-Triplet	[ij|ab], [ia|jb] and [ab]

Comparing this integral classification with the one for the total Hamiltonian [52] it is clear that H_0 does not behave as a true one-electron operator in this basis. The appearance of two-electron integrals leads to a rather expensive matrix-vector multiplication. In less detail this was pointed out by Roos *et al.* [30] also.

In the contracted formalism the Fock operator behaves as a true one-electron operator, but the contracted states differ from the configurations used by the original CI code in such a way that special multiplicators have to be written. However, these multiplicators may be derived from the CSF-based multiplicators. Following the MP2, MP3 merely constitutes a single iteration of the Direct CI process, using Ψ_1 as the CI vector. Because the CI part of the program still works in CSF space, Ψ_1 must be transformed back to the CSF basis. Therefore the transformations from contracted state basis to CSF basis and vice versa are implemented. These transformations together with the CI matrix multiplication effectively offer a single CI iteration in the contracted space. As a spin-off, this allows one to implement a contracted CI approach also.

Our contracted formalism differs slightly from the formalisms implemented by Werner and Andersson. Andersson [26] implemented an H_0 such that no internally excited states will contribute to Ψ_1, whereas our H_0 may result in contributions in the entire first order interacting space. Werner [29] has for efficiency reasons implemented an approach in which the internal and semi-internal states are left uncontracted. Only the external states are internally contracted. In our approach all states are contracted.

The projected and the contracted implementation differ in various ways. First, the projected approach is less work to implement than the contracted one because many of the CSF based routines available may be reused. Although this may seem attractive there is a runtime penalty in using the projected approach. This is shown in table **II** were we compare the timings of the projected and contracted approach for one of our test examples, the dissociation of N_2 in a 6-311G+g(d,p) basis set, using a 6 orbital - 6 electron CASSCF wavefunction as reference (a comparison of the efficiency of MRCI and MRMP is given in the 1,1’-bicyclohexylidene example). This produces a total of 219042 symmetry allowed CSFs and 6964 excited states for model 1. All these calculations are performed on an Apollo DN10000 workstation, a rather modest machine by today’s standards.

Table II. Timing comparison for MRMP models and MRCI

	Symbolic* (seconds)	Time/Iteration (seconds)	# Iterations	Total (seconds)
MRSDCI	210	270	12	3500
Proj. MP2-model 1	320	110	31	3700
Proj. MP2-model 0	270	73	10	1000
Contr. MP2-model 1	220	1.5	10	235

* The symbolic includes the calculation of the MRCI and MRMP coupling constants, and the generation and orthogonalisation of excited states.

The projected MP2 is more efficient than the CI because H_0 contains fewer interactions than H. Specifically, integral types having three or four external orbital indices (e.g. [ab|cd]) are absent in H_0. Furthermore, projected MP2 model 0 is more efficient than model 1. This is partly because there are no projectors between the doubles and the singlets and triplets in model 0 so there are fewer interactions among those CSFs. Also, the model 0 H_0 is more diagonal dominant thus facilitating the convergence of the linear system solver. However, the contracted MP2 is much more efficient than the projected MP2. Because in the contracted formalism only the true first order interacting space of the reference function is taken into account the contracted space is much smaller than the CSF space. For example, the CSF reference space in this case even holds six fold excited states whereas in the contracted formalism these states were removed while transforming to the contracted basis.

In an MP3 calculation however the contracted formalism has the disadvantage that the contracted Ψ_1 must be transformed to CSF base which is not needed in the projected approach. In practice, this transformation costs a small fraction of the total CPU time needed to calculate the MP3 energy.

Concluding, the comparison between projected MP and contracted MP shows that the contracted approach is the method of choice. Thus all results in the following sections are calculated using the contracted formalism.

2.4 Applications

We show the application of the contracted MRMP method to various problems. We concentrate on comparing the various models. We compare the Møller-Plesset results to MRCI [52], with and without a multi-reference Davidson size consistency correction and MRCEPA(0) [60]. The Davidson correction [61] was calculated using the overlap of the reference function and the CI wavefunction as c_0 and the difference of the energies of the two wavefunctions as correlation energy [62].

As a first application we studied the singlet-triplet splitting of methylene to asses the amount of correlation energy recovered in the multi-reference perturbation approach. The other applications, except for the last one, are concerned with various aspects of potential energy surfaces. The dispersion coefficient of helium dimer is discussed as an application to pure dynamic correlation interactions. The potential energy curves of N_2 and O_2 are discussed as well as the dissociation energies of N_2, O_2 and ethene. Finally, the transition energy in 1,1'-bicyclohexylidene between the A_g and B_u state is calculated to demonstrate the performance in large molecules.

2.4.1 Singlet-triplet splitting of methylene

The energy separation between the $^3B_1$ ground state and the $^1A_1$ first excited state of CH_2 was calculated by Bauschlicher and Taylor [63] as a calibration of other correlation approaches. We therefore use their DZP basis and their geometry. The results for both models 1 and 3 are given in table **III**.

For the reference functions a 2-configuration MCSCF function for the singlet and the associated single-configuration function for a triplet, and a CASSCF function where all electrons are distributed over the 5 active orbitals were chosen.

Table III Singlet-triplet splitting for methylene

	Singlet (hartree)	Triplet (hartree)	Delta E (kcal/mole)	Error (kcal/mole)
SCF	-38.886296	-38.927947	26.14	14.17
SCF/SDCI	-39.018284	-39.041602	14.63	2.66
2-MCSCF	-38.907659	-38.927947	12.73	0.76
2-MP2-1	-39.005802	-39.031047	15.84	3.87
2-MP3-1	-39.021856	-39.042263	12.81	0.83
2-MP2-3	-39.005858	-39.031192	15.90	3.93
2-MP3-3	-39.021953	-39.042350	12.80	0.83
2-CI	-39.022155	-39.041602	12.20	0.23
2-CI+Dav	-39.027748	-39.046911	12.03	0.05
2-MRCEPA(0)	-39.028373	-39.047386	11.93	-0.04
CASSCF	-38.945528	-38.965954	12.82	0.85
CASMP2-1	-39.013013	-39.037543	15.39	3.42
CASMP3-1	-39.023437	-39.043826	12.79	0.82
CASMP2-3	-39.013245	-39.037852	15.44	3.47
CASMP3-3	-39.023633	-39.043901	12.72	0.75
CASCI	-39.025803	-39.044929	12.00	0.03
CASCI+Dav.	-39.028541	-39.047389	11.83	-0.14
CASCEPA(0)	-39.028030	-39.046816	11.79	-0.18
Full CI	-39.027181	-39.046260	11.97	0.00

Firstly, one notices that the single-reference approach (SCF/SDCI) is inadequate and therefore multi-reference techniques are essential. From the perturbation approaches model 1 and model 3 yield essentially the same results. Second order perturbation theory is slightly worse than the two-configuration MCSCF/ROHF level, due to the inadequacy of the spin-averaged Fock matrix as H_0 for the triplet state. This agrees with the results of other groups [27]. The modified H_0 recommended by Andersson [31] performs significantly better in this case. Third order perturbation theory is an improvement over the second order but the resulting singlet-triplet splitting is no more accurate than that obtained from the MCSCF/ROHF calculations. The MRCI calculation with the Davidson correction and the MRCEPA(0) calculation, both starting with the 2-configuration MCSCF/ROHF wavefunctions, agree with the full CI results for the splitting, although their energy goes slightly below the full CI result.

In [29] Werner gives a comprehensive comparison of CAS reference approaches to this case. Our CASMP results are in agreement with Werner's overview.

2.4.2 The C_6 dispersion coefficient of the He-dimer

The attractive component of the weak interaction energy of the He-dimer is caused solely by the correlation energy. Whereas the single-configuration SCF wavefunction for the system, does describe a proper dissociation to separate He atoms, no attractive minimum is found in the Hartree-Fock approximation. The C_6 van der Waals coefficient, which corresponds to the dominant induced-dipole-induced-dipole term in the multipole expansion, is calculated by fitting the C_6 and the C_8 coefficients to energies calculated at 10, 20 and 2000 Bohr inter-nuclear distance. We used the [6s3p2d] C1D2 basis set [64,65]. The full CI calculations were performed using the program of Harrison [66] .

For this property the normal MP2 method fails badly. All approaches using a single-configuration as reference function severely underestimate the dispersion energy. To obtain the dispersion energy between correlated He atoms triple excitations have to be included [67]. In the multi-reference calculations this is achieved using the direct product of a 5-configuration double excitation wavefunction for each He atom as the reference space.

The results in table **IV** show that the multi-reference CI calculations indeed reach almost the basis set limit for C_6. Furthermore, the MRCI and the Contracted MRCI (MRCCI) yield the same results. This shows that all essential contributions to the dispersion energy are contained in the contracted multi-reference singles-doubles space. However, the multi-reference Møller-Plesset theory performs no better than the single-reference approach regardless of the model used, so the multi-configuration character of the reference function does not yield improved results. Therefore, higher order perturbation theory is required to obtain a proper C_6 coefficient.

Table IV C_6 Dispersion coefficient for He_2.

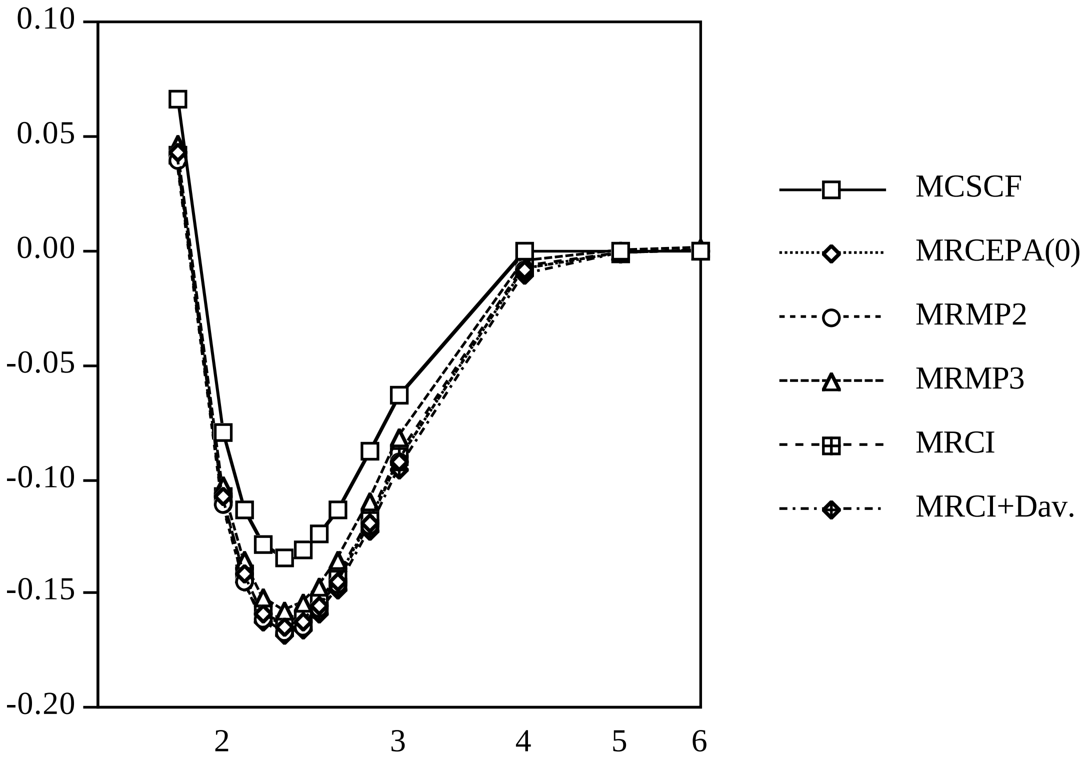
	C_6 (au)		C_6 (au)
MP2	1.11	MRMP2	1.05
MP3	1.32	MRMP3	1.32
SDCI	1.32	MRCI	1.44
SDCI+Dav.	1.39	MRCI+Dav.	1.44
CEPA-0	1.39	MRCEPA(0)	1.44
		MRCCI	1.44
Full CI	1.45	Full CI	1.45

2.4.3 Dissociation of O_2 and N_2

Any performance study of a generalised perturbation method should at least consider an open shell system. The dissociation of the triplet ground state O_2 molecule into two triplet ground state O atoms is a typical example of such systems. Therefore, we calculated the potential energy surface of the molecule using MRMP2 and MRMP3 and compare the results to MRCI, MRCI+Davidson, and MRCEPA(0) results. The calculations were performed within D_{2h} symmetry B_{1u} using a DZP-basis with spherical harmonic d-functions (exponent 1.2) (see [68] for the DZ part of the basis). A CAS in the 5σ, $6\sigma^*$, $1\pi_x$, $2\pi_x^*$, $1\pi_y$, and $2\pi_y^*$ orbitals is used for the reference, and the 1S orbitals are frozen at the MCSCF level.

The results are shown in figure **I**. The MCSCF curve has a maximum at 4.0 bohr. This causes the strong curvature around that point. This maximum has been found in other studies [69-73] also, and there has been much discussion about its nature. Because perturbation theory has been plagued by convergence problems we feared for divergences at this maximum. However, the MRMP2 and MRMP3 behave nicely at that point. Further, the plot shows that correlating methods improve the bonding energy considerably. All correlating methods seem to give comparable results.

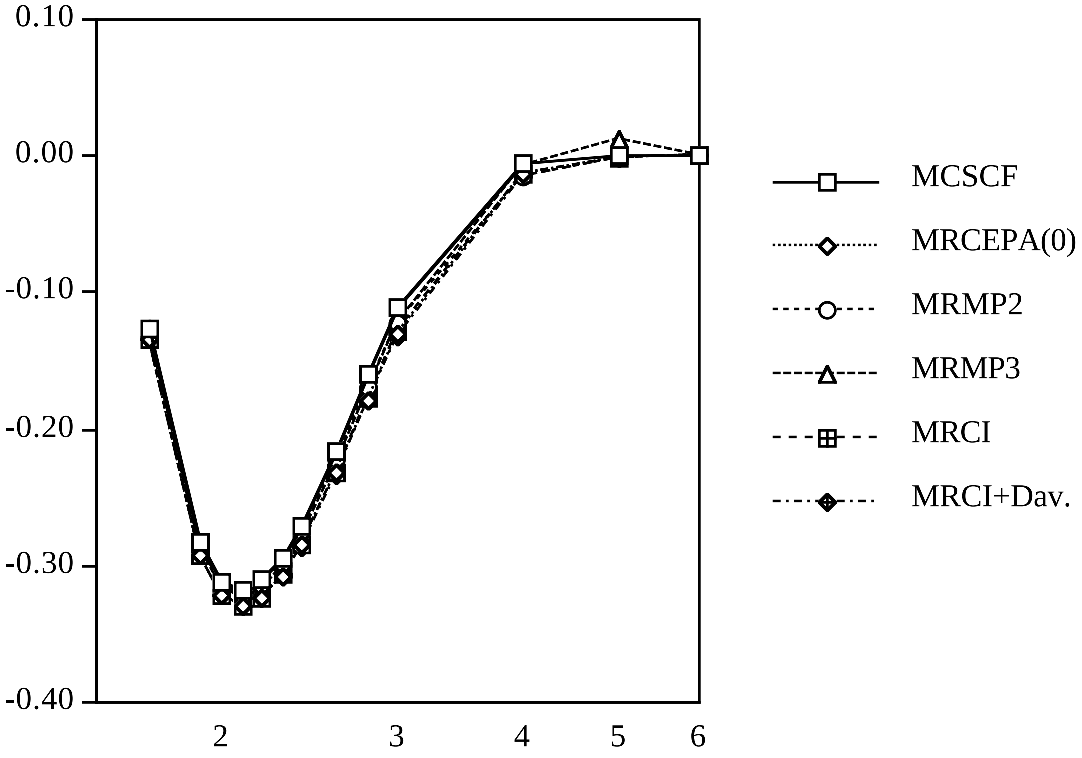
Figure I: Bonding energy of O_2 in DZP-basis plotted against the internuclear distance



The bonding energy is plotted vertically in hartree. The internuclear distance *R* is plotted horizontally in bohr on a logarithmic scale. The horizontal scale starts at 1.5 bohr. The distance at infinity R_{inf} is 200 bohr.

As a reference molecule to compare the O_2 calculations to we chose N_2. This singlet molecule dissociates into two quartet atoms, breaking a triple bond without further complications. The calculations on N_2 were performed using the 6-311g* [74] basis, and a CAS in the 5σ, $6\sigma^*$, $1\pi_x$, $2\pi_x{}^*$, $1\pi_y$, and $2\pi_y{}^*$ orbitals as reference space. The 1S orbitals are frozen at the MCSCF level and not correlated. The bonding energies of N_2 are plotted in figure **II**.

Figure II: Bonding energy of N_2 in 6-311g* basis plotted against the internuclear distance



The bonding energy is plotted vertically in hartree. The internuclear distance R is plotted horizontally in bohr on a logarithmic scale. The horizontal scale starts at 1.5 bohr. The distance at infinity R_{inf} is 200 bohr.

In contrast to the O_2 plot the N_2 curves show that correlating methods do not seem to be so important for the bonding energy. Again all correlating methods seem to give comparable results, except for the MRMP3 at 5.0 Bohr which shows a “peak”. The “peak” in this curve suggests that the perturbation series is divergent at that geometry. In a forthcoming article we will consider a test for the convergence behaviour of perturbation series at dubious geometries [75] and study this case in more detail.

The dissociation energies of O_2 and N_2 obtained from the calculations are presented in tables **V** and **VI** respectively.

Table V Dissociation energy of O_2

	$E(R_{eq})$ (hartree)	$E(R_{inf})$ (hartree)	Delta E (kcal/mole)
MCSCF	-149.732663	-149.598813	83.99
MRCEPA(0)	-150.004409	-149.839666	103.38
MRMP2	-149.991294	-149.824555	104.63
MRMP3	-149.992384	-149.835002	98.76
MRCI	-149.990020	-149.825954	102.95
MRCI+Dav.	-150.007358	-149.839323	105.45
Experiment	n.a.	n.a.	114.13

The equilibrium distance R_{eq} is 2.3081604 bohr, and the infinite distance R_{inf} is 200 bohr. The experimental values were calculated from thermodynamic data [85] and MCSCF zero point vibration energies in the DZP basis set calculated with GAMESS-UK [82]

Table VI Dissociation energy of N_2

	$E(R_{eq})$ (hartree)	$E(R_{inf})$ (hartree)	Delta E (kcal/mole)
MCSCF	-109.107223	-108.789473	199.39
MRCEPA(0)	-109.314009	-108.984356	206.86
MRMP2	-109.288746	-108.968016	201.26
MRMP3	-109.305037	-108.981735	202.88
MRCI	-109.304247	-108.975167	206.50
MRCI+Dav.	-109.315312	-108.985580	206.91
Experiment	n.a.	n.a.	220.26

The equilibrium distance R_{eq} is 2.0929693 bohr, and the infinite distance R_{inf} is 200 bohr. The experimental values were calculated from thermodynamic data [85] and MCSCF zero point vibration energies in the DZP basis set calculated with GAMESS-UK [82]

In O_2 the dissociation energies calculated with MRCEPA(0), MRCI and MRMP2 are all quite close to each other. However the MRMP3 results in a dissociation energy that is to low. It underestimates the correlation energy in the molecule. In N_2 the MRCEPA(0) and MRCI calculation agree even more closely upon the dissociation energy than in O_2. But the perturbation theory finds dissociation energies to low both at second order and third order. Again it seems that the perturbation theory underestimates the correlation energy at the equilibrium geometry. This is in accordance with the conclusions by Werner [29] although his results are much more accurate due to the use of much larger basis sets.

Table VII Dissociation energy of ethene

	E CH_2 (hartree)	E C_2H_4 (hartree)	Delta-E (kcal/mol)
SCF	-38.927946	-78.050543	-122.15
MCSCF	-38.927946	-78.103024	-155.08
MRMP2	-39.031047	-78.327354	-166.46
MRMP3	-39.042263	-78.351993	-167.84
MRCI	-39.041700	-78.342357	-162.50
MRCI+Dav.	-39.047028	-78.361232	-167.66
MRCEPA(0)	-39.047507	-78.364167	-168.90
Experiment	n.a.	n.a.	-182.17

The experimental values were calculated from thermodynamic data [85] and MCSCF zero point vibration energies in the DZP basis set calculated with GAMESS-UK [82]

2.4.4 Dissociation of ethene

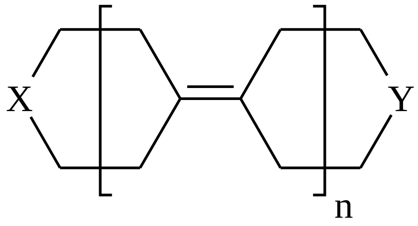
The dissociation of ethene into two $^3B_1$ fragments requires a 4 orbital 4 electron CASSCF reference function. We used the triplet basis set from the methylene calculations. The methylene geometries were also taken to be the $^3B_1$ geometries. The ethene geometry was optimised at the Hartree-Fock level using this basis set (R_{CC} 1.325 Å, R_{CH} 1.077Å, σ_{HCC} 121.56°). The results are given in table **VII**.

The CASSCF is quite an improvement with respect to the SCF. The MRMP2 reduces the difference with CI an order of magnitude further, at much less cost, and the MP3 is virtually indistinguishable from the Davidson corrected CI.

2.4.5 The A_g and B_u ground states of 1,1'-bicyclohexylidene

1,1'-bicyclohexylidene is the smallest member of the family of oligo(cyclohexylidenes), shown in figure **III**.

Figure III: Structure of oligo(cyclohexylidenes)



In oligo(cyclohexylidenes) the X and Y groups are CH_2 by default.

These molecules may be used as molecular building blocks for functional materials, in particular materials showing non-linear optical activity [76]. As a consequence there is quite some interest in the interpretation of the UV-spectrum. Recently, Hoogesteeger *et al.* [77] have studied this spectrum using results obtained with the MRD-CI [78-81] program of Buenker *et al.* included in GAMESS-UK [82]. However, due to the program limitations the size of the configuration space was severely restricted. The program presented in this article allows for much larger configuration spaces and therefore offers a way to validate the results by Hoogesteeger *et al.* In this context the energy of the $1A_g$ and the $1B_u$ states were recalculated at the MRMP2, MRMP3, MRCI, and MRCEPA(0) level with the MRCI program of Saunders *et al.* [52] using the geometry from Hoogesteeger and the same basis set, i.e. 6-31G [83]. Only the 5 most important configurations were selected as reference configurations as shown in table 4 in [77]. The results are shown in table **VIII**.

Tabel VIII: Comparison of MRD-CI energies for the $1A_g$ and $1B_u$ states of 1,1'-bicyclohexylidene to the MRMP2/3, MRCI and MRCEPA(0) energies.

	Energy $1A_g$ (hartree)	Energy $1B_u$ (hartree)	Transition energy (eV)
MRD-CI	-465.938378	-465.597163	9.28
MRMP2	-466.960643	-466.698324	7.14
MRMP3	-467.046972	-466.752633	8.01
MRCEPA(0)	-467.105770	-466.812577	7.98
MRCI	-466.762505	-466.446343	8.60
MRCI+Davidson	-466.951134	-466.644702	8.34
MRCI+Pople	-467.075290	-466.781485	8.00
Experiment	n.a.	n.a.	5.95

The Davidson corrections were calculated according to [61], and the Pople corrections were calculated according to [84].

Because Hoogesteeger *et al.* performed CI calculations using only 5431 CSFs for the $1A_g$ state and 3413 CSFs for the $1B_u$ state the energies were relatively high. MRCI, size consistency corrected MRCI, and MRCEPA(0) calculations involving 11446335 CSFs and 17997035 CSFs for the $1A_g$ and the $1B_u$ state gave improved estimates of the transition energy. However MRCI performed relatively poor compared to MRMP3, MRCEPA(0) and MRCI+Pople [84] suggesting that its lack of size consistency introduces a significant error. The MRMP2/3 calculations employed 1786521 and 1834052 contracted excited states for the $1A_g$ state and $1B_u$ state respectively, i.e. 6-10 fewer coefficients than in MRCI or MRCEPA(0). In spite of this, results agree closely with those of higher level calculations. We consider this encouraging. The MRMP2 transition energy is a bit low relative to MRCEPA(0) and size-consistency corrected MRCI but the MRMP3 agrees nicely with MRCEPA(0) and MRCI+Pople correction. Note that the MRMP2 and MRMP3 calculations on the $1A_g$ state took only 734 and 3440 seconds of CPU-time respectively compared to 42200 seconds for the MRCI and MRCEPA(0) calculations on a 90 MHz R8000 SGI PowerChallenge. Thus stressing the efficiency of the perturbation approaches. The basis set in this calculation was rather small valence double zeta, i.e. no polarization or diffuse functions were included. This helps to avoid complications from Rydberg states mixing in. However, it is also probably responsible for the 2 eV discrepancy in the transition energy, compared to experiment.

2.5 Conclusions

We have implemented a generalised MRMP method using both a contracted and an uncontracted formalism. The contracted formalism is much more efficient than the uncontracted formalism and is therefore the method of choice. Different zeroth order hamiltonians may be chosen through selecting different projections. Comparing model 1 and model 3 it should be noted that the results obtained are essentially the same. This indicates that model 1 is preferable because it is the computationally more efficient one and has smaller size consistency errors.

The applications show that the MRMP performs reasonably well on calculating the singlet-triplet splitting of methylene and the dissociation energies of O_2, N_2 and ethene. However, the MRMP3 dissociation energy for O_2 was less accurate than the MRMP2 result. Also, potential energy surfaces could be calculated quite well although a problem was found in the N_2 surface. However, MRMP2/3 fails in calculating the dispersion energy of the Helium dimer. Higher order perturbation theory is required for this property. In the calculation of the transition energy of 1,1’-bicyclohexylidene the size consistency corrected MRCI, the MRCEPA(0) and the MRMP3 agreed very closely on the transition energy. This means that the MRMP3 calculations led to significant improvements over the MRD-CI calculations by Hoogesteeger *et al.* at a much lesser cost than the MRCEPA(0) and the MRCI calculations.

Acknowledgements

We want to thank professor H.-J. Werner for sending copies of his manuscripts and helpful discussions.

The investigations were supported (in part) 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, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller, Karen Schuchardt, or Don Jones for further information.

References

[1] A. C. Wahl and G. Das, "Recent Progress in the *Ab Initio* Calculation of Potential Energy Surfaces for Chemical Reactions" in: "Proceedings on the Conference on Potential Energy Surfaces in Chemistry", (IBM Research Laboratory, San José, California, 1971).

[2] A. C. Wahl, in: "Energy, Structure and Reactivity", eds D. W. Smith and W. B. McRae (1973).

[3] A. C. Wahl and G. Das, in: "Methods of Electronic Structure Theory" (Plenum Press, New York, 1977).

[4] J. Goldstone, Proc. Roy. Soc. **A239**, 267 (1957).

[5] B. H. Brandow, Rev. Mod. Phys. **39**, 771 (1967).

[6] Z. Gershgorn and I. Shavitt, Int. J. Quant. Chem. **2,** 751 (1968).

[7] R. J. Cave and E. R. Davidson, J. Chem. Phys. **88,** 5770 (1988).

[8] R. J. Cave and E. R. Davidson, J. Chem. Phys. **89**, 6798 (1988).

[9] K. G. Dyall, J. Chem. Phys. **102,** 4909 (1995).

[10] J. P. Finley, R. K. Chaudhuri and K. F. Freed, J. Chem. Phys. **103,** 4990 (1995).

[11] K. Hirao, Chem. Phys. Lett. **196**, 397 (1992).

[12] K. Hirao, Chem. Phys. Lett. **201**, 59 (1993).

[13] G. Hose and U. Kaldor, J. Phys. B **12**, 3827 (1979).

[14] G. Hose, J. Chem. Phys. **84**, 4505 (1986).

[15] G. Hose, "Multi-reference-state Rayleigh-Schrödinger perturbation theory", in: "Many-body methods in quantum chemistry", Lecture Notes in Chemistry 52, ed. U. Kaldor (Springer-Verlag, Berlin, 1989).

[16] B. Huron, J. P. Malrieu and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[17] U. Kaldor, Phys. Rev. Lett. **31**, 1338 (1973).

[18] P. M. Kozlowski and E. R. Davidson, J. Chem. Phys. **100**, 3672 (1994).

[19] H. Nakano, J. Chem. Phys. **99,** 7983 (1993).

[20] I. Shavitt and L. T. Redmon, J. Chem. Phys. **73**, 5711 (1980).

[21] H. J. Silverstone and R. K. Moats, Phys. Rev. A **23**, 1645 (1981).

[22] A. Zaitsevskii and J.-P. Malrieu, Chem. Phys. Lett. **233**, 597 (1995).

[23] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[24] K. Wolinski, H. L. Sellers and P. Pulay, Chem. Phys. Lett. **140**, 225 (1987).

[25] K. Wolinski and P. Pulay, J. Chem. Phys **90**, 3647 (1989).

[26] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[27] K. Andersson, P.-Å. Malmqvist and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[28] B. O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995).

[29] H.-J. Werner, Mol. Phys. **89**, 645 (1996).

[30] B. O. Roos, P. Linse, P. E. M. Siegbahn and M. A. Blomberg, Chem. Phys. **66**, 197 (1982).

[31] K. Andersson, Theor. Chim. Acta. **91**, 31 (1995).

[32] I. Hubac and P. Carsky, Phys. Rev. A **22**, 2392 (1980).

[33] C. W. Murray and N. C. Handy, J. Chem. Phys. **97**, 6509 (1992).

[34] T. J. Lee and D. Jayatilaka, Chem. Phys. Lett. **201**, 1 (1993).

[35] W. Meyer, in: "Methods of Electronic Structure Theory" (Plenum, New York, 1977).

[36] K. Andersson and B. O. Roos, Chem. Phys. Lett. **191**, 507 (1992).

[37] K. Andersson and B. O. Roos, Int. J. Quantum Chem. **45**, 591 (1993).

[38] K. Andersson, B. O. Roos, P.-A. Malmqvist and P.-O. Widmark, Chem. Phys. Lett. **230,** 391 (1994).

[39] R. González-Luque, M. Merchán and B. O. Roos, Chem. Phys. **171**, 107 (1993).

[40] J. Lorentzon, P.-A. Malmqvist, M. P. Fülscher and B. O. Roos, Chim. Acta. in press, (1995).

[41] P.-Å. Malmqvist eds., "The RASSCF, RASSI, and CASPT2 methods used on small molecules of astrophysical interest.", in: "Molecules in the stellar environment", ed. U. G. Jorgensen, Lecture Notes in Physics 428 (Springer-Verlag, Berlin Heidelberg New York, 1994).

[42] R. Pou-Amérigo, M. Merchán, I. Nebot-Gil, P.-Å. Malmqvist and B. O. Roos, J. Chem. Phys. **101**, 4893 (1994).

[43] B. O. Roos, in: "New Challenges in Computational Quantum Chemistry," (Dept. of Chem. Phys. and Mat. Sci. Centre, Univ. of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands, 1994).

[44] B. O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot and M. Merchán, in: "Advances in Chemical Physics: New Methods in Computational Quantum Mechanics" (John Wiley & Sons, New York, 1995).

[45] B. O. Roos, L. Serrano-Andrés and M. Merchán, Pure & Appl. Chem. **65**, 1693 (1993).

[46] R. B. Murphy and R. P. Messmer, Chem. Phys. Lett. **5**, 443 (1991).

[47] B. O. Roos, P. R. Taylor and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).

[48] K. Ruedenberg and K. R. Sundberg, in: "Quantum science: methods and structure" (Plenum Press, New York, 1976).

[49] K. Ruedenberg, L. M. Cheung and S. T. Elbert, Int. J. Quant. Chem. **16**, 1069 (1979).

[50] P. Pulay and S. Saebo, Theor. Chim. Acta. **69**, 357 (1986).

[51] K. Wolinski, Theor. Chim. Acta. **82**, 459 (1992).

[52] V. R. Saunders and J. H. v. Lenthe, Mol. Phys. **48**, 923 (1983).

[53] J. Werner and E. A. Reinsch, J. Chem. Phys. **76**, 3144 (1982).

[54] P.-O. Löwdin, Advances in Physics **5**, 1 (1956).

[55] G. H. Golub and C. F. Van Loan, "Matrix Computations" (The Johns Hopkins University Press, Baltimore, London, 1993).

[56] C. G. J. Jacobi, Astronomische Nachrichten 297 (1845).

[57] R. Barrett, M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, "Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods" (SIAM, Philadelphia, 1994).

[58] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[59] P. Pulay, J. Comput. Chem. **3,** 556 (1982).

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

[61] S. R. Langhoff and E. R. Davidson, Int. J. Quant. Chem. **8**, 61 (1974).

[62] J. Verbeek, J. H. v. Lenthe, P. J. J. A. Timmermans, A. Mackor and P. H. M. Budzelaar, J. Organ. Chem. **52**, 2955 (1987).

[63] C. W. Bauschlicher and P. R. Taylor, J. Chem. Phys. **85**, 6510 (1986).

[64] M. Gutowski, J. H. v. Lenthe, J. Verbeek, F. B. v. Duijneveldt and G. Chalasinski, Chem. Phys. Lett. **124**, 370 (1986).

[65] M. Gutowski, G. C. van Duijneveldt-van de Rijdt, J. H. van Lenthe and F. B. van Duijneveldt, J. Chem. Phys. **98**, 4728 (1993).

[66] R. J. Harrison and S. Zarrabian, Chem. Phys. Lett. **158**, 193 (1989).

[67] W. Meyer, P. C. Hariharan and W. Kutzelnigg, J. Chem. Phys. **73**, 1880 (1980).

[68] T. H. Dunning Jr., J. Chem. Phys. **53**, 2823 (1970).

[69] C. P. Byrman, "Non-orthogonal Orbitals in Chemistry", thesis, Utrecht University (1995).

[70] S. L. Guberman, J. Chem. Phys. **67**, 1125 (1977).

[71] G. C. Lie and E. Clementi, J. Chem. Phys. **60**, 1288 (1974).

[72] R. McWeeny, Int. J. Quantum Chem. Symp. **24**, 733 (1990).

[73] B. J. Moss, F. W. Bobrowics and W. J. Goddard III, J. Chem. Phys. **63**, 4632 (1975).

[74] R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[75] H. J. J. van Dam and J. H. van Lenthe, Mol. Phys. **90**, 1007 (1997)

[76] F. J. Hoogesteeger, "Oligo(cyclohexylidenes): Development of Novel Functional and Organized Materials", thesis, Utrecht University (1996).

[77] F. J. Hoogesteeger, J. H. van Lenthe and L. W. Jenneskens, Chem. Phys. Lett. **259**, 178 (1996).

[78] R. J. Buenker, S. D. Peyerimhoff and W. Butscher, Molec. Phys. **35**, 771 (1978).

[79] R. J. Buenker, S.-K. Shih and S. D. Peyerimhoff, Chem. Phys. **36**, 97 (1979).

[80] R. J. Buenker, in: "The Workshop on Quantum Chemistry and Molecular Physics", (Wollongong, Australia, 1980).

[81] R. J. Buenker, in: "Studies in Physical and Theoretical Chemistry", **21**, 11 (1982).

[82] M. F. Guest, P. Fantucci, R. J. Harrison, J. Kendrick, J. H. van Lenthe, K. Schoeffel and P. Sherwood, GAMESS-UK (Daresbury, 1993).

[83] W. J. Hehre, R. Ditchfield and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[84] J. A. Pople, R. Seeger and R. Krishnan, Quantum Chemistry Symposium **11**, 149 (1977).

[85] D. D. Wagman, W. H. Evans, V. B. Parker, R. H. Schumm, I. Halow, S. M. Bailey, K. L. Churney and R. L. Nutall, in: "Journal of Physical and Chemical Reference Data" (National Bureau of Standards, Washington, DC, 1982).

Chapter 3

Convergence behaviour of multi-reference perturbation theory

An indicator

Huub J.J. van Dam, Joop H. van Lenthe
Mol. Phys. **90**, 1007 (1997)

Abstract

An approach to investigate the convergence behaviour of a (multi-reference) MP 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 perturbations series for stretched H_2 molecules and H_2O molecules. The MRMP approach does show improved convergence characteristics compared to single-reference MP for stretched H_2 molecules. Applications of the approach to O_2 at 4.0 bohr, and N_2 at 5.0 bohr show that the MRMP series is divergent for these cases. Furthermore, in the latter case it was found that the multi-reference series shows much stronger divergence than the single-reference series.

3.1 Introduction

Recent developments in multi-reference perturbation theory (MRPT) and complete active space perturbation theory (CASPT) [1-4] in particular have lead to programs which allow correlating calculations on relatively large molecular systems. As a consequence MRPT has become rather popular. However, to limit the computational cost most programs can not evaluate the perturbation series beyond second order, which is only the first term in the correlation energy. Only some programs that are based on a multi-reference 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 to consider the truncated series as an approximation to some full series.

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

3.2 Theory

The program [5] is based on the theory presented in papers 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 diagonalise then perturb approach, thus applying Møller-Plesset perturbation theory to a general contracted multi-configurational reference function. The singly and doubly excited states are obtained by (repeated) action of single excitation operators on the reference function and subsequent orthonormalisation, thus generating a contracted CI space.

The zeroth order Hamiltonian was chosen to be

$$H_0 = P_0FP_0 + P_1FP_1 + P_2FP_2 + \dots \tag{1}$$

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 the one proposed by Andersson [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 denegerate 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 O_2 molecule. The perturbation series [9,10] is generated by

$$\begin{aligned} \left(H_0 - E^{(0)}\right)\Psi^{(k)} &= \sum_{r=1}^{k} E^{(r)}\Psi^{(k-r)} - \left(H - H_0\right)\Psi^{(k-1)} \\ E^{(0)} &= \left\langle \Psi^{(0)} \middle| H_0 \middle| \Psi^{(0)} \right\rangle \\ E^{(k)} &= \left\langle \Psi^{(0)} \middle| H - H_0 \middle| \Psi^{(k-1)} \right\rangle, k > 0 \end{aligned} \tag{2}$$

where H is the Hamiltonian matrix in the basis of contracted states, $E^{(k)}$ 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 to 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

$$\left(H_0 - E^{(0)}\right)^{-1}.$$

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

$$\begin{aligned} \left(H_0 - E^{(0)}\right)^{-1} = &\left(P_0FP_0 - E^{(0)}\right)^{-1} + \left(P_1FP_1 - E^{(0)}\right)^{-1} \\ &+ \left(P_2FP_2 - E^{(0)}\right)^{-1} + \ldots \end{aligned} \tag{3}$$

is block diagonal too. Note that the first term is undefined because (P_0FP_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, 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.

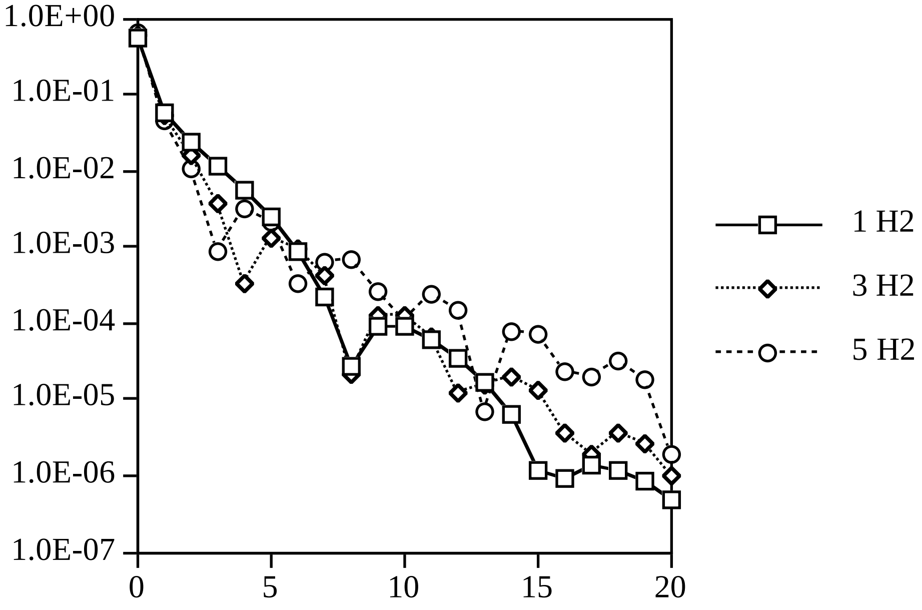
To obtain the limiting energies the perturbation series should converge to, we implemented a contracted CI option also. For this we used the contracted matrix-vector product $H\Psi$ from the PT code and plugged it into a Davidson diagonalisation [11,12] routine.

3.3 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 some *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 bondlength, 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 bohr apart. 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 multi-reference calculations 1, 2 and 3 molecules are involved, using a CAS reference containing the 1σ and $2\sigma^*$ orbitals of each H_2 molecule. In figures **I** and **II** the deviation of the *N*-th order corrected perturbation energy from the (contracted) CI energy per molecule is plotted on a logarithmic scale.

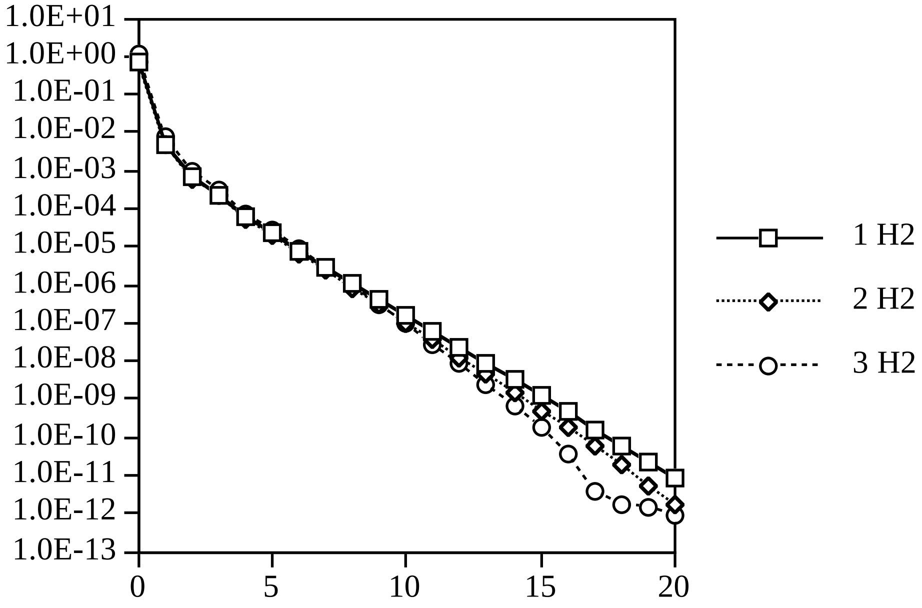
Figure I: MPPT calculations on non-interacting H_2 molecules R_{HH} is 2.776 bohr and R_{H2H2} is 50.0 bohr



The difference between the energy obtained from the perturbation series accumulated to *i*-th order and the infinite order CI energy are plotted in absolute value on the vertical axis. The scale along this axis is logarithmic and the unit is a hartree. Horizontally the orders in perturbation theory are plotted.

Figure **I** 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 **II** shows that the multi-reference approach removes the degeneracy and the series convergences smoothly. Also, for the multi-reference series the convergence is similar for 1, 2 and 3 H_2 molecules.

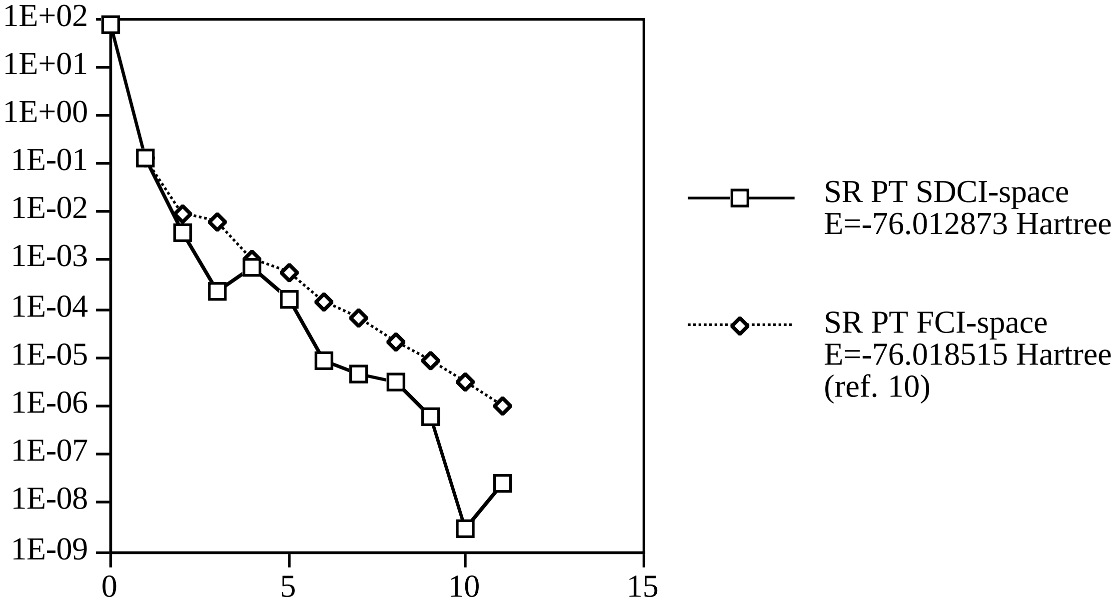
Figure II: CAS-MPPT calculations on non-interacting H_2 molecules R_{HH} is 2.776 bohr and R_{H2H2} is 50.0 bohr



The axes are defined equivalently to those in figure **I**.

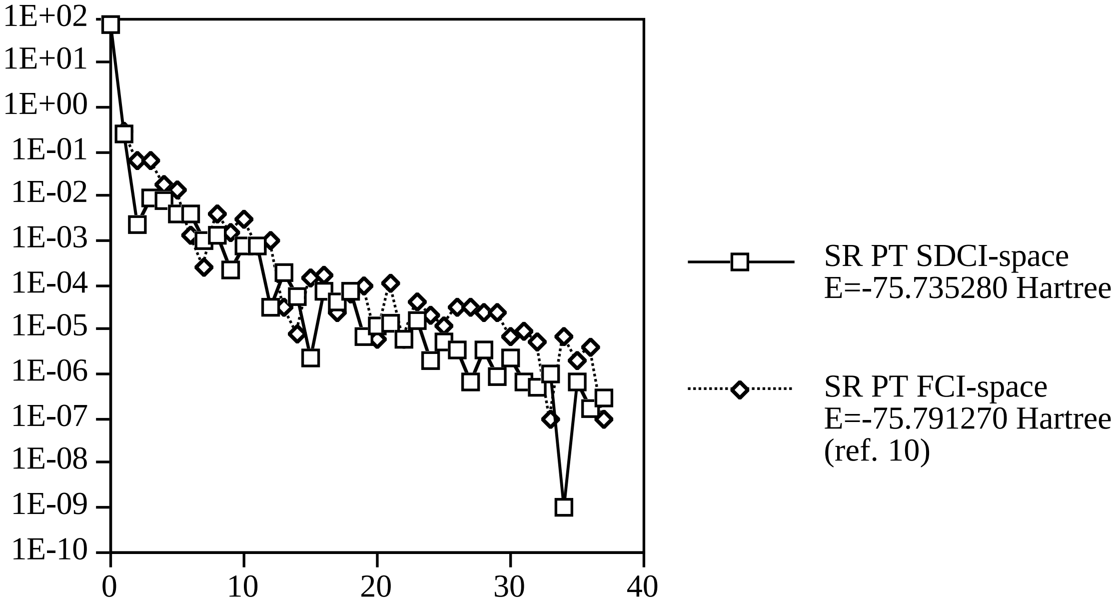
The H_2 test may seem an academic one; We now turn to a more realistic test involving the H_2O molecule. We use the results of Handy *et.al.* [10] in a study on the convergence of single-reference MP for the groundstate of the water molecule. We use their geometries at R_e and $2R_e$ ($R_e = 0.967$ Å and $\theta = 107.6°$) and basis set (6-21G) [14-18] to calculate the MP series in the singles-doubles space only. The results are shown in figures **III** and **IV**. In these figures 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 reasonbly 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 bondlength 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 23-rd order at $2.5R_e$.

Figure III: MPPT calculations on H_2O in C_{2v} symmetry at R_{OH} is 0.967 Å and θ_{HOH} is 107.6°



The axes are defined equivalently to those in figure **I**.

Figure IV: MPPT calculations on H_2O in C_{2v} symmetry at R_{OH} is 1.934 Å and θ_{HOH} is 107.6°



The axes are defined equivalently to those in figure **I**.

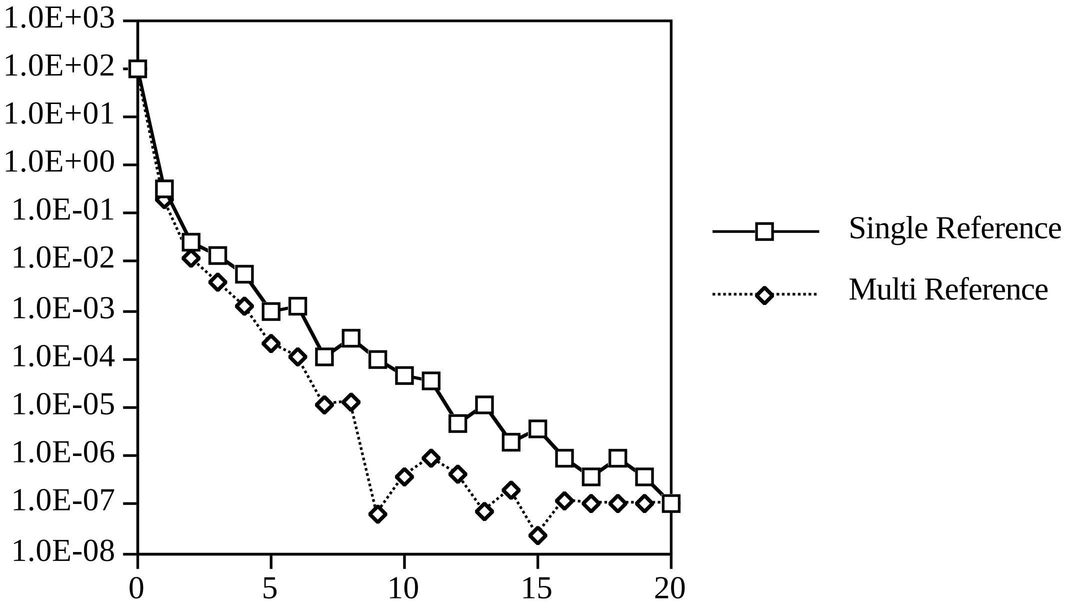
It has come to our attention that recently Olsen and coworkers showed that the convergence behaviour of the Møller-Plesset perturbation series is basisset dependent [19,20]. For example the single-reference MP series in the full CI space for the neon atom diverges if a basisset with diffuse functions is used. The approach we suggest in this article failed to reproduce these results both in single- and multi-reference 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 H_2 molecules both in single-reference and multi-reference calculations, as well as in single-reference calculations on H_2O. From the tests performed we conclude that the chosen approach seems to be satisfactory in studying the convergence behaviour. Using this result we apply the test to the nitrogen molecule at the equilibrium geometry (R = 2.1 bohr) and at a stretched geometry (R = 5.0 bohr) comparing the single-reference and the multi-reference series. Also, the test is applied to the oxygen molecule at R = 4.0 bohr.

3.4 Application to N_2

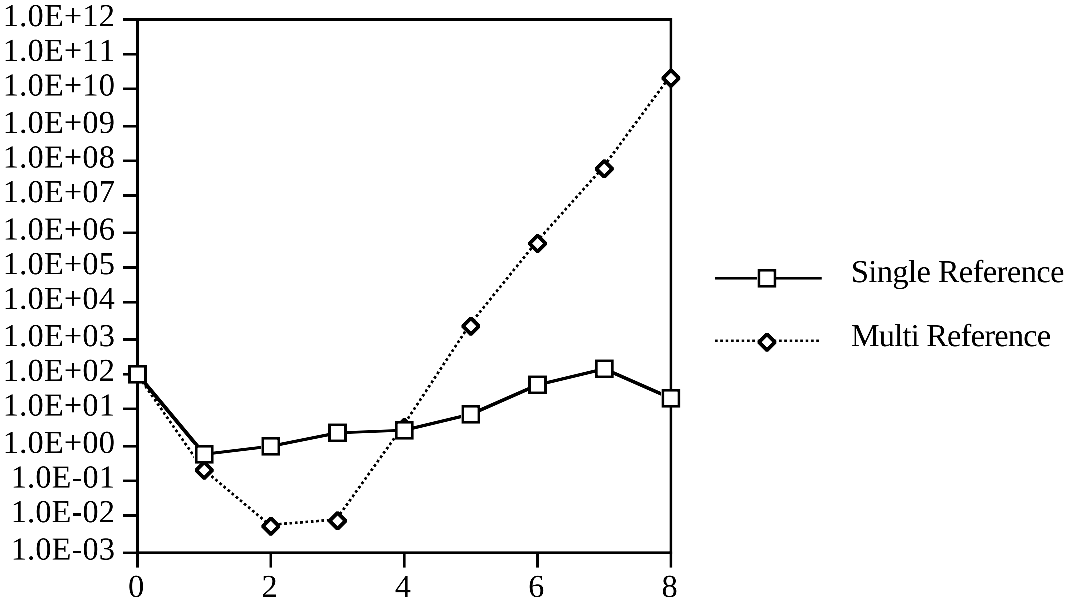
In a recent article where we presented our implementation of multi-reference Møller-Pleset perturbation theory [6] an application to the nitrogen molecule suggested that the multi-reference series is divergent at R = 5.0 bohr. Therefore, a closer inspection seems appropriate. Also, we want to comment on the "classical" idea that perturbation theory is allowed when the perturbation (H - H_0) is small enough, whereas we emphasised 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 = 2.1 bohr) and the stretched geometry (R = 5.0 bohr) using the 6-311g* [21] basis-set. Both multi-reference calculations using a CAS in the 5σ, $6\sigma^*$, $1\pi_x$, $2\pi_x^*$, $1\pi_y$ and $2\pi_y^*$ orbitals as a reference and single-reference calculations are performed. The 1S orbitals are frozen at the MCSCF level and not correlated. The results are shown in figures **V** and **VI**.

Figure V: Convergence behaviour of the MPPT series of the nitrogen molecule at R_{NN} is 2.1 bohr



The axes are defined equivalently to those in figure **I**.

Figure VI: Convergence behaviour of the MPPT series of the nitrogen molecule at R_{NN} is 5.0 bohr



The axes are defined equivalently to those in figure **I**.

From the calculations at the equilibrium geometry it is clear that the multi-reference 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 multi-reference series the first term (MRMP2) improves the energy and divergence sets in 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. Tabulating these values together with the zeroth-order energies table **I** is obtained.

Table I: Zeroth-order and first-order energies for single- and multi-reference calculations on the nitrogen molecule.

		$E^{(0)}$ (hartree)	$E^{(1)}$ (hartree)
SR	R = 2.1 bohr	-70.91866361	-38.04781784
MR	R = 2.1 bohr	-70.38351456	-38.72366820
SR	R = 5.0 bohr	-69.62628236	-38.46756589
MR	R = 5.0 bohr	-67.38477070	-41.40496354

At both geometries the perturbation is smaller for the single-reference approach than for the multi-reference approach. Also, the zeroth-order energy is lower for the single-reference calculations than for the multi-reference calculations suggesting that the single-reference zeroth-order approximation is the better one. This suggestion is odd because the single-reference zeroth-order wavefunction is lacking an essential physical quality, i.e. proper dissociation. Furthermore, at the equilibrium geometry the multi-reference series converges faster than the single-reference, even though the perturbation is larger. At the stretched geometry the multi-reference 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 $E^{(1)}$.

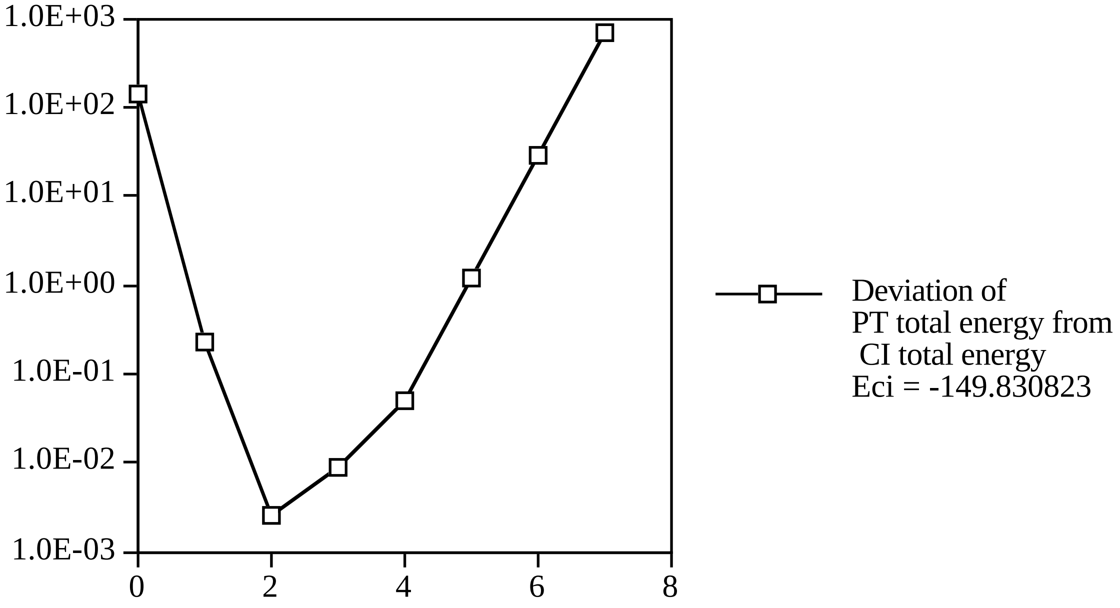
3.5 Application to O_2

Besides the N_2 molecule we also studied the potential energy surface of O_2 in our previous paper [6]. Because the N_2 series diverges at $R = 5.0$ bohr we revisit the O_2 molecule here to validate our original results.

The triplet groundstate 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, C.P. Byrman studied this maximum using Valence Bond calculations [22]. In these calculations the wavefunction was written as a superposition of structures describing the triplet 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 at O-O distances larger than 3.0 bohr, 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 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 7-th 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 1S oribtals are frozen at the MCSCF level and not correlated. In figure **VII** the deviation of the perturbation energy is plotted and in figure **VIII** the 2-norm of the *i*-th order correction to the wavefunction is plotted. Both plots are on a logarithmic scale.

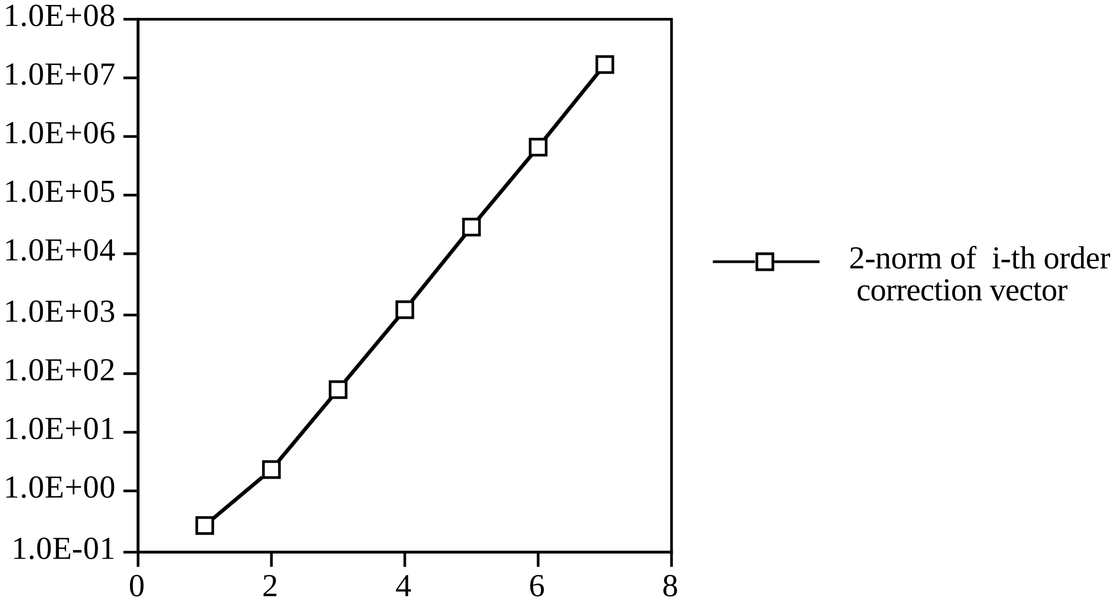
From the plots it is clear that the perturbation series is divergent in this case. The norm of the correction vector is monotonously growing. Although the series is divergent the MRMP2/3 estimates of the potential surface we published recently [6] show no sign of any problems. This suggests that it is rather dangerous to judge a perturbation series by its first few terms only.

Figure VII: The convergence behaviour of the MR-MPPT series of oxygen molecule at R_{OO} is 4.0 bohr



The axes are defined equivalently to those in figure **I.**

Figure VIII: The 2-norm of the MR-MPPT correction vector in each order of perturbation theory for the oxygen molecule at R_{OO} is 4.0 bohr



3.6 Conclusions

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 multi-reference, and the single-reference H_2O molecule at the equilibrium geometry and a stretched geometry. Applying the same approach to the oxygen molecule stretched to a bondlength of 4.0 bohr shows that the MRPT series is divergent at that geometry. Also we confirm that the MRPT series is divergent in nitrogen molecule stretched to a bondlength of 5.0 bohr, as we suspected from calculations published recently. While we failed to show the diffuse basisset effects of Olsen and coworkers on the neon atom, we do show that divergences may appear even with small basissets in typical multi-reference 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.

Acknowledgements

The authors want to thank Jeppe Olsen for sending copies of his manuscripts and 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, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institue for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller, Karen Schuchardt, or Don Jones for further information.

References

[1] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[2] K. Andersson, P.-Å. Malmqvist and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[3] B. O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995).

[4] H.-J. Werner, Mol. Phys. **89**, 645 (1996).

[5] J. H. van Lenthe, P. Pulay and H. J. J. van Dam, Multi-Reference MØller-Plesset Perturbation Theory Option (Fayetteville, Arkansas / Utrecht, The Netherlands; 1989-1996).

[6] H. J. J. van Dam, J. H. van Lenthe and P. Pulay, Chapter 2

[7] K. Wolinski and P. Pulay, J. Chem. Phys **90**, 3647 (1989).

[8] K. Wolinski, H. L. Sellers and P. Pulay, Chem. Phys. Lett. **140**, 225 (1987).

[9] P. J. Knowles, K. Somasundram, N. C. Handy and K. Hirao, Chem. Phys. Lett. **113**, 8 (1985).

[10] N. C. Handy, P. J. Knowles and K. Somasundram, Theor. Chim. Acta **68**, 87 (1985).

[11] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).

[12] E. R. Davidson, Comput. Phys. Comm. **53**, 49 (1989).

[13] T. H. Dunning Jr., J. Chem. Phys. **53**, 2823 (1970).

[14] J. S. Binkley, J. A. Pople and W. J. Hehre, J. Am. Chem. Soc. **102**, 939 (1980).

[15] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro and W. J. Hehre, J. Am. Chem. Soc. **104**, 2797 (1982).

[16] M. J. Frisch, J. A. Pople and J. S. Binkley, J. Chem. Phys. **80**, 3265 (1984).

[17] K. D. Dobbs and W. J. Hehre, J. Comput. Chem. **7**, 359 (1986).

[18] K. D. Dobbs and W. J. Hehre, J. Comput. Chem. **8**, 861 (1987).

[19] J. Olsen, O. Christiansen, H. Koch and P. Jørgensen, The Journal of Chemical Physics **105**, 5082 (1996).

[20] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch and P.-Å. Malmqvist, Chem. Phys. Lett. **261**, 369 (1996).

[21] R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[22] C. P. Byrman, "Non-orthogonal Orbitals in Chemistry", thesis, Utrecht University (1995).

[23] S. L. Guberman, J. Chem. Phys. **67**, 1125 (1977).

[24] G. C. Lie and E. Clementi, J. Chem. Phys. **60**, 1288 (1974).

[25] R. McWeeny, Int. J. Quantum Chem. Symp. **24**, 733 (1990).

[26] B. J. Moss, F. W. Bobrowics and W. J. Goddard III, J. Chem. Phys. **63**, 4632 (1975).

Chapter 4

Exact size consistency of multi-reference Møller-Plesset perturbation theory

Theory

Huub J. J. van Dam, Joop H. van Lenthe, and Paul J. A. Ruttink

Submitted to Int. J. Quant. Chem. (1997)

Abstract

Single-reference closed shell Møller-Plesset perturbation theory is well known for its size consistency, a quality that is essential for consistent comparisons of calculations on molecules of different size. However, it is far from obvious whether this quality can be retained in the multi-reference case. In this paper it is shown that an exactly size consistent generalisation to multi-reference perturbation theory can be constructed.

The central result is that the zeroth-order Hamiltonian should be constructed using separate projection operators for each excitation level, i.e. it should contain no couplings between different excitation levels.

4.1 Introduction

Single-reference Møller-Plesset perturbation theory (MP) has the important property that it is size consistent [1–3]. Basically this means that the calculated energy of a system scales correctly with the size of the system [4]. This property is becoming ever more important as more powerful computers and better *ab initio* codes allow calculations on ever larger molecules. This together with the fact that MP is very efficient at low orders has made the method popular over limited CI approaches (which are not size consistent) and Coupled Cluster (CC) methods [5–11] (which are computationally rather involved). These properties have become strong motivations for the development of more advanced approaches in perturbation theory, e.g. Multi-Reference Møller-Plesset (MRMP) perturbation theory. This generalisation of perturbation theory to the multi-reference case is important to enable the calculation of potential energy surfaces and the calculation of low lying excited states. As a consequence numerous proposals of such methods have been published (see [4, 12, 13] and the references therein). These methods may be organized into two families [13]. One in which the reference functions are perturbed first by the functions outside the reference space and the resulting effective Hamiltonian in the reference space is diagonalised. The other in which a reference function is created by diagonalizing the Hamiltonian in the reference space first and then perturbing this function. The latter methods seem to be less sensitive to convergence problems due to intruder states [14] and have become rather popular recently.

The popularity of the first-diagonalize-then-perturb methods is for a large part due to the enormous efforts invested by Roos and Andersson in particular, and their collaborators (see e.g. [15–29]). Apart from this work there have been significant contributions by others [30–43], and in particular by Wolinski *et al.* [44–46], Murphy *et al.* [47–49], and most recently by Werner [50] who implemented the approach to the MRMP3 level and studied the performance of CASMP3. The methods proposed by Andersson *et al.*, Wolinski *et al.*, and Werner all have in common that the first order correction to the wavefunction Ψ_1 is expanded in a space of internally contracted excited states. This space is typically obtained by applying the unitary group generators [51] to the multi-configurational reference function. The Møller-Plesset zeroth-order Hamiltonian $H^{(0)}$ is in general not diagonal in this space, therefore Ψ_1 is calculated through iteratively solving a linear system of equations. The three methods differ, however, in the chosen $H^{(0)}$ and therefore may yield different results. Although all these methods limit to the size consistent Møller-Plesset perturbation theory in the single-reference closed shell case there is discussion about the question whether or not size consistency is retained in the multi-reference case.

Wolinski *et al.* [44] proposed an approach which is strictly equivalent to the canonical MP except for the use of a general reference function. In this approach the $H^{(0)}$ is defined as

$$H^{(0)} = P_0 F P_0 + P_\mathrm{S} F P_\mathrm{S} + P_\mathrm{D} F P_\mathrm{D} + \dots \tag{1}$$

where F is the N-electron Fock operator, and P_0, P_S and P_D are the projection operators onto the reference function, the singly excited states and the doubly excited states respectively (for further details see [44, 45]). Throughout this paper this zeroth-order Hamiltonian will be denoted as $H^{(0)}_{S,D}$. Because the method is equivalent to canonical MP Wolinski claims that the $H^{(0)}$ they chose guarantees size consistency in the limited sense of Pople *et al.* [8] if a size consistent reference function is used. Later Wolinski *et al.* [45] stated that this may easily be proven using localized orbitals. In that case it should be possible to show that $H^{(0)}$ is a sum of local operators.

Andersson *et al.* [16, 17] and recently Werner [50] chose an $H^{(0)}$ of the form

$$H^{(0)} = P_0 F P_0 + P_\mathrm{SD} F P_\mathrm{SD} + \dots \tag{2}$$

where P_SD is the projection operator onto the space of the singly and doubly excited states. This zeroth-order Hamiltonian will be denoted as $H^{(0)}_{S\cup D}$. The space in which Werner expands $H^{(0)}$ is somewhat different from the space that is used by Wolinski because the internal and semi-internal excitations are left uncontracted, whereas the doubly external configurations are contracted. Andersson [17] treats the singles and the doubles within the reference space separately from the other excited states. Because of this and the fact that these states have no interaction with the reference function over the total Hamiltonian these states do not show up in CASMP2. Malmqvist [50, 52] pointed out that this form of $H^{(0)}$ does not yield strictly size consistent energies because of the choice of the projection operators. Indeed, Werner found non-zero size consistency errors in calculations on N_2, O_2 and F_2 (see [50], page 658).

In this article we study the size consistency of MRMP. Here size consistency means that the energy of a system consisting of two non-interacting subsystems should be the sum of the energies of the separate subsystems, c.f. [8]. First we reproduce Rayleigh-Schrödinger perturbation theory for the Schrödinger equation in a non-orthogonal N-electron space. From this the requirement that $H^{(0)}$ has to satisfy in order to obtain a size consistent perturbation theory is derived. Then we study the role of the projection operators to show why $H^{(0)}_{S,D}$ yields a size consistent perturbation theory and $H^{(0)}_{S\cup D}$ does not. The insight gained from this discussion finally allows for the definition of a zeroth-order Hamiltonian that yields exactly size consistent energies also but has less off-diagonal elements. As a consequence this zeroth-order Hamiltonian is easier to implement and more efficient to use.

Because in practice MRMP is restricted to second and at most third order due to the computational cost, we restrict the equations to the first order interacting space (i.e. singles and doubles). However, the basic results are completely general and not limited to any particular order in perturbation theory.

4.2 Generalized perturbation theory

In multi-reference perturbation theory the space in which the first order correction to the wavefunction is calculated is orthogonalized at least at an intermediate stage. However it is not required to express the orthogonalisation explicitly in the equations when writing out the Rayleigh-Schrödinger equations. After all orthogonalisation of the space should not yield physically different results.

Therefore we will here consider the Schrödinger equation expanded in a non-orthogonal N-electron basis

$$H\Psi = E\mathrm{S}\Psi \tag{3}$$

where S is the metric. Although the N-electron basis is non-orthogonal it is tacitly assumed in this article that the orbitals are orthogonal. In principle the metric may be expanded in a way similar to the expansion of the Hamiltonian. However, in this study there is no real need for that, so the metric will be used as is. The Hamiltonian, the energy, and the wavefunction may be expanded in the usual way

$$\begin{aligned}
\hat{H} &= \hat{H}^{(0)} + \lambda\hat{V} && (4)\\
E &= E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \ldots && (5)\\
\Psi &= \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \ldots && (6)
\end{aligned}$$

and substituted in the Schrödinger equation, collecting terms of the same order in λ yields for the corrections to the wavefunction

$$(H^{(0)} - E^{(0)}\mathrm{S})\Psi^{(k)} \;=\; \sum_{r=1}^{k} E^{(r)}\mathrm{S}\Psi^{(k-r)} - V\Psi^{(k-1)} \tag{7}$$

Note that equation 7 evaluated to the first order is essentially the same as equation (5) in [16] published by Andersson *et al.*

From the above equation evaluated for k equal to zero it follows that $\Psi^{(0)}$ should be an eigenfunction of the generalized zeroth-order equation

$$(H^{(0)} - E^{(0)}\mathrm{S})\Psi^{(0)} = 0 \tag{8}$$

This implies that all corrections $\Psi^{(k)}$ to the wavefunction should be expanded in a space orthogonal to $\Psi^{(0)}$, i.e. $\langle\Psi^{(0)}|\mathrm{S}|\Psi^{(k)}\rangle = 0$ if $k \neq 0$. Using this the energy expressions derived through projecting equation 7 with $\Psi^{(0)}$ from the left and integrating become

$$E^{(0)} \;=\; \langle\Psi^{(0)}|\mathrm{S}|\Psi^{(0)}\rangle^{-1}\langle\Psi^{(0)}|H^{(0)}|\Psi^{(0)}\rangle \tag{9}$$

$$E^{(k)} \;=\; \langle\Psi^{(0)}|\mathrm{S}|\Psi^{(0)}\rangle^{-1}\langle\Psi^{(0)}|\mathrm{V}|\Psi^{(k-1)}\rangle, k > 0. \tag{10}$$

We are interested in the size consistency of these energies. Therefore the total energy of a system AB consisting of two non-interacting subsystems A and B should be, for physical reasons, the sum of the subsystem total energies. For comparison of the subsystem equations with the total system equation the equations of these systems have to be expressed in the same function space. For the subsystems the generalized eigenvalue equations are

$$H_A\Psi_A \;=\; E_A\mathrm{S}_A\Psi_A \tag{11}$$

$$H_B\Psi_B \;=\; E_B\mathrm{S}_B\Psi_B \tag{12}$$

Multiplying equation 11 by $\mathrm{S}_B\Psi_B$ on both sides, and multiplying equation 12 by $\mathrm{S}_A\Psi_A$ on both sides, subsequently adding the equations yields

$$(H_A\mathrm{S}_B + \mathrm{S}_A H_B)\Psi_A\Psi_B = (E_A + E_B)\mathrm{S}_A\mathrm{S}_B\Psi_A\Psi_B \tag{13}$$

which is the equation for the system AB.

So in a non-orthogonal space the wavefunction describing a system consisting of two non-interacting subsystems is a direct product function

$$\Psi_{AB} = \Psi_A \Psi_B \tag{14}$$

as in an orthogonal basis. Substituting the expansion for the wavefunction and insisting that the equation should hold for any λ gives

$$\Psi_{AB}^{(i)} = \sum_{k=0}^{i} \Psi_A^{(k)} \Psi_B^{(i-k)} \tag{15}$$

for the *i*-th correction to the wavefunction of the total system.

From equation 13 it is concluded that the total Hamiltonian of the total system is

$$H_{AB} = H_A \mathrm{S}_B + \mathrm{S}_A H_B. \tag{16}$$

Expanding the Hamiltonians of the total system and the subsystems gives

$$H_{AB}^{(0)} + \lambda \mathrm{V}_{AB} = (H_A^{(0)} \mathrm{S}_B + \mathrm{S}_A H_B^{(0)}) + \lambda(\mathrm{V}_A \mathrm{S}_B + \mathrm{S}_A \mathrm{V}_B) \tag{17}$$

which defines $H_{AB}^{(0)}$ and V_{AB}.

Substituting the expression for the perturbationV_{AB} resulting from 17 and the wavefunction given by equation 15 in the k-th order correction to the energy now yields

$$\begin{aligned}
E_{AB}^{(k)} &= \langle \Psi_{AB}^{(0)} | \mathrm{S}_{AB} | \Psi_{AB}^{(0)} \rangle^{-1} \langle \Psi_{AB}^{(0)} | \mathrm{V}_{AB} | \Psi_{AB}^{(k-1)} \rangle, k > 0. && (18)\\
&= (\langle \Psi_A^{(0)} | \mathrm{S}_A | \Psi_A^{(0)} \rangle \langle \Psi_B^{(0)} | \mathrm{S}_B | \Psi_B^{(0)} \rangle)^{-1} \cdot \\
&\quad \sum_{i=0}^{k-1} \langle \Psi_A^{(0)} \Psi_B^{(0)} | \mathrm{V}_A \mathrm{S}_B + \mathrm{S}_A \mathrm{V}_B | \Psi_A^{(i)} \Psi_B^{(k-1-i)} \rangle && (19)\\
&= \langle \Psi_B^{(0)} | \mathrm{S}_B | \Psi_B^{(0)} \rangle^{-1} \langle \Psi_B^{(0)} | \mathrm{V}_B | \Psi_B^{(k-1)} \rangle + \\
&\quad \langle \Psi_A^{(0)} | \mathrm{S}_A | \Psi_A^{(0)} \rangle^{-1} \langle \Psi_A^{(0)} | \mathrm{V}_A | \Psi_A^{(k-1)} \rangle && (20)\\
&= E_A^{(k)} + E_B^{(k)} && (21)
\end{aligned}$$

which is exactly the size consistent result. Here we have used that $\langle \Psi^{(0)} | \mathrm{S} | \Psi^{(i)} \rangle = 0$ if $i > 0$.

So if $\Psi^{(0)}$ is an eigenfunction of the generalized zeroth-order equation and the corrections to the wavefunction are expanded in a space orthogonal to $\Psi^{(0)}$ and the Hamiltonian can be written as equation 17 then the resulting Rayleigh-Schrödinger perturbation series is size consistent in every order. To write the Hamiltonian as in equation 17 implies that

$$H_{AB}^{(0)} = H_A^{(0)} \mathrm{S}_B + \mathrm{S}_A H_B^{(0)}. \tag{22}$$

So to prove that the perturbation series is size consistent is to prove that the zeroth-order Hamiltonian satisfies equation 22. This equation requires that the dimer space can be expressed in terms of the monomer spaces.

In the next section it is shown how the choice of the projection operators may influence the size consistency. To this end $H_{S,D}^{(0)}$ and $H_{S\cup D}^{(0)}$ are written out for system AB. It is checked whether they satisfy equation 22. Following this exercise a way to simplify the zeroth-order Hamiltonian will be discussed.

4.3 The choice of the projection operators and the size consistency of multi-reference Møller-Plesset perturbation theory

In the internally contracted multi-reference perturbation theory considered here the reference function is defined as

$$\Psi^{(0)} = \sum_k c_k \Phi_k \tag{23}$$

where the Φ_k are Slater determinants spanning the reference space. The coefficients c_k and the orbitals are optimized in a MCSCF calculation. The singly, doubly, etc. substituted spaces are generated through the excitation operators $E_{rs} - E_{sr}$, $(E_{rs} - E_{sr})(E_{tu} - E_{ut})$, etc. E_{rs} denotes the unitary group generator [51], and r and s are orbital labels. This use of the unitary group generators guarantees that the singly substituted states are orthogonal to the reference state and that they are Brillioun states, i.e. have no interaction with the reference state over the total Hamiltonian. In the following discussions we will use the shorthand notation

$$\hat{E}_{rs} = E_{rs} - E_{sr}. \tag{24}$$

The higher excited subspaces will be orthogonalized to the lower ones to obtain well defined excitation levels as suggested by Wolinski et al. [44].

The zeroth-order Hamiltonian may be based on the Fock-operator

$$\begin{aligned}\hat{F} &= \sum_{rs} f_{rs}\hat{E}_{rs} &(25)\\ f_{rs} &= h_{rs}+\sum_{ij}\langle\Psi^{(0)}|\hat{E}_{ij}|\Psi^{(0)}\rangle\left[(rs|ij)-\frac{1}{2}(ri|js)\right] &(26)\end{aligned}$$

that Werner [50] uses but there is some freedom to choose a non-closed shell like Fock operator. Essential is that the Fock operator difinition for a system AB consisting of two non-interacting subsystems A and B satisfies

$$\hat{F}_{AB} = \hat{F}_A + \hat{F}_B \tag{27}$$

To ensure that $\Psi^{(0)}$ is an eigenfunction of the zeroth-order Hamiltonian projection operators will be used. In a non-orthogonal basis they are defined as

$$P_{\mathrm{X}} = XS^{-1}X^{\dagger} \tag{28}$$

where it is assumed for the moment that S is not singular. Because the discussion will be concentrated on the spaces of different excitation levels it is useful to introduce some notation to denote these spaces. The singly substituted space is defined through

$$X_1 = \sum_{it}\hat{E}_{it}\Psi^{(0)} + \sum_{ia}\hat{E}_{ia}\Psi^{(0)} + \sum_{tu}\hat{E}_{tu}\Psi^{(0)} + \sum_{ta}\hat{E}_{ta}\Psi^{(0)} \tag{29}$$

where the inactive orbitals are labeled i, and j, the active orbitals are labeled s, t, u, and v, and the external orbitals are labeled a, and b in accordance with the notation used by Andersson *et al.* [16]. The doubly substituted space is defined through

$$\begin{aligned}X_2 &= \sum_{ijtu}\hat{E}_{it}\hat{E}_{ju}\Psi^{(0)} + \sum_{ituv}\hat{E}_{it}\hat{E}_{uv}\Psi^{(0)} + \sum_{stuv}\hat{E}_{st}\hat{E}_{uv}\Psi^{(0)}\\ &+ \sum_{ijta}\hat{E}_{it}\hat{E}_{ja}\Psi^{(0)} + \sum_{itua,t\neq u}\hat{E}_{ia}\hat{E}_{tu}\Psi^{(0)} + \sum_{itua}\hat{E}_{ta}\hat{E}_{iu}\Psi^{(0)} + \sum_{tuva}\hat{E}_{tu}\hat{E}_{va}\Psi^{(0)}\\ &+ \sum_{ijab}\hat{E}_{ia}\hat{E}_{jb}\Psi^{(0)} + \sum_{itab}\hat{E}_{ia}\hat{E}_{tb}\Psi^{(0)} + \sum_{tuab}\hat{E}_{ta}\hat{E}_{ub}\Psi^{(0)} \qquad (30)\end{aligned}$$

Note that spin recoupled single excitations correspond to the $\hat{E}_{ta}\hat{E}_{it}$ contributions which are contained in the 6-th term in equation 30.

Due to the choice of the excitation operators the singly substituted states are orthogonal to the reference function by construction. However, with the doubly substituted states a projection is needed to orthogonalize them to the reference function and the singly substituted space. Here, the doubly substituted states will be written as

$$X_D = (1 - P_0 - P_\mathrm{S})X_2 \tag{31}$$

and P_D projects to this space. In other words, if P_2 is the idempotent projector on X_2 then P_D may be expressed as

$$\begin{aligned} P_\mathrm{D} &= P_2 - P_0P_2 - P_\mathrm{S}P_2 && (32)\\ &= P_2 - P_2P_0 - P_2P_\mathrm{S} && (33) \end{aligned}$$

In the next two paragraphs we will check the additivity of $H^{(0)}_{S,D}$ and $H^{(0)}_{S\cup D}$ respectively. This is done through writing out the zeroth-order Hamiltonians of the subsystems and the total system in the space of the total system so that they may be compared.

4.3.1 Additivity of $H^{(0)}_{S,D}$

For size consistency $H^{(0)}_{S,D}$ should satisfy equation 22 where the matrices are truncated to the first order interacting space since we are interested in MRMP2/3 only. So at most double excitations are allowed. Writing $H^{(0)}_A \mathrm{S}_B$ in the space of the total system means writing $H^{(0)}_{S,D}$ for system A irrespective of what happens on system B. Truncated to singles and doubles this gives

$$\begin{aligned} H^{(0)}_A \mathrm{S}_B &= P^A_0 P^B_0 \hat{F}_A P^A_0 P^B_0 \\ &+ P^A_0 P^B_\mathrm{S} \hat{F}_A P^A_0 P^B_\mathrm{S} \\ &+ P^A_0 P^B_\mathrm{D} \hat{F}_A P^A_0 P^B_\mathrm{D} \\ &+ P^A_\mathrm{S} P^B_0 \hat{F}_A P^A_\mathrm{S} P^B_0 \\ &+ P^A_\mathrm{S} P^B_\mathrm{S} \hat{F}_A P^A_\mathrm{S} P^B_\mathrm{S} \\ &+ P^A_\mathrm{D} P^B_0 \hat{F}_A P^A_\mathrm{D} P^B_0 \end{aligned} \tag{34}$$

whereas for the truncated $H^{(0)}_{AB}$ matrix we have

$$\begin{aligned} H^{(0)}_{AB} &= P^{AB}_0 (\hat{F}_A + \hat{F}_B) P^{AB}_0 \\ &+ P^{AB}_\mathrm{S} (\hat{F}_A + \hat{F}_B) P^{AB}_\mathrm{S} \\ &+ P^{AB}_\mathrm{D} (\hat{F}_A + \hat{F}_B) P^{AB}_\mathrm{D} \end{aligned} \tag{35}$$

$$
\begin{aligned}
&= P_0^{AB}\hat{F}_A P_0^{AB} \\
&+ P_{\mathrm{S}}^{AB}\hat{F}_A P_{\mathrm{S}}^{AB} \\
&+ P_{\mathrm{D}}^{AB}\hat{F}_A P_{\mathrm{D}}^{AB} \\
&+ H_{AB}^{(0)}(\hat{F}_B) \qquad (36)
\end{aligned}
$$

where $H_{AB}^{(0)}(\hat{F}_B)$ represents the contributions due to the Fock-operator on subsystem B. To compare equations 34 and 36 the projection operators for the total system should be expressed in terms of projection operators on the subsystems. That is (using equation 32)

$$
\begin{aligned}
P_0^{AB} &= P_0^A P_0^B \qquad (37) \\
P_{\mathrm{S}}^{AB} &= P_0^A P_{\mathrm{S}}^B + P_{\mathrm{S}}^A P_0^B \qquad (38) \\
P_{\mathrm{D}}^{AB} &= P_2^{AB} - P_0^{AB} P_2^{AB} - P_{\mathrm{S}}^{AB} P_2^{AB} \qquad (39)
\end{aligned}
$$

To remove the P_2 terms from equation 39 the projection operators on the total system are expressed in terms of the projection operators on the subsystems and the terms are recollected yielding

$$
\begin{aligned}
P_{\mathrm{D}}^{AB} &= P_2^{AB} - P_0^{AB} P_2^{AB} - P_{\mathrm{S}}^{AB} P_2^{AB} \qquad (40) \\
&= P_0^A P_2^B + P_{\mathrm{S}}^A P_{\mathrm{S}}^B + P_2^A P_0^B \\
&- P_0^A P_0^B (P_0^A P_2^B + P_{\mathrm{S}}^A P_{\mathrm{S}}^B + P_2^A P_0^B) \\
&- (P_0^A P_{\mathrm{S}}^B + P_{\mathrm{S}}^A P_0^B)(P_0^A P_2^B + P_{\mathrm{S}}^A P_{\mathrm{S}}^B + P_2^A P_0^B) \qquad (41) \\
&= P_0^A P_{\mathrm{D}}^B + P_{\mathrm{S}}^A P_{\mathrm{S}}^B + P_{\mathrm{D}}^A P_0^B \qquad (42)
\end{aligned}
$$

where the last equation is obtained using the idempotency of the projection operators and the fact that $P_0 P_{\mathrm{S}} = P_{\mathrm{S}} P_0 = 0$. So the projection operator onto the doubles of the total system can be expressed in terms of the projections on the reference, singles and doubles of the subsystems. Substituting these projection operators in 36 we have

$$
\begin{aligned}
H_{AB}^{(0)} &= P_0^A P_0^B \hat{F}_A P_0^A P_0^B \\
&+ P_0^A P_{\mathrm{S}}^B \hat{F}_A P_0^A P_{\mathrm{S}}^B \\
&+ P_{\mathrm{S}}^A P_0^B \hat{F}_A P_{\mathrm{S}}^A P_0^B \\
&+ P_0^A P_{\mathrm{D}}^B \hat{F}_A P_0^A P_{\mathrm{D}}^B \\
&+ P_{\mathrm{D}}^A P_0^B \hat{F}_A P_{\mathrm{D}}^A P_0^B \\
&+ P_{\mathrm{S}}^A P_{\mathrm{S}}^B \hat{F}_A P_{\mathrm{S}}^A P_{\mathrm{S}}^B \\
&+ H_{AB}^{(0)}(\hat{F}_B) \qquad (43)
\end{aligned}
$$

Comparing equation 34 and equation 43 they are found to be identical except for the term on subsystem B. The term on subsystem B is just that part of the zeroth-order Hamiltonian that corresponds to the Fock-operator on B. It corresponds to $H_B^{(0)}$ in the same way that the subsystem A part of equation 43 corresponds to $H_A^{(0)}$ given in equation 34. This means that this approach to perturbation theory satisfies equation 22 and therefore is size consistent.

4.3.2 Additivity of $H^{(0)}_{S\cup D}$

Here the same reasoning as in the previous section is applied. Expressing $H^{(0)}_{S\cup D}$ in the basis of system AB leads to

$$\begin{aligned}
H_A^{(0)}\mathrm{S}_B &= P_0^A P_0^B \hat{F}_A P_0^A P_0^B \\
&+ P_0^A (P_\mathrm{S}^B + P_\mathrm{D}^B) \hat{F}_A P_0^A (P_\mathrm{S}^B + P_\mathrm{D}^B) \\
&+ (P_\mathrm{S}^A + P_\mathrm{D}^A) P_0^B \hat{F}_A (P_\mathrm{S}^A + P_\mathrm{D}^A) P_0^B \\
&+ (P_\mathrm{S}^A + P_\mathrm{D}^A)(P_\mathrm{S}^B + P_\mathrm{D}^B) \hat{F}_A (P_\mathrm{S}^A + P_\mathrm{D}^A)(P_\mathrm{S}^B + P_\mathrm{D}^B) \qquad (44)
\end{aligned}$$

Writing out and truncating to at most double excitations because we are interested in the first order interacting space only, gives

$$\begin{aligned}
H_A^{(0)}\mathrm{S}_B &= P_0^A P_0^B \hat{F}_A P_0^A P_0^B \\
&+ P_0^A P_\mathrm{S}^B \hat{F}_A P_0^A P_\mathrm{S}^B \\
&+ P_0^A P_\mathrm{D}^B \hat{F}_A P_0^A P_\mathrm{D}^B \\
&+ P_\mathrm{S}^A P_0^B \hat{F}_A P_\mathrm{S}^A P_0^B \\
&+ P_\mathrm{S}^A P_0^B \hat{F}_A P_\mathrm{D}^A P_0^B \\
&+ P_\mathrm{D}^A P_0^B \hat{F}_A P_\mathrm{S}^A P_0^B \\
&+ P_\mathrm{D}^A P_0^B \hat{F}_A P_\mathrm{D}^A P_0^B \\
&+ P_\mathrm{S}^A P_\mathrm{S}^B \hat{F}_A P_\mathrm{S}^A P_\mathrm{S}^B \qquad (45)
\end{aligned}$$

whereas for the truncated $H_{AB}^{(0)}$ matrix we have

$$\begin{aligned}
H_{AB}^{(0)} &= P_0^{AB} (\hat{F}_A + \hat{F}_B) P_0^{AB} \\
&+ (P_\mathrm{S}^{AB} + P_\mathrm{D}^{AB})(\hat{F}_A + \hat{F}_B)(P_\mathrm{S}^{AB} + P_\mathrm{D}^{AB}) \qquad (46) \\
&= P_0^{AB} \hat{F}_A P_0^{AB} \\
&+ P_\mathrm{S}^{AB} \hat{F}_A P_\mathrm{S}^{AB} \\
&+ P_\mathrm{S}^{AB} \hat{F}_A P_\mathrm{D}^{AB}
\end{aligned}$$

$$\begin{aligned}
&+ P^{AB}_{\mathrm{D}}\hat{F}_A P^{AB}_{\mathrm{S}} \\
&+ P^{AB}_{\mathrm{D}}\hat{F}_A P^{AB}_{\mathrm{D}} \\
&+ H^{(0)}_{AB}(\hat{F}_B) \qquad (47)
\end{aligned}$$

Substituting the equations 37, 38,and 42 gives

$$\begin{aligned}
H^{(0)}_{AB} &= P^A_0 P^B_0 \hat{F}_A P^A_0 P^B_0 \\
&+ (P^A_0 P^B_{\mathrm{S}} + P^A_{\mathrm{S}} P^B_0)\hat{F}_A (P^A_0 P^B_{\mathrm{S}} + P^A_{\mathrm{S}} P^B_0) \\
&+ (P^A_0 P^B_{\mathrm{S}} + P^A_{\mathrm{S}} P^B_0)\hat{F}_A (P^A_0 P^B_{\mathrm{D}} + P^A_{\mathrm{D}} P^B_0 + P^A_{\mathrm{S}} P^B_{\mathrm{S}}) \\
&+ (P^A_0 P^B_{\mathrm{D}} + P^A_{\mathrm{D}} P^B_0 + P^A_{\mathrm{S}} P^B_{\mathrm{S}})\hat{F}_A (P^A_0 P^B_{\mathrm{S}} + P^A_{\mathrm{S}} P^B_0) \\
&+ (P^A_0 P^B_{\mathrm{D}} + P^A_{\mathrm{D}} P^B_0 + P^A_{\mathrm{S}} P^B_{\mathrm{S}})\hat{F}_A (P^A_0 P^B_{\mathrm{D}} + P^A_{\mathrm{D}} P^B_0 + P^A_{\mathrm{S}} P^B_{\mathrm{S}}) \\
&+ H^{(0)}_{AB}(\hat{F}_B) \qquad (48) \\
&= P^A_0 P^B_0 \hat{F}_A P^A_0 P^B_0 \\
&+ P^A_0 P^B_{\mathrm{S}} \hat{F}_A P^A_0 P^B_{\mathrm{S}} \\
&+ P^A_{\mathrm{S}} P^B_0 \hat{F}_A P^A_{\mathrm{S}} P^B_0 \\
&+ P^A_0 P^B_{\mathrm{S}} \hat{F}_A P^A_{\mathrm{S}} P^B_{\mathrm{S}} \\
&+ P^A_{\mathrm{S}} P^B_0 \hat{F}_A P^A_{\mathrm{D}} P^B_0 \\
&+ P^A_{\mathrm{D}} P^B_0 \hat{F}_A P^A_{\mathrm{S}} P^B_0 \\
&+ P^A_{\mathrm{S}} P^B_{\mathrm{S}} \hat{F}_A P^A_0 P^B_{\mathrm{S}} \\
&+ P^A_0 P^B_{\mathrm{D}} \hat{F}_A P^A_0 P^B_{\mathrm{D}} \\
&+ P^A_{\mathrm{D}} P^B_0 \hat{F}_A P^A_{\mathrm{D}} P^B_0 \\
&+ P^A_{\mathrm{S}} P^B_{\mathrm{S}} \hat{F}_A P^A_{\mathrm{S}} P^B_{\mathrm{S}} \\
&+ H^{(0)}_{AB}(\hat{F}_B) \qquad (49)
\end{aligned}$$

Comparing equation 44 and equation 49 we find that

$$\begin{aligned}
H^{(0)}_{AB}(\hat{F}_A) &= H^{(0)}_A \mathrm{S}_B \\
&+ P^A_0 P^B_{\mathrm{S}} \hat{F}_A P^A_{\mathrm{S}} P^B_{\mathrm{S}} \\
&+ P^A_{\mathrm{S}} P^B_{\mathrm{S}} \hat{F}_A P^A_0 P^B_{\mathrm{S}} \qquad (50)
\end{aligned}$$

So unless the terms that couple the reference function and the singles of a subsystem over the N-electron Fock operator are zero this approach to perturbation theory is not strictly size consistent. Note that $P^A_0 P^B_{\mathrm{S}}$ and $P^A_{\mathrm{S}} P^B_{\mathrm{S}}$ are projectors onto singles and doubles of the total system, respectively. Therefore, it is not unreasonable that such terms may arise in a zeroth-order Hamiltonian based on projection operators that encompass both single and double excitations.

4.3.3 A nearly diagonal size consistent zeroth-order Hamiltonian

The results from the last two sections are consistent with the fact that single-reference Møller-Plesset perturbation theory is size consistent for closed shell systems. In that case the orbitals are eigenfunctions of the Fock-operator so that the terms that disturb the size consistency in $H^{(0)}_{S\cup D}$ will become zero.

The important result from the above derivations concerns the choice of the projection operators to construct the zeroth-order Hamiltonian from. It was found that if the projection operator on a subspace encompasses more than one excitation level then there may be some cross product terms in $H^{(0)}$ that destroy strict size consistency. For example, in $H^{(0)}_{S\cup D}$ the cross product terms of the reference function and the singly excited states over the Fock operator destroyed the size consistency. With this result we can understand how $H^{(0)}_{S,D}$ may be simplified without losing the size consistency. If we assume that $\Psi^{(0)}$ is a CAS reference function, the subspaces that make up X_1 and X_2 are equivalent to the subspaces $V_K, V_A, \ldots, V_H$ that Andersson *et al.* [17] defined except that we distinguish between singles and doubles within these subspaces also. If the projection operators in $H^{(0)}_{S,D}$ are split up a little more then the size consistency should be retained while obtaining a more diagonal zeroth-order Hamiltonian.

Denoting the subspaces using the excitation class notation of Ruttink *et al.* [53] and the excitation levels as defined by Wolinski *et al.* [44] we get fourteen subspaces $V_{(\#holes,\#particles)excitation}$ listed in table **I** where $\#holes$ is the number of holes in the doubly occupied orbitals, $\#particles$ is the number of particles in the virtual orbitals, and $excitation$ is the excitation level according to the recipe of Wolinski *et al.* This notation is more semantic than the one proposed by Andersson.

The subspaces are invariant for unitary transformations among the inactive orbitals, among the active orbitals, and among the virtual orbitals. Therefore a zeroth-order Hamiltonian defined as

$$
\begin{aligned}
H^{(0)} &= \hat{P}_{(0,0)0} F \hat{P}_{(0,0)0} \\
&+ \sum_{i=0}^{1}\sum_{j=0}^{1} \hat{P}_{(i,j)\mathrm{S}} F \hat{P}_{(i,j)\mathrm{S}} \\
&+ \sum_{i=0}^{2}\sum_{j=0}^{2} \hat{P}_{(i,j)\mathrm{D}} F \hat{P}_{(i,j)\mathrm{D}} \qquad (51)
\end{aligned}
$$

Andersson *et al.*		This paper
internal	V_0	$V_{(0,0)0}$
	V_K	$V_{(0,0)\mathrm{S}}$
		$V_{(0,0)\mathrm{D}}$
		$V_{(0,0)\mathrm{TQ...}}$
	V_A	$V_{(1,0)\mathrm{S}}$
		$V_{(1,0)\mathrm{D}}$
	V_B	$V_{(2,0)\mathrm{D}}$
semiinternal	V_C	$V_{(0,1)\mathrm{S}}$
		$V_{(0,1)\mathrm{D}}$
	V_D	$V_{(1,1)\mathrm{S}}$
		$V_{(1,1)\mathrm{D}}$
	V_E	$V_{(2,1)\mathrm{D}}$
external	V_F	$V_{(0,2)\mathrm{D}}$
	V_G	$V_{(1,2)\mathrm{D}}$
	V_H	$V_{(2,2)\mathrm{D}}$

Table I: The excitation classes used to improve the zeroth-order Hamiltonian represented as a subdivision of the subspaces defined by Andersson *et al.* [17].

is invariant for these orbital rotations also. A more severe truncation would destroy these invariances which is unacceptable [52]. A less severe truncation leaves more types of interactions to be implemented and computed. In this sense the above zeroth-order Hamiltonian is optimal.

Since this zeroth-order Hamiltonian has a minimal number of off-diagonal terms it will be referred to as $H^{(0)}_{S,D-minimal}$. Essentially this approach (when applied to a CAS reference function) is equivalent to the CASPT2D defined by Andersson *et al.* [17] but now with separate projectors onto singles and doubles also. Disadvantages of the CASPT2D method may therefore show up if this $H^{(0)}$ is used. For example if an orbital is almost doubly occupied it may easily mix with the doubly occupied orbitals under the influence of a perturbing field leading to a discontinuous change in $H^{(0)}$. However if efficiency and easy implementation are required and the disadvantages referred to do not occur $H^{(0)}_{S,D-minimal}$ may be used.

The difference between the zeroth-order Hamiltonians discussed here will be examplified with a calculation on Beryllium dimer.

Method		E monomer (a.u.)	E dimer (a.u.)	δ_{sc} (a.u.)
	CASSCF	-14.6156077572	-29.2312155144	0.00E-00
	CASSDCI	-14.6333357883	-29.2665681785	1.03E-04
	CASSDCI+Dav.	-14.6333912609	-29.2667822456	2.76E-07
	CASCEPA	-14.6333894658	-29.2667790315	-9.98E-08
$H^{(0)}_{S\cup D}$	CASMP2	-14.6312047590	-29.2624100676	-5.50E-07
	CASMP3	-14.6329875254	-29.2659717121	3.34E-06
$H^{(0)}_{S,D}$	CASMP2	-14.6312015484	-29.2624030967	0.00E-00
	CASMP3	-14.6329906432	-29.2659812864	0.00E-00
$H^{(0)}_{S,D-minimal}$	CASMP2	-14.6312088309	-29.2624176618	0.00E-00
	CASMP3	-14.6329876830	-29.2659753660	0.00E-00

Table II: Size consistency of multi-reference Møller-Plesset perturbation theory tested on a system of two non-interacting beryllium atoms using the 6-311G* basisset where in the multi-reference case the reference space was a CAS in the $2s, 2p_x, 2p_y$, and $2p_z$ orbitals of each atom. The size consistency error δ_{sc} is defined as the energy calculated for the dimer minus twice the energy of the monomer.

4.4 Results

To demonstrate the performance of $H^{(0)}_{S\cup D}$, $H^{(0)}_{S,D}$, and $H^{(0)}_{S,D-minimal}$ we studied a system of two Beryllium atoms. The 6-311G* basis set was used and the $2S$ and $2P_{x,y,z}$ orbitals were chosen as active orbitals to construct the CAS reference space. All orbitals were correlated. The size consistency error was calculated by subtracting twice the atomic energy from the dimer energy at a large separation of 1000 bohr. The results we obtained using ATMOL [54–57] are listed in table **II**. In this table the energies of the corresponding CASCI, CASCI corrected for the size consistency error with Davidson's correction [58], and CASCEPA [53] are given also. The latter energies are an indication of the performance of alternatives to perturbation theory.

In accordance with the proof given in the above sections $H^{(0)}_{S\cup D}$ yields a non-size consistent result, whereas $H^{(0)}_{S,D}$ and $H^{(0)}_{S,D-minimal}$ are exactly size consistent.

4.5 Conclusions

We have shown that Møller-Plesset perturbation theory generalized to a multi-configurational reference function is strictly size consistent if an appropriate zeroth-order Hamiltonian is used. The essence is that if $H^{(0)}$ is defined using projection operators that encompass more than one excitation level cross-product terms over the N-electron Fock-operator may appear that spoil the size consistency. This is a general result that may be used as a guideline to devise various zeroth-order Hamiltonians that retain the size consistency in the multi-reference case. In particular such a zeroth-order Hamiltonian is given for multi-reference Møller-Plesset perturbation theory restricted to second and third order.

Acknowledgements

The authors would like to thank professor Pulay 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).

References

[1] K. A. Brueckner. *Phys. Rev.* **100,** 36, 1955.

[2] J. Goldstone. *Proc. Roy. Soc. (London)* **A239,** 267, 1957.

[3] B. H. Brandow. *Rev. Mod. Phys.* **39,** (4) 771–828, 1967.

[4] R. J. Bartlett. *Ann. Rev. Phys. Chem.* **32,** 359–401, 1981.

[5] F. Coester and H. Kummel. *Nuc. Phys.* **17,** 477, 1960.

[6] J. Cizek. *J. Chem. Phys.* **45,** 4256, 1966.

[7] J. Cizek and J. Paldus. *Int. J. Quant. Chem.* **5,** 359, 1971.

[8] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley. *Int. J. Quant. Chem.* **14,** 545, 1978.

[9] R. J. Bartlett and G. D. Purvis. *Int. J. Quant. Chem.* **14,** 561, 1978.

[10] G. D. Purvis III and R. J. Bartlett. *J. Chem. Phys.* **76,** 1910, 1982.

[11] J. Noga, R. J. Bartlett, and M. Urban. *Chem. Phys. Lett.* **134,** 126, 1987.

[12] K. F. Freed. Tests and applications of complete active space quasidegenerate many-body perturbation theory for molecules. In U. Kaldor, editor, *Many-Body Methods in Quantum Chemistry*, volume 52 of *Lecture Notes in Chemistry*, page 1. Springer-Verlag, Berlin Heidelberg New York, 1989.

[13] P. M. Kozlowski and E. R. Davidson. *J. Chem. Phys.* **100,** (5) 3672, 1994.

[14] J. P. Finley, R. K. Chaudhuri, and K. F. Freed. *J. Chem. Phys.* **103,** (12) 4990, 1995.

[15] B. O. Roos, P. Linse, P. E. M. Siegbahn, and M. A. Blomberg. *Chem. Phys.* **66,** 197–207, 1982.

[16] K. Anderson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. *J. Chem. Phys.* **94,** 5483–5488, 1990.

[17] K. Andersson, P.-Å. Malmqvist, and B. O. Roos. *J. Chem. Phys.* **96,** 1218, 1992.

[18] K. Andersson. *Multiconfigurational Perturbation Theory*. PhD thesis, University of Lund, Theoretical Chemistry, Chemistry Center, P.O. Box 124, S-221 00 Lund, 1992.

[19] K. Andersson and B. O. Roos. *Chem. Phys. Lett.* **191,** 507, 1992.

[20] K. Andersson and B. O. Roos. *Int. J. Quant. Chem.* **45,** 591, 1993.

[21] B. O. Roos, L. Serrano-Andrés, and M. Merchán. *Pure & Appl. Chem.* **65,** 1693, 1993.

[22] K. Andersson and B. O. Roos. Multiconfigurational second-order perturbation theory. In R. Yarkony, editor, *Modern electron structure theory*, volume 1. World Scientific Publishing, New York, 1994.

[23] P.-A. Malmqvist. The rasscf, rassi, and caspt2 methods used on small molecules of astrophysical interest. In U. G. Jorgensen, editor, *Molecules in the Stellar Environment*, volume 428 of *Lecture Notes in Physics*, page 338. Springer-Verlag, Berlin Heidelberg New York, 1994.

[24] R. Pou-Amérigo, M. Merchán, I. Nebot-Gil, P.-Å. Malmqvist, and B. O. Roos. *J. Chem. Phys.* **101,** 4893, 1994.

[25] B. O. Roos. The casscf/caspt2 approach applied to the study of transition-metal compounds and their properties. In P. J. C. Aerts, R. Broer, and P. S. Bagus, editors, *New Challenges in Computational Quantum Chemistry*, Lecture Notes in Physics. Univ. of Groningen, Dept. of Chem. Phys. and Mat. Sci. Centre., Nijenborgh 4, NL-9747 AG Groningen, The Netherlands, 1994.

[26] K. Andersson, B. O. Roos, P.-Å. Malmqvist, and P.-O. Widmark. *Chem. Phys. Lett.* **230,** 391, 1994.

[27] K. Andersson. *Theor. Chim. Acta.* **91,** 31–46, 1995.

[28] K. Andersson. *Chem. Phys. Lett.* **237,** 212, 1995.

[29] J. Lorentzon, P.-Å. Malmqvist, M. P. Fülscher, and B. O. Roos. *Theor. Chim. Acta.* **91,** 91–108, 1995.

[30] C. F. Bender and E. R. Davidson. *J. Chem. Phys.* **47,** 4972, 1967.

[31] C. F. Bender and E. R. Davidson. *Phys. Rev.* **183,** 23, 1969.

[32] E. R. Davidson. *J. Chem. Phys.* **46,** 3320, 1967.

[33] E. R. Davidson and C. F. Bender. *Chem. Phys. Lett.* **59,** 369, 1978.

[34] B. Huron, J. P. Malrieu, and P. Rancurel. *J. Chem. Phys.* **58,** (12) 5745–5759, 1973.

[35] S. Evangelisti, J. P. Daudey, and J. P. Malrieu. *Chem. Phys.* **75,** 91, 1983.

[36] J. J. W. McDouall, K. Peasley, and M. A. Robb. *Chem. Phys. Lett.* **148,** 183, 1988.

[37] W. Duch and G. H. F. Diercksen. *Phys. Rev. A* **46,** 95, 1992.

[38] J. B. Foresman, M. Head-Gordon, and J. A. Pople. *J. Phys. Chem.* **96,** 135, 1992.

[39] K. Hirao. *Int. J. Quant. Chem.* **S26,** 517, 1992.

[40] K. Hirao. *Chem. Phys. Lett.* **190,** 374, 1992.

[41] K. Hirao. *Chem. Phys. Lett.* **196,** 397, 1992.

[42] K. Hirao. *Chem. Phys. Lett.* **201,** 59, 1993.

[43] K. G. Dyall. *J. Chem. Phys.* **102,** 4909, 1995.

[44] K. Wolinski, H. L. Sellers, and P. Pulay. *Chem. Phys. Lett.* **140,** 225, 1987.

[45] K. Wolinski and P. Pulay. *J. Chem. Phys.* **90,** (7) 3647, 1989.

[46] K. Wolinski and P. Pulay. *Theor. Chim. Acta.* **82,** 459, 1992.

[47] R. B. Murphy and R. P. Messmer. *Chem. Phys. Lett.* **183,** (5) 443–448, 1991.

[48] R. B. Murphy and R. P. Messmer. *J. Chem. Phys.* **97,** 4170, 1992.

[49] R. B. Murphy and R. P. Messmer. *J. Chem. Phys.* **92,** 4974, 1992.

[50] H.-J. Werner. *Mol. Phys.* **89,** 645–661, 1996.

[51] J. Paldus. *Theoret. Chem. Advan. Perspect.* **,** 131, 1976.

[52] P.-A. Malmqvist. *Chem. Phys. Lett.* **241,** 429–431, 1995.

[53] P. J. A. Ruttink, J. H. van Lenthe, R. Zwaans, and G. C. Groenenboom. *J. Chem. Phys.* **94,** (11) 7212–7220, 1991.

[54] V. R. Saunders and M. F. Guest. *ATMOL.* Daresbury, 1974.

[55] V. R. Saunders and J. H. van Lenthe. *Mol. Phys.* **48,** 923, 1983.

[56] P. Knowles and H.-J. Werner. *MULTI*. Cambridge University, Cambridge, 1984.

[57] J. H. van Lenthe, P. Pulay, and H. J. J. van Dam. *Multi-reference Møller-Plesset perturbation theory option*. Utrecht University, Fayetteville, Arkansas / Utrecht, The Netherlands, 1989-1996.

[58] S. R. Langhoff and E. R. Davidson. *Int. J. Quant. Chem.* **8,** 61, 1974.

Chapter 5
Exact size consistency of multi-reference Møller-Plesset perturbation theory

Practice

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

Peter Pulay
Department of Chemistry
University of Arkansas

Submitted to Molecular Physics (1997)

Abstract

The size consistency of multi-reference Møller-Plesset perturbation theory as a function of the structure of the zeroth-order Hamiltonian is studied. In calculations it is shown that the choice of projection operators to define the zeroth-order Hamiltonian is crucial. In essence whenever such a projection operator can be written as the sum of projection operators onto particular subspaces, cross-product terms may appear in the zeroth-order Hamiltonian that spoil the size consistency. This problem may be solved using a separate projection operator for each subspace spanning an excitation level. In principle a zeroth-order Hamiltonian based on these projection operators results in a size consistent perturbation theory. However, it was found that some non-local spin recoupling effects remain. A new zeroth-order Hamiltonian formulated recently circumvents this problem and is shown to be exactly size consistent.

Apart from the choice of projection operators, the orthogonalisation of the excited states is crucial also. It was found that modified Gramm-Schmidt in quadruple precision was not sufficient. A pivotted Householder QR factorisation (in double precision) offered the numerical stability needed to obtain size consistent results.

5.1 Introduction

Single-reference Møller-Plesset [1] perturbation theory (MP) is well known as an efficient and size extensive [2,3] way to treat electron correlation. Ever since it was recognised [4] that size extensivity is an essential prerequisite to allow consistent comparison of energies of molecules of different sizes this method has become widely used. However, for the calculation of potential energy surfaces or low lying excited states a single determinant is an inadequate zeroth-order description. Therefore, much research has been invested to develop a multi-reference approach, where the reference space contains all configurations needed for a proper zeroth-order order description (see [5-7] and the references therein). However, multi-reference perturbation theory did not become very popular until recently the so called first-diagonalise-then-perturb methods were efficiently implemented. In this approach the total Hamiltonian in the reference space is diagonalised to obtain the zeroth-order wavefunction. This multi-configurational reference function is then corrected using perturbation theory. In particular the variant where the reference space is a complete active space (CASMP) has become rather popular. This is mainly due to the overwhelming amount of work done by Roos and Andersson *et al*. [8-17], but there have been significant contributions by Wolinski *et al*. [18-20], Murphy *et al*. [21], Dyall [22], and most recently by Werner [23].

Although, MRMP is very efficient at low orders like the single-reference approach, there is discussion about the size consistency of the method. Wolinski claims that his approach is size consistent because it is an exact extension of canonical MP. Werner documented that Malmqvist pointed out that the zeroth-order Hamiltonian that Werner and Andersson use cannot be strictly size consistent because of the projection operators. Davidson [24] wrote referring to Huron *et al*. [25] that multi-reference perturbation theory cannot be size consistent because no multi-determinantal wavefunction can be the eigenfunction of a Hamiltonian containing only one-body operators. Before Huron *et al*. Claverie *et al*. ([26] p. 761) stated that one of the conditions of the linked cluster theorem is that the transition energy to a multi-excited state must be expressed as a sum of monoelectronic transition energies according to all possible ways of constructing the pluriexcited state. This may be construed as to imply that since it is improbable that this condition can be met using a contracted multi-configurational reference state and contracted excited states, the linked cluster theorem cannot hold for a multi-reference perturbation treatment. So the literature mainly suggests that although size consistency is very desirable it is not retained in the multi-reference case.

A proper study of the size consistency should consider both a mathematical treatment of the subject and examine the possibility of obtaining size consistency in practice. Here we show that these size consistent results may actually be obtained. The mathematical proof that size consistency may be obtained is given elsewhere [27]. Indeed as Malmqvist suggested (in Werner [23]) the projection operators used to define the zeroth-order Hamiltonian play a crucial role. However, we have shown [27] that these projection operators need not hamper the size consistency if they are chosen appropriately. The essential result is that projection operators onto subspaces that encompass more than one excitation level may lead to cross product terms over the Fock-operator that spoil the size consistency. Formally, projecting onto subspaces of different excitation levels separately is sufficient to construct a zeroth-order Hamiltonian to obtain size consistency. However, there still is a problem involved with non-local effects due to spins. A new more diagonal zeroth-order Hamiltonian circumvents this problem and will yield size consistent results in practice. In this article we compare the additivity of CASMP energies using a number of different zeroth-order Hamiltonians. It is shown that the size consistency errors vary considerably with the choice of these Hamiltonians and even may be nullified.

To actually nullify the size consistency errors in practice requires very accurate calculations. In particular the orthogonalisation of the initial contracted excited states to obtain an orthonormal space of contracted states requires a high precision algorithm. Numerical errors in this step in the calculation change the basis in which the perturbation theory is applied and may lead to different results. This problem is amplified by the fact that numerical errors differ for calculations of different size. As a result it becomes impossible to distinguish between systematic errors in the method and numerical errors. Therefore we will comment on the orthogonalisation method we used to minimize numerical errors.

5.2 Theory

In this paper we restrict ourselves to third order MRMP. In practice this means that only the single and double excitations are needed to evaluate the energy expressions. At higher orders the triples, quadruples, etc. come in and the dimension of the problem in general becomes prohibitive. To third order the perturbation expressions are

$$\begin{aligned}
&(H_0 - E_0)\Psi_0 = 0\\
&(H_0 - E_0)\Psi_1 = E_1\Psi_0 - H\Psi_0\\
&E_0 = \langle\Psi_0|H_0|\Psi_0\rangle\\
&E_1 = \langle\Psi_0|H|\Psi_0\rangle\\
&E_2 = \langle\Psi_0|H|\Psi_0 + \Psi_1\rangle\\
&E_3 = \langle\Psi_0 + \Psi_1|H|\Psi_0 + \Psi_1\rangle - E_1\langle\Psi_1|\Psi_1\rangle
\end{aligned} \tag{1}$$

where Ψ_0 is the multi-configurational reference function.

The total Hamiltonian is partitioned as usual as

$$H = H_0 + V \tag{2}$$

and we have assumed that the wavefunction is intermediately normalised, i.e.

$$\langle\Psi_0|\Psi_0 + \Psi_1\rangle = 1. \tag{3}$$

For more details see, e.g. [20].

In a MP approach the zeroth-order Hamiltonian H_0 is usually derived from the N-electron Fock-operator F defined as

$$\hat{F} = \sum_{rs} f_{rs}\hat{E}_{rs} \tag{4}$$

where the 1-electron Fock-operators are defined in matrix representation as

$$f_{rs} = h_{rs} + \sum_{ij}\langle\Psi_0|\hat{E}_{ij}|\Psi_0\rangle\left[(rs|ij) - \frac{1}{2}(ri|js)\right] \tag{5}$$

and the $\hat{E}_{xy}$ are the unitary group generators, cf. [23].

The theory requires that Ψ_0 is an eigenfunction of H_0; In single-reference MP this condition is automatically satisfied. In the multi-reference case this can be arranged through projection operators onto subspaces of the total first order interacting space. The simplest way to do this is to use an projection operator onto Ψ_0 and one onto the orthogonal complement of Ψ_0, leading to

$$H_0 = |\Psi_0\rangle\langle\Psi_0|F|\Psi_0\rangle\langle\Psi_0| + \left(I - |\Psi_0\rangle\langle\Psi_0|\right)F\left(I - |\Psi_0\rangle\langle\Psi_0|\right). \tag{6}$$

Restricted to singles and doubles only this is essentially what Werner uses [23].

To define the other zeroth-order Hamiltonians we distinguish between doubly occupied, active and virtual orbitals. The doubly occupied orbitals are doubly occupied in all reference states, the active orbitals may have different occupations for different reference states, and the virtual orbitals are empty in all reference states. We denote projection operators as $P_{(\#\text{holes},\#\text{particles})\text{excitation level}}$ where #holes is the number of electrons removed from the doubly occupied orbitals, #particles is the number of electrons in the virtual orbitals, and the excitation level is the number of single-electron substitution operators needed to create the states in the subspace. Note that the higher excited states are orthogonalised to the lower ones, in accordance with the recipe by Wolinski [18,20]. The projection operators defined in this manner are invariant to orbital rotations that leave Ψ_0 invariant. This is essential to define a H_0 that will yield a Ψ_1 that is independent of the particular orbital representation of Ψ_0. In terms of these projection operators the projection operators used by Wolinski may be defined as

$$\begin{aligned}
P_0 &= P_{(0,0)0} \\
P_\text{S} &= P_{(0,0)\text{S}} + P_{(0,1)\text{S}} + P_{(1,0)\text{S}} + P_{(1,1)\text{S}} \\
P_\text{D} &= P_{(0,0)\text{D}} + P_{(0,1)\text{D}} + P_{(1,0)\text{D}} + P_{(1,1)\text{D}} \\
&\quad + P_{(0,2)\text{D}} + P_{(1,2)\text{D}} + P_{(2,0)\text{D}} + P_{(2,1)\text{D}} + P_{(2,2)\text{D}}
\end{aligned} \quad . \tag{7}$$

The projection operators used by Andersson, who employs a projector P_K onto the orthogonal complement of the reference function within the CAS reference space, may be expressed as

Table I: Definition of projection operators onto subspaces of the total *N*-electron basis represented as sums of projection operators on mutually orthogonal minimal invariant subspaces.

	Andersson	Werner	Wolinski	This article
internal	P_0	P_0	P_0	$P_{(0,0)0}$
	P_K	$(I\text{-}P_0)$	P_S	$P_{(0,0)S}$
	P_K	$(I\text{-}P_0)$	P_D	$P_{(0,0)D}$
	P_K	$(I\text{-}P_0)$		$P_{(0,0)TQ...}$
	P_{SD}	$(I\text{-}P_0)$	P_S	$P_{(1,0)S}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(1,0)D}$
		$(I\text{-}P_0)$		$P_{(1,0)TQ...}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(2,0)D}$
		$(I\text{-}P_0)$		$P_{(2,0)TQ...}$
semiinternal	P_{SD}	$(I\text{-}P_0)$	P_S	$P_{(0,1)S}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(0,1)D}$
		$(I\text{-}P_0)$		$P_{(0,1)TQ...}$
	P_{SD}	$(I\text{-}P_0)$	P_S	$P_{(1,1)S}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(1,1)D}$
		$(I\text{-}P_0)$		$P_{(1,1)TQ...}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(2,1)D}$
		$(I\text{-}P_0)$		$P_{(2,1)TQ...}$
external	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(0,2)D}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(1,2)D}$
	P_{SD}	$(I\text{-}P_0)$	P_D	$P_{(2,2)D}$

The projection operators by Andersson *et al.* are defined in [9,11], the ones by Werner come from [23], and those by Wolinski *et al.* are taken from [18]. The projection operators used in this article are denoted using the excitation class notation by Ruttink *et al.* [33] and the generalised excitation level definition by Wolinski *et al.* [18,20].

$$\begin{aligned}P_0 &= P_{(0,0)0}\\ P_K &= P_{(0,0)S} + P_{(0,0)D} + P_{(0,0)TQ\ldots}\\ P_{SD} &= P_{(0,1)S} + P_{(1,0)S} + P_{(1,1)S} + P_{(0,1)D} + P_{(1,0)D} + P_{(1,1)D}\\ &\quad + P_{(0,2)D} + P_{(1,2)D} + P_{(2,0)D} + P_{(2,1)D} + P_{(2,2)D}\end{aligned} \tag{8}$$

P_{SD} projects onto the complete singly and doubly excited space outside the reference space.

Table II: Definitions of zeroth-order Hamiltonians.

Andersson	$H_0 = P_0FP_0 + P_KFP_K + P_{SD}FP_{SD}$	$H_0(1)$
Werner	$H_0 = P_0FP_0 + (I-P_0)F(I-P_0)$	
Wolinski	$H_0 = P_0FP_0 + P_SFP_S + P_DFP_D$	$H_0(2)$
This article	$H_0 = P_0FP_0 + \sum_{i,j=0}^{1} P_{(i,j)S}FP_{(i,j)S} + \sum_{i,j=0}^{2} P_{(i,j)D}FP_{(i,j)D}$	$H_0(3)$

Instead of projecting onto sums of subspaces we propose to project onto the separate subspaces. I.e. we use the projections $P_{(i,j)k}$ operators defined here as such. All projection operators referred to in this paper are listed in table **I** and the zeroth-order Hamiltonians derived from them are defined in table **II**.

5.3 Orthogonalisation

The H_0's used by Wolinski and ourselves require that the singles and the doubles be treated separately. In practice this requires a very accurate orthogonalisation method to separate the singles and the doubles. We found that a repeated modified Gramm-Schmidt orthogonalisation was inadequate; Although it succeeds in producing orthogonal vectors it does not retain the same space. A symmetric or Löwdin orthogonalisation could not be used without adding a Gramm-Schmidt step to project the singles out of the doubles. A better orthogonalisation is offered by pivoted Householder QR [28,29]. The

Householder QR method first calculates a Q-R factorisation in which all the sub diagonal elements of the matrix R are zero. The pivots are calculated as the 2-norm of the column vectors. Each time a column vector is selected the dimension of the remaining column vectors is reduced by one and the pivots are adjusted correspondingly. To separate the singles from the doubles we first select pivots from the singles space until this space is exhausted, i.e. until no pivots greater than machine precision remain. Then the algorithm is allowed to take pivots from the doubles space. This procedure yields spaces orthogonal to machine precision with the singles and doubles separated.

5.4 Calculations

In this article we give results obtained using H_0's of the form proposed by Andersson denoted as H_0(1), Wolinski [H_0(2)], and our H_0 [H_0(3)]. Our program [30,31] cannot use the H_0 of Werner because of the TQ... states that appear. In Werner's implementation these states arise because the internal and semiinternal excited states are left uncontracted. Besides MRMP results we give the corresponding MRCI, MRCI with Davidsons size consistency correction (MRCI+Q) [32] and MRCEPA(0) [33] results also. The MRCI results are given because this method is a variational alternative that is close to the Full CI. The MRCI+Q and MRCEPA(0) methods are variants of this method corrected for the size consistency error that is inevitable in any truncated variational CI approach. Using Davidsons correction the energy is corrected a posteriori, whereas in the MRCEPA(0) method the correction is effected through shifting the diagonal elements of the total Hamiltonian during the diagonalisation process. All MRMP results were obtained using Householder QR orthogonalisation unless stated otherwise.

In this article we present a set of applications to test the size consistency of various Møller-Plesset approaches using different H_0's. The set includes closed shell and open shell valence CAS reference calculations, with and without frozen cores. Open shell calculations on high spin states and other spin states are given. Also, a calculation on a restricted (non-CAS) reference function is given.

We give the size consistency errors of CAS reference calculations on Be, C, N, O, and F. The Be system serves as an example of a dimer of closed shell systems. The calculations on C, N and F are applications to closed shell systems dissociated to open shell subsystems. The calculation on O is an example of the dissociation of an open shell system to open shell systems. In the calculations on C, N, O, and F mentioned so far, the core electrons are frozen at the MCSCF level and not correlated. This means there are no doubly occupied orbitals in the correlation calculation. As an example of a system with doubly occupied orbitals in the reference states we present all-electron valence-CAS reference calculations on C, *vide infra*.

Although CASMP may yield excellent results CAS reference calculations soon become prohibitively large. In such cases a method that may be applied to a smaller restricted reference space becomes desirable. Typically to obtain a size consistent zeroth-order description such a restricted reference space will be a direct product space. The size consistency for such a case is demonstrated for Carbon-dimer.

Because size consistency is not the only important aspect we finally compare spectroscopic constants for O_2 calculated with different H_0's. The results of these calculations are compared to the results of Werner.

In the beryllium calculations the 6-311G* basis set [34] was used, the active space was a CAS constructed from the 2S and $2P_{x,y,z}$ orbitals and all electrons were correlated.

The calculations on C, N, O and F were performed in the cc-pV5Z basis set [35] and the 2S and $2P_{x,y,z}$ were selected as active orbitals. In the CASMP calculations the 1S orbitals were not correlated, as in the calculations by Werner [23].

The all-electron calculation on C_2 was performed in the cc-pCVQZ basis set [36,37] . The CAS reference space was constructed taking the 2S and $2P_{x,y,z}$ as active orbitals.

To show the size consistency of our H_0 in case of non-CAS reference spaces we give the results of the following high-spin direct-product reference wavefunction for C_2.

$$\begin{aligned}\Psi_0 &= \left(\left[2S^2 2P_x^1 2P_y^1\right] + \left[2P_x^1 2P_y^1 3S^2\right]\right)_A \\ &\otimes\left(\left[2S^2 2P_x^1 2P_y^1\right] + \left[2P_x^1 2P_y^1 3S^2\right]\right)_B\end{aligned} \tag{9}$$

The 1S electrons were frozen at the MCSCF level.

All size consistency errors listed here are evaluated as the dimer energy at R=1000 bohr minus twice the atomic energy. The calculations were performed with the ATMOL [38] program package. The orbitals for the multi-configurational reference function were optimized using MULTI [39]. To minimize numerical errors the orbitals used in the dimer calculation were identical to the monomer orbitals.

The calculations of spectroscopic constants of O_2 were performed in the cc-pV5Z basis set using a CAS constructed from the 2S and $2P_{x,y,z}$ orbitals of each atom. The spectroscopic constants are obtained from a quartic polynomial fitted through five points, i.e. $R_{O\text{-}O}$ is 1.1076, 1.1576, 1.2076, 1.2576, and 1.3076 Å. The equilibrium geometries listed are the calculated minima of the polynomial fits.

5.5 Results and discussion

The CASMP results on Be, C, N, O and F are listed in tables **III** to **VII**.
To appreciate the importance of a proper orthogonalisation method, compare the CASMP2 energy obtained for beryllium dimer (table **III**) with $H_0(2)$ and Householder QR to the CASMP2 energy of -29.2624032987 a.u. we obtained using modified Gramm-Schmidt in quadruple precision (REAL*16). Note that there is a difference of 0.0000002020 a.u. due to using a different orthogonalisation method. This means that even for a small system as beryllium dimer a suboptimal choice for the orthogonalisation method may lead to errors in the energy in the order of 10^{-6} eV.

Turning to the size consistency errors we find that for N, O and F (tables **V**, **VI**, and **VII**) with $H_0(1)$ they are of the same order of magnitude as the size consistency errors reported by Werner. The deviations from Werners results are due to the fact that Werner used a slightly different H_0 because he left the internal and semiinternal excited states uncontracted and he did not separate the excitations in the reference space from the other excitations. Thus we may assume that Werner's approach yields results that are similar to those that may be obtained using $H_0(1)$ as far as size consistency is concerned.

Table III: Size consistency errors for beryllium

Method		Be (Hartree)	2 Be (Hartree)	δ_{SC} (eV)
	CASCI	-14.6333357882	-29.2665681785	0.00281353
	CASCI+Q	-14.6333912609	-29.2667822456	0.00000752
	CASCEPA	-14.6333894658	-29.2667790315	-0.00000272
$H_0(1)$	CASMP2	-14.6312046713	-29.2624098148	-0.00001285
	CASMP3	-14.6329909096	-29.2659805466	0.00003463
$H_0(2)$	CASMP2	-14.6312015484	-29.2624030967	0.00000000
	CASMP3	-14.6329906432	-29.2659812864	0.00000000
$H_0(3)$	CASMP2	-14.6312088309	-29.2624176618	0.00000000
	CASMP3	-14.6329876830	-29.2659753660	0.00000000

Table IV: Size consistency errors for carbon

Method		C (Hartree)	2 C (Hartree)	δ_{SC} (eV)
	CASCI	-37.7874077807	-75.5711102337	0.10082456
	CASCI+Q	-37.7903202833	-75.5815578313	-0.02495941
	CASCEPA	-37.7895516074	-75.5792308340	-0.00347261
$H_0(1)$	CASMP2	-37.7834200454	-75.5668875699	-0.00129194
	CASMP3	-37.7869026531	-75.5736360002	0.00460693
$H_0(2)$	CASMP2	-37.7832757985	-75.5665600908	-0.00023112
	CASMP3	-37.7870336873	-75.5740629298	0.00012095
$H_0(3)$	CASMP2	-37.7837910135	-75.5675820269	0.00000000
	CASMP3	-37.7871157268	-75.5742314536	0.00000000

Table V: Size consistency errors for nitrogen

Method		N	2 N	δ_{SC}
		(Hartree)	(Hartree)	(eV)
Werner	CASCI	-54.5249631497	-109.0430734829	0.18646994
	CASCI+Q	-54.5297471870	-109.0598731911	-0.01030788
	CASCEPA	-54.5291464443	-109.0581632276	0.00352816
	CASMP2	n.a.	n.a.	-0.0023
	CASMP3	n.a.	n.a.	0.0086
$H_0(1)$	CASMP2	-54.5244497299	-109.0489626776	-0.00172020
	CASMP3	-54.5253087196	-109.0503213249	0.00805748
$H_0(2)$	CASMP2	-54.5240400265	-109.0480918296	-0.00032045
	CASMP3	-54.5261216441	-109.0522389427	0.00011824
$H_0(3)$	CASMP2	-54.5247780502	-109.0495561004	0.00000000
	CASMP3	-54.5263172760	-109.0526345519	0.00000000

Table VI: Size consistency errors for oxygen

Method		O	2 O	δ_{SC}
		(Hartree)	(Hartree)	(eV)
Werner	CASCI	-74.9934860305	-149.9750162868	0.32532498
	CASCI+Q	-75.0006312491	-149.9992280562	0.05535859
	CASCEPA	-75.0011104249	-150.0014836721	0.02005912
	CASMP2	n.a.	n.a.	-0.00034
	CASMP3	n.a.	n.a.	0.00084
$H_0(1)$	CASMP2	-74.9928726807	-149.9857623220	-0.00046151
	CASMP3	-74.9961125598	-149.9921836092	0.00112953
$H_0(2)$	CASMP2	-74.9926815473	-149.9853717267	-0.00023489
	CASMP3	-74.9965406936	-149.9930787647	0.00007136
$H_0(3)$	CASMP2	-74.9931582774	-149.9863165548	0.00000000
	CASMP3	-74.9967136703	-149.9934273406	0.00000000

Note that going from $H_0(1)$ to $H_0(3)$ the zeroth-order Hamiltonian becomes more diagonal, i.e. more off-diagonal elements are set to zero due to the choice of projection operators. Correspondingly, from $H_0(1)$ to $H_0(2)$ the sizeconsistency error becomes smaller. From $H_0(2)$ to $H_0(3)$ the size consistency error is nullified. In case the system dissociates to two closed shell subsystems (see Be) separating the projection operators onto the singles and the doubles is sufficient to achieve size consistency. This result is in accordance with theory [27]. However, $H_0(2)$ does not yield exactly size consistent results in those cases were the system dissociates into two open shell systems, *vide infra*. $H_0(3)$, where the projection operators are split up even more, remedies the size consistency defect in all cases to within the numerical precision.

Table VII: Size consistency errors for fluorine

Method		F (Hartree)	2 F (Hartree)	δ_{SC} (eV)
	CASCI	-99.6504789882	-199.2852295243	0.42798219
	CASCI+Q	-99.6592172239	-199.3148576591	0.09732692
	CASCEPA	-99.6601018811	-199.3198878304	0.00859673
Werner	CASMP2	n.a.	n.a.	-0.00053
	CASMP3	n.a.	n.a.	0.00041
$H_0(1)$	CASMP2	-99.6527928661	-199.3055893018	-0.00009713
	CASMP3	-99.6554293798	-199.3108546469	0.00011191
$H_0(2)$	CASMP2	-99.6527426104	-199.3054873122	-0.00005691
	CASMP3	-99.6556344771	-199.3112681678	0.00002140
$H_0(3)$	CASMP2	-99.6529696147	-199.3059392294	0.00000000
	CASMP3	-99.6557216939	-199.3114433878	0.00000000

The CASCI method fails in the cases discussed leading to size consistency errors that may become three orders of magnitude larger than the least size consistent perturbation approach. Even MRCI+Q or MRCEPA(0) do not get the size consistency right. However it is remarkable that the MRMP3 results often fall in between the MRCI and the MRCI+Q, MRCEPA(0) results. This underlines the conclusion by Werner that CASPT3 is often close to the corresponding MRCI.

Table VIII: Size consistency errors for carbon with a direct product reference space

Method		C (Hartree)	2 C (Hartree)	δ_{SC} (eV)
	MRCI	-37.7806064437	-75.5541991685	0.19084820
	MRCI+Q	-37.7846103057	-75.5674426405	0.04837983
	MRCEPA	-37.7841858616	-75.5670072048	0.03712950
$H_0(1)$	MRMP2	-37.7707614170	-75.5415350695	-0.00033294
	MRMP3	-37.7770624650	-75.5540841876	0.00110863
$H_0(2)$	MRMP2	-37.7720040846	-75.5440154843	-0.00019905
	MRMP3	-37.7786078405	-75.5572121440	0.00009624
$H_0(3)$	MRMP2	-37.7721954366	-75.5443908731	0.00000000
	MRMP3	-37.7773852226	-75.5547704452	0.00000000

Table IX: Size consistency errors for carbon in a all electron calculation

Method		C (Hartree)	2 C (Hartree)	δ_{SC} (eV)
$H_0(1)$	CASMP2	-37.8321246112	-75.6642835663	-0.00093452
	CASMP3	-37.8382958139	-75.6765208005	0.00192726
$H_0(2)$	CASMP2	-37.8319925917	-75.6639934443	-0.00022478
	CASMP3	-37.8382868606	-75.6765693337	0.00011939
$H_0(3)$	CASMP2	-37.8325658306	-75.6651316611	0.00000000
	CASMP3	-37.8383050047	-75.6766100095	0.00000000

The high-spin all-electron calculations on carbon-dimer yield similar size consistency errors as the other calculations on carbon-dimer. These calculations demonstrate the efficiency of contracted MRMP also. In the configuration space generation a total of $62.5 \cdot 10^6$ configuration state functions (CSF) were generated. After contraction only 150 000 states remained. The preparations in the calculation took 1827 seconds of CPU time on a 90 MHz R8000 SGI PowerChallenge, solving the linear system and evaluating the CASMP2 energy additionaly took 23 seconds of CPU time. Finally performing one CI iteration to evaluate the CASMP3 energy took an extra 7446 seconds of CPU time. The fact that the CASMP3 is so expensive is partly due to our implementation were the contracted Ψ_1 is transformed back to the large CSF basis to perform the CI-iteration. Other implementations (e.g. Werner) may be more efficient. However even though our implementation may not be optimal it is obvious that a CASMP2 calculation is extremely efficient. Solving the non-diagonal system of linear equations takes only a small amount of the total CPU time and can hardly be considered a problem. However the CI-iteration to evaluate the CASMP3 energy is an expensive step. Nevertheless CASMP3 may be expected to be an order of magnitude more efficient than the corresponding CASCI. These results underline the efficiency of MRMP methods but stress the fact that approaches more involved than MRMP3 soon become prohibitively time consuming also.

Table X: Spectroscopic constants for O_2 ($X^3\Sigma_g^+$)

Method		$E(r_e)$ (Hartree)	r_e (Å)	ω_e (cm^{-1})	D_e (eV)
Werner	CASPT2	-150.163783	1.2126	1566.1	5.12
	CASPT3	-150.159563	1.2076	1590.9	4.89
$H_0(1)$	CASMP2	-150.1808568125	1.2117	1585.4	5.31
	CASMP3	-150.1741201684	1.2063	1584.0	4.95
$H_0(2)$	CASMP2	-150.1806326095	1.2116	1587.7	5.31
	CASMP3	-150.1742452718	1.2069	1579.6	4.93
$H_0(3)$	CASMP2	-150.1822085790	1.2127	1587.2	5.33
	CASMP3	-150.1731496065	1.2070	1583.9	4.89

The size consistency results for the direct-product reference space demonstrated for carbon-dimer are listed in table **VIII.** These results basically show the same charasteristics as the CASMP results. In particular $H_0(3)$ gives size consistent energies in this case also.

Finally the spectroscopic constants we found for O_2 are close to the values that Werner found. However, we used only five points of the potential energy surface whereas Werner calculated ten points. Furthermore the results found with different zeroth-order Hamiltonians are very close to each other. This means that spectroscopic constants do not seem to be very sensitive to the exact form of the zeroth-order Hamiltonian.

5.6 The size consistency error of $H_0(2)$ for open-shell systems

The reason that $H_0(2)$ fails when dissociating to open shell systems is a consequence of the unitary group excitation operators. To explain this, consider a system consisting of two identical non-interacting atoms with each three orbitals and three electrons. The orbitals 1, 2 and 3 reside on the one atom whereas the orbitals 4, 5 and 6 reside on the other atom. It is assumed that each atom is in a doublet spin state and the atoms together are coupled to a triplet state. The branching diagram for the dimer reference state is shown in figure **I**.

Figure I: Branching diagram for the reference function of a non-interacting dimer of doublet atoms with 3 electrons in 3 orbitals each.

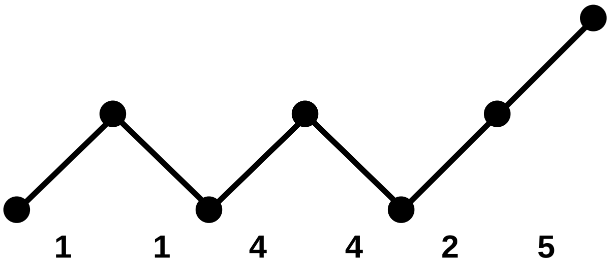
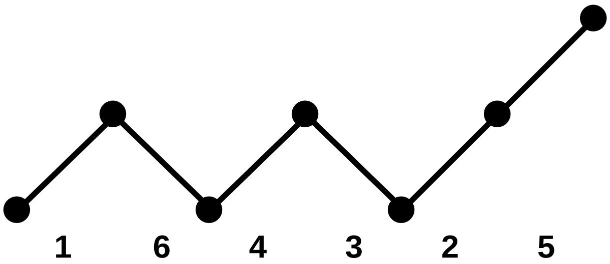


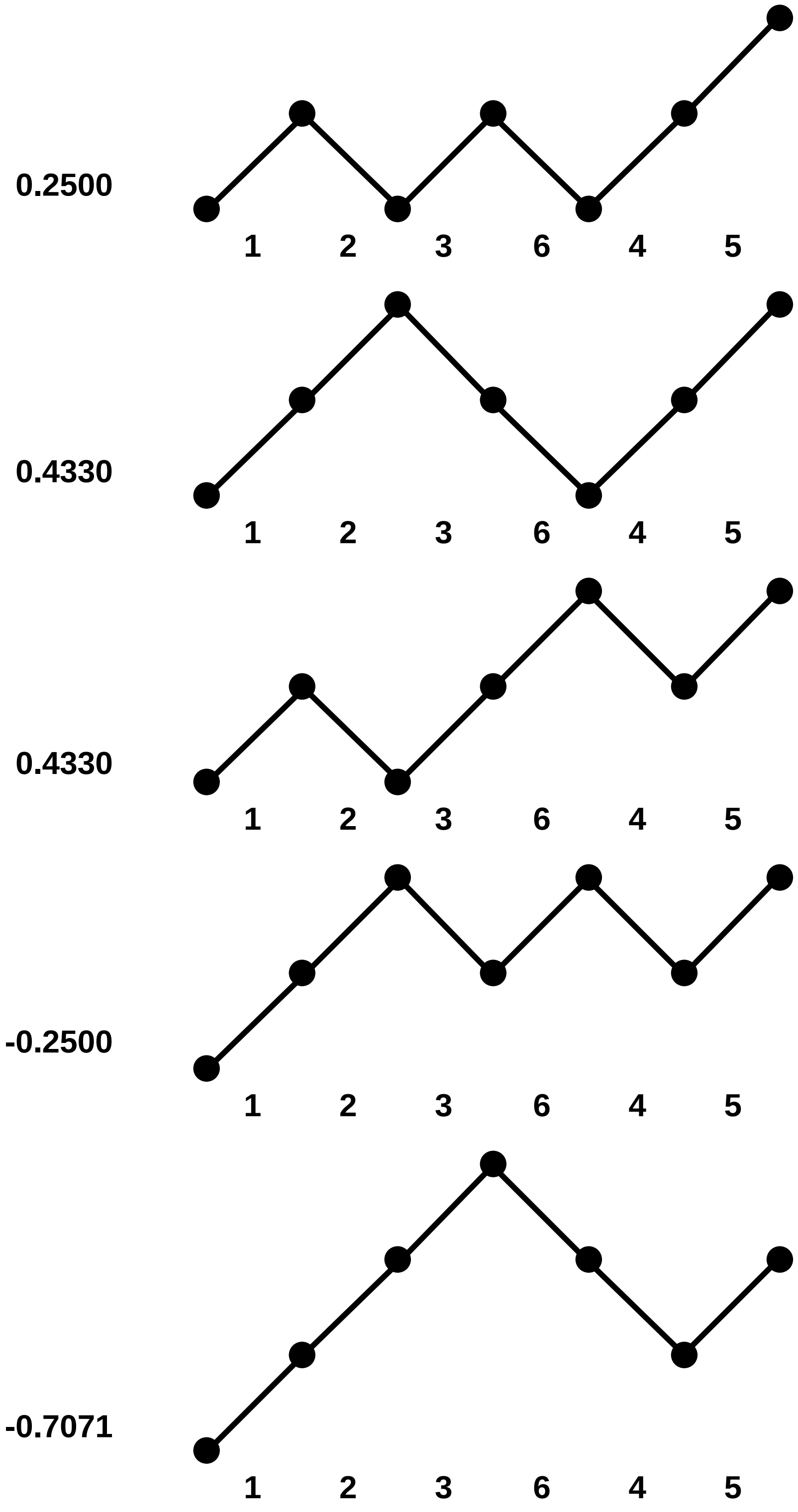
Figure II: Branching diagram for a doubly excited and spin recoupled state of a non-interacting dimer of doublet atoms with 3 electrons in 3 orbitals each.



Applying unitary group generators to build the excited state space, a doubly excited state is created in which two electrons are interchanged among the atoms. This state depicted in figure **II** is generated by exciting an electron from orbital 1 to orbital 6 and an exciting electron from orbital 4 to orbital 3.

Rearranging the orbitals in figure **II** to gather the orbitals on one atom, the linear combination shown in figure **III** is obtained.

Figure III: Branching diagram of figure **II** after reordering orbitals.



The last diagram in figure **III** is a quartet state on the first atom instead of a doublet. Therefore this excited state can not be represented in terms of two doublet monomers. So if this excited state gets a non-zero coefficient a wavefunction is obtained that can no longer be related to two doublet atoms, unless there are some other states that cancel the quartet contributions. Although this excited state does not have a direct interaction with the reference state over the Hamiltonian, it is however coupled with other doubly excited states over the N-electron Fock-operator. So it may get a non-zero coefficient in solving the linear system with H_0 to obtain Ψ_1 depending on the form of H_0. Therefore $H_0(2)$, where all couplings between states of the same excitation level are kept, may give non-size consistent results in open-shell cases. In $H_0(3)$ many couplings between doubly excited states have been removed, in particular those coupling the semiinternal and external excitations which seem to cause the problems. This way the size consistency problems may be circumvented. At the moment it is not clear yet if this problem can be solved for $H_0(2)$.

5.7 Conclusions

The size consistency error even for small diatomic systems may vary by as much as 6 orders of magnitude depending on the chosen H_0. Especially the use of $H_0(3)$ consistently gives size consistency errors within the numerical precision of our calculations. That means that $H_0(3)$ yields exactly size consistent energies. Although Werner stated that all the size consistency errors found with the perturbation treatment can be safely neglected for most practical applications we feel that exact size consistency is important for two reasons. First, size consistency is now recognised as an essential prerequisite for consistent comparison of molecular calculations of different size. This insight made perturbation theory more popular than restricted CI, even though CI yields an upper bound to the energy. Therefore size consistency is a fundamental motivation for generalising perturbation theory to the multi-reference case. So compromising on this aspect taints the quest for a proper multi-reference correlation treatment. Second, it is not clear how the size consistency errors found with non-size consistent H_0's scale with the size of the system. They may still be negligible for small diatomic systems but will they be negligible for larger systems?

We have shown that a zeroth-order Hamiltonian constructed using separate projector operators onto the singles and doubles may yield non-zero size consistency errors with open-shell dimers that dissociate to open-shell monomers. It was shown that these errors are due to spin-recouplings leading to non-local effects. The proof for size consistency which depends on the assumption that the dimer space can be expressed in terms of the spaces of the monomers does not apply in that case. It is not yet clear if this problem can be solved.

However, we have shown that a zeroth-order Hamiltonian that is as diagonal as possible without distroying the unitary invariance for important orbital rotations yields exactly size consistent results. The above mentioned spin problem is avoided using this H_0.

Also, the problem of orthogonalising large vector spaces that are strongly linearly dependent and meanwhile separating single and double excitations can be solved accurately. We have shown that a pivotted Householder QR factorisation performs this task to high precision. Therefore we see no reason for using a perturbation theory that is not strictly size consistent.

Acknowledgements

The investigations were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

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, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institue for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller, Karen Schuchardt, or Don Jones for further information.

References

[1] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[2] B. H. Brandow, Rev. Mod. Phys. **39**, 771 (1967).

[3] J. Goldstone, Proc. Roy. Soc. **A239**, 267 (1957).

[4] J. A. Pople, R. Krishnan, H. B. Schlegel and J. S. Binkley, Int. J. Quant. Chem. **14**, 545 (1978).

[5] R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).

[6] K. F. Freed, in: "Many-Body Methods in Quantum Chemistry" (Springer-Verlag, Berlin Heidelberg New York, 1989).

[7] P. M. Kozlowski and E. R. Davidson, J. Chem. Phys. **100**, 3672 (1994).

[8] B. O. Roos, P. Linse, P. E. M. Siegbahn and M. A. Blomberg, Chem. Phys. **66**, 197 (1982).

[9] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[10] K. Andersson, Theor. Chim. Acta. **91**, 31 (1995).

[11] K. Andersson, P.-Å. Malmqvist and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[12] K. Andersson and B. O. Roos, Chem. Phys. Lett. **191**, 507 (1992).

[13] K. Andersson and B. O. Roos, Int. J. Quantum Chem. **45**, 591 (1993).

[14] K. Andersson, B. O. Roos, P.-A. Malmqvist and P.-O. Widmark, Chem. Phys. Lett. **230**, 391 (1994).

[15] B. O. Roos, L. Serrano-Andrés and M. Merchán, Pure & Appl. Chem. **65**, 1693 (1993).

[16] B. O. Roos, in: "New Challenges in Computational Quantum Chemistry," (Dept. of Chem. Phys. and Mat. Sci. Centre, Univ. of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands, 1994).

[17] B. O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot and M. Merchán, in: "Advances in Chemical Physics: New Methods in Computational Quantum Mechanics" (John Wiley & Sons, New York, 1995).

[18] K. Wolinski, H. L. Sellers and P. Pulay, Chem. Phys. Lett. **140**, 225 (1987).

[19] K. Wolinski, Theor. Chim. Acta. **82**, 459 (1992).

[20] K. Wolinski and P. Pulay, J. Chem. Phys **90**, 3647 (1989).

[21] R. B. Murphy and R. P. Messmer, Chem. Phys. Lett. **5**, 443 (1991).

[22] K. G. Dyall, J. Chem. Phys. **102**, 4909 (1995).

[23] H.-J. Werner, Mol. Phys. **89**, 645 (1996).

[24] E. R. Davidson, Chem. Phys. Lett. **241**, 432 (1995).

[25] B. Huron, J. P. Malrieu and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[26] P. Claverie, S. Diner and J. P. Malrieu, Int. J. Quant. Chem. **1**, 751 (1967).

[27] H. J. J. van Dam, J. H. van Lenthe and P. J. A. Ruttink, Chapter 4, submitted to Int. J. Quant. Chem., (1997).

[28] G. H. Golub and C. F. Van Loan, "Matrix Computations" (The Johns Hopkins University Press, Baltimore, London, 1993).

[29] G. Sleijpen, Utrecht University, Mathematical Institute, Personal communication (1997).

[30] J. H. van Lenthe, P. Pulay and H. J. J. van Dam, Multi-reference Møller-Plesset perturbation theory option (Utrecht University, Fayetteville, Arkansas / Utrecht, The Netherlands, 1989-1996).

[31] V. R. Saunders and J. H. v. Lenthe, Mol. Phys. **48**, 923 (1983).

[32] S. R. Langhoff and E. R. Davidson, Int. J. Quant. Chem. **8**, 61 (1974).

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

[34] R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[35] D. E. Woon and T. H. Dunning Jr. , J. Chem. Phys. **100**, 2975 (1994).

[36] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

[37] D. E. Woon and T. H. Dunning Jr., J. Chem. Phys. **103**, 4572 (1995).

[38] V. R. Saunders and M. F. Guest, ATMOL (Utrecht University, Daresbury, 1974).

[39] P. Knowles and H.-J. Werner, MULTI (Cambridge University, Cambridge, 1984).

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

J. Comput. Chem. **17** (3), 267-272 (1996)

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 H_2O and O_2 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

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

Therefore, if the desired eigenvector is approximated by a vector $|\mathrm{u}\rangle$ one is interested in a perturbation vector $|\delta\rangle$ that minimises $\mathrm{R}(|\mathrm{u}\rangle + |\delta\rangle)$. 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).

$$(\mathrm{A}-\lambda\mathrm{I})|\delta\rangle = -|\mathrm{r}\rangle \tag{2}$$

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

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

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

$$\left|\left(\mathrm{A}_{jj}-\lambda\right)\right| > \sum_{k\neq j}\left|\mathrm{A}_{jk}\right| \tag{3}$$

then equation (2) may be approximated by

$$(\mathrm{D}-\lambda\mathrm{I})|\delta\rangle = -|\mathrm{r}\rangle. \tag{4}$$

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 $|\mathrm{u}\rangle$ 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 $|\delta\rangle = -|\mathrm{u}\rangle$. 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 $|\mathrm{u}\rangle$ 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 $|\mathrm{u}\rangle$. However, to assure that $|\delta\rangle$ is an effective perturbation vector one should determine $|\delta\rangle$ under the *condition* that $|\delta\rangle$ is orthogonal to $|\mathrm{u}\rangle$. This may be effected by computing the perturbation vector from the projection of A onto the subspace orthogonal to $|\mathrm{u}\rangle$. Sleijpen *et al.* showed that this is equivalent to substituting the projected matrix

$$\mathrm{B} = \left(\mathrm{I} - |\mathrm{u}\rangle\langle\mathrm{u}|\right)\mathrm{A}\left(\mathrm{I} - |\mathrm{u}\rangle\langle\mathrm{u}|\right) \tag{5}$$

for A in equation (2) yielding

$$\left(\mathrm{B} - \lambda\mathrm{I}\right)|\delta\rangle = -|\mathrm{r}\rangle. \tag{6}$$

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

$$\left(\mathrm{D}_{\mathrm{B}} - \lambda\mathrm{I}\right)|\delta\rangle = -|\mathrm{r}\rangle \tag{7}$$

where D_{B} is the diagonal of B. This equation is analogous to equation (4). Moreover, if the matrix $\left(\mathrm{A} - \lambda\mathrm{I}\right)$ is strongly diagonally dominant then a unit vector is a good approximation to the eigenvector and to $|\mathrm{u}\rangle$. 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_2O 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 H_2O at 2.0 R_e.

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

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 Davidsons 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 $\left(\mathrm{B}-\lambda\mathrm{I}\right)$. 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 $"\mathrm{v}"$ and an external part with a subscript $"\mathrm{e}"$.

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

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

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

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

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

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

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

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, 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 $(\mathrm{B}-\lambda\mathrm{I})$ is performed in every iteration. These multiplications may be efficiently computed as follows. Assume the vacuum matrix $(\mathrm{B}-\lambda\mathrm{I})$ is to be multiplied with a vector $|\mathrm{q}\rangle$. Then, using equation (5) we have

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

Writing

$$\lambda = \langle\mathrm{u}|\mathrm{A}|\mathrm{u}\rangle,$$

defining

$$|\mathrm{z}\rangle = \mathrm{A}|\mathrm{u}\rangle,$$

and using that A is hermitian equation (13) becomes

$$(\mathrm{B}-\lambda\mathrm{I})|\mathrm{q}\rangle = \mathrm{A}|\mathrm{q}\rangle + |\mathrm{u}\rangle\left(\langle\mathrm{u}|\mathrm{q}\rangle\lambda - \langle\mathrm{z}|\mathrm{q}\rangle\right) - |\mathrm{z}\rangle\langle\mathrm{u}|\mathrm{q}\rangle - |\mathrm{q}\rangle\lambda\,. \tag{14}$$

Equation (14) involves a matrix-vector product with the original vacuum matrix A and some simple vector-vector operations. The vector $|\mathrm{z}\rangle$ 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 Ruttink *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_2O at various geometries in MRSDCI.

OH Bond Length (R_e)	Energy (Hartree)	Number 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 $\tilde{E} - E < 0.1\text{E} - 6$, where $\tilde{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

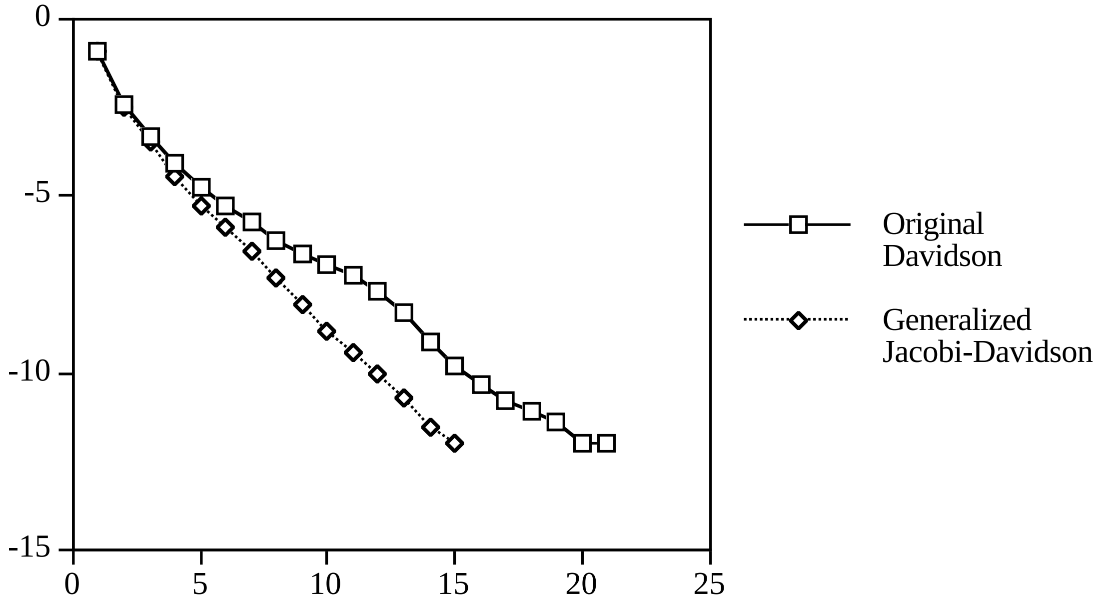
Table III: Comparison of the convergence of Davidson and the generalized Jacobi-Davidson for H_2O at various geometries in MRCEPA(0).

OH Bond Length (R_e)	Energy (Hartree)	Number of Iterations	
		Davidson	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 $3a_1$, $4a_1$, $1b_1$, $2b_1$, $1b_2$ and $2b_2$ orbitals. The convergence threshold is $\tilde{E} - E < 0.1\text{E} - 6$, 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 $(\mathrm{B}-\lambda\mathrm{I})$ 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 $(\mathrm{B}-\lambda\mathrm{I})$ 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 $(\mathrm{B}-\lambda\mathrm{I})$. 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



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 $(\mathrm{B}-\lambda \mathrm{I})$ 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 **I**. 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)=>[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 $3a_g$, $1b_{2g}$, $1b_{3g}$, $3b_{1u}$, $1b_{2u}$, and $1b_{3u}$ orbitals. The initial CI vector is computed as described for the H_2O test case. The MRCI and MRCEPA(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_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 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 O_2 at various geometries in MRSDCI.

O_2 Bond Length (R_e)	Energy (Hartree)	Number 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 $\tilde{E} - E < 0.1\mathrm{E} - 6$, where $\tilde{E}$ 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(0).

O_2 Bond Length (R_e)	Energy (Hartree)	Number 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 $\tilde{E} - E < 0.1\mathrm{E} - 6$, where $\tilde{E}$ is the current approximation to the eigenvalue and E is the fully converged eigenvalue.

References

[1] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).

[2] E. R. Davidson, Comput. Phys. Comm. **53**, 49 (1989).

[3] C. W. Murray, S. C. Racine and E. R. Davidson, J. Comput. 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, Preprint: 856, Dept. Math., Utrecht University (1994).

[6] E. Kreyszig, "Advanced Engineering Mathematics" (John Wiley & Sons, Inc., Singapore, 1993).

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

[8] R. J. Buenker, in: "Current Aspects of Quantum Chemistry 1981" (Elsevier Scientific Publishing Company, Amsterdam, 1982).

[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. Comput. Phys. **48**, 23 (1982).

[11] G. H. Golub and C. F. van Loan, "Matrix Computations" (North Oxford Academic, Oxford, 1986).

[12] C. W. Bauschlicher and P. R. Taylor, The Journal of Chemical Physics **85**, 2779 (1986).

[13] V. R. Saunders and M. F. Guest, ATMOL (Daresbury, 1974).

[14] F. B. v. Duijneveldt, Internal Report: RJ945, IBM Research San José (1971).

Chapter 7

Exoduction

This thesis is mostly concerned with MRMP, and some consideration is given to MRCI and MRCEPA(0). Looking back upon the previous chapters it is learnt that an efficient and size consistent multi-reference perturbation method for calculating correlation energies can be constructed. To achieve size consistency the form of the zeroth-order Hamiltonian has to be constructed from projection operators that project onto subspaces of a single excitation level only. In practice it is possible to define a whole family of zeroth-order Hamiltonians that meet this requirement. In particular a zeroth-order Hamiltonian with as few as possible off-diagonal terms may be chosen. Besides choosing a proper zeroth-order Hamiltonian the choice for the orthogonalisation method to construct the excited state space is essential. Perturbation theory proves to be rather sensitive to numerical errors in the orthogonalisation. Because the errors are different for calculations of different size they can lead to considerable size consistency errors.

Despite the fact that a size consistent multi-reference approach to correlation energy calculations has been obtained the correlation problem is not solved yet. It should be realised that perturbation theory is not stable enough to be used in all cases; On the other hand MRCI corrected for size consistency errors and MRCEPA(0) are only approximately size consistent. However they may be applicable in cases where perturbation theory fails.

Although not too much thought was given to the size consistency corrected MRCI (MRCI+Q) and the MRCEPA(0), the MRMP chapters allow for a comparison. For closed shell cases MRCI+Q and MRCEPA(0) yield small size consistency errors. However, in the open-shell cases the behaviour was rather poor. In every case MRCEPA(0) performed better with respect to size consistency than MRCI+Q. Regarding the stability of these methods it should be noted that the MRCI approaches as well as MRCEPA(0) converged in all calculations. This meant that MRCEPA(0) at the moment is the best alternative to MRMP we have available, and this method should be considered in cases where MRMP is hampered by divergences.

Reconsidering these results we may relate them to the list of requirements that was compiled in the introduction to this thesis.

1) A problematic point with perturbation theory is that it is not well defined for all chemically reasonable geometries. This problem was long known for single-reference perturbation theory. One might have hoped that this problem could be solved using a multi-reference approach since one expects that the perturbation would be smaller in such a treatment. However, the contrary may be the case.

   In calculations on O_2 the multi-reference approach yielded a larger difference between E_0 and E_1 than the single-reference treatment. So the perturbation is larger in the multi-reference case. Indeed, the multi-reference calculations are more strongly divergent than the single-reference ones.

   Although a method is suggested to detect divergences so that these points may be avoided in calculating potential energy surfaces, a MRMP study may be impossible if too many divergent points are present. MRCI is not subject to divergence problems. MRCEPA(0) shows to be well behaved in the calculations documented in this thesis. However, it is to be expected that MRCEPA(0) may diverge if an inadequate reference space is used.

2) MRMP, MRCI, and MRCEPA(0) are able to dissociate molecules properly if the reference wavefunction is chosen judiciously.

3) MRMP, MRCI, and MRCEPA(0) are invariant for orbital rotations that leave the reference wavefunction unchanged.

4) MRMP may be implemented in a such way that exactly size consistent results may be obtained. This is an important result because there was a growing tendency to believe that multi-reference approaches could not be size consistent except for some special cases. MRCI and MRCEPA(0) fail with respect to this requirement especially in open shell cases.

5) The MRMP approaches sometimes have problems in calculating energies of a similar quality at a wide range of geometries. For example MRMP2 may not be accurate enough in calculating dissociation energies. In most cases this could be mended in the next order of perturbation theory. However, one case was found in which the MRMP3 gave less accurate results than MRMP2 (see chapter 2, table **V**). With respect to dissociation energies MRCI and MRCEPA(0) behave very well.

6) Considering efficiency MRMP is at least an order of magnitude more efficient than MRCI or MRCEPA(0). MRMP3 costs about as much as a single CI-iteration. This means that for large molecules MRMP may be the only reasonable approach to take correlation energy into account. This was one of the strong motivations for Andersson *et al.* [1] to develop CASMP2.

Concluding, we have found that MRMP satisfies the requirements we listed except that it may diverge and the quality of the energy may be insufficient. The most appropriate alternative currently available, MRCEPA(0) fails in obtaining size consistency in open shell cases and it is relatively expensive. A reasonable scheme for calculating potential energy surfaces therefore seems to first try MRMP and then use MRCEPA(0) if MRMP fails. A problem with this scheme is in detecting whether the MRMP fails. Convergence problems in MRMP may be hard to detect at low orders. Inspection of the coefficients in Ψ_1 and the singles-doubles part of Ψ_2 should be helpful. When in doubt the convergence of the series should be tested.

7.1 Future developments

At the 9th International Conference on Quantum Chemistry (Atlanta, 1997) Morokuma explained that quantum chemistry has come a long way. Along this way a number of landmarks were passed, cf. [2]. These include the introduction of Slater determinants [3], the Hartree-Fock method [4,5], the use of Gaussian basis functions [6], the Roothaan parametrisation of the wavefunction [7], the generalisation of Hartree-Fock to the multi-configuration case [8], the development of multi-reference approaches like MRCI [9,10], and multi-reference perturbation theory (MRPT) [11,12]. The multi-reference coupled cluster (MRCC) method is underway. Finally, relativistic effects are being included in many codes, and much effort is being invested in coupling quantum-mechanical and classical models. Although this all seems very well it does not mean that we have come to the end of the road.

Looking at approaches to obtain a reasonable estimate of the correlation energy we find that there is a problem with all of them. For example MRCI is not size consistent, MRPT and MRCC may be hampered by divergences [13], and MRCI and MRCC are computationally expensive. Besides these problems, multi-reference methods in general have a tendency to lead to such large configuration spaces that they may not be applicable to large molecules without severe approximations. However, it is well known that dynamical correlation energy has to be accounted for to some extent. MCHF is insufficient to calculate potential energy surfaces accurate enough for qualitatively correct chemical predictions. At this stage the fundamental problem seems to be that everything that can be done using a determinantal basis has been tried, but the results are not really satisfactory.

Basically the paths that lead onwards lay within two extremes. The first extreme is to use the knowledge we have obtained so far and to develop approaches that allow to direct the computational power we have available to the heart of the problems of interest. Mainly this is a matter of tailoring the current methods to allow more flexibility in defining problems. For example, one may want to use MRCI to account for electron correlation in only a part of a molecule. In such a case localisation of orbitals, selections of orbitals and configurations, use of a different basis set for the MCHF calculation and the correlation treatment, and other advanced approximations will be needed. The basic methods will be the same but the new developments will be in defining meaningful approximations.

An other extreme is to develop methods that extend the current approaches to non-determinantal wavefunctions. This is not new as ventures in this direction were already performed by Hylleraas [14]. However, up to a few years ago there was always the promise that new many-configuration methods would cure the problems at hand. I think that is no longer true. This perspective suggests that the time has come to reconsider whether Slater determinants are a good idea. Essentially one would like to have an *ab initio* method (in contrast to DFT) that accounts for a significant part of the correlation energy while not being more involved than MCHF. This suggests to consider geminal type approaches to

MCHF of some sort. There have been some investigations on the use of geminals in many-configurational methods (cf. [15-28]) but the computational complexity limited their use. Recently, Persson *et al.* [29] proposed a new way to expand the electron interaction that will reduce the costs of these approaches considerably. Introducing explicitly correlated wavefunctions into MCHF type calculations may be an efficient alternative in describing potential energy surfaces.

In practice of course the final solution will be somewhere in between these extremes. For the next decade I think significant contributions to the latter approach may be expected. Thus although quantum chemistry has come a long way, it may still have a long way to go.

References

[1] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[2] H. F. Schaefer III, "Quantum Chemistry: The development of *ab initio* methods in molecular electronic structure theory" (Clarendon Press, Oxford, 1984).

[3] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[4] V. Fock, Z. Physik **61**, 126 (1930).

[5] D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

[6] S. F. Boys, Proc. Roy. Soc. (London) **A200**, 542 (1950).

[7] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[8] G. Das and A. C. Wahl, J. Chem. Phys. **44**, 87 (1966).

[9] P. E. M. Siegbahn, J. Chem. Phys. **72**, 1647 (1980).

[10] V. R. Saunders and J. H. v. Lenthe, Mol. Phys. **48**, 923 (1983).

[11] B. O. Roos, P. Linse, P. E. M. Siegbahn and M. A. Blomberg, Chem. Phys. **66**, 197 (1982).

[12] K. Wolinski, H. L. Sellers and P. Pulay, Chem. Phys. Lett. **140**, 225 (1987).

[13] L. Meissner, S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **91**, 6187 (1989).

[14] E. A. Hylleraas, Z. Physik **48**, 469 (1928).

[15] W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985).

[16] W. Klopper, Chem. Phys. Lett. **134**, 17 (1987).

[17] W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).

[18] W. Klopper, Chem. Phys. Lett. **186**, 583 (1991).

[19] W. Klopper and J. Almlöf, J. Chem. Phys. **99**, 5167 (1993).

[20] W. Cencek and J. Rychlewski, J. Chem. Phys. **98**, 1252 (1993).

[21] J. Rychlewski, W. Cencek and J. Komasa, Chem. Phys. Lett. **229**, 657 (1994).

[22] J. Komasa, W. Cencek and J. Rychlewski, Physical Review A. **52**, 4500 (1995).

[23] J. Komasa and J. Rychlewski, Chem. Phys. Lett. **249**, 253 (1996).

[24] K. Szalewics, B. Jeziorski, H. J. Monkhorst and J. G. Zabolitzky, Chem. Phys. Lett. **91**, 169 (1982).

[25] K. B. Wenzel, J. G. Zabolitzky, K. Szalewics, B. Jeziorski and H. J. Monkhorst, J. Chem. Phys. **85**, 3964 (1986).

[26] S. A. Alexander, H. J. Monkhorst and K. Szalewicz, J. Chem. Phys. **89**, 355 (1988).

[27] R. Bukowski, B. Jeziorski and K. Szalewicz, J. Chem. Phys. **100**, 1366 (1994).

[28] R. Bukowski, B. Jeziorski, S. Rybak and K. Szalewicz, J. Chem. Phys. **102**, 888 (1995).

[29] B. J. Persson and P. R. Taylor, J. Chem. Phys. **105**, 5915 (1996).

Summary

This thesis is about methods for electronic structure calculations on molecular systems. The ultimate goal is to construct methods that yield potential energy surfaces of sufficient accuracy to allow a qualitatively correct description of the chemistry of these systems; i.e. heat of formation, isomerisation barriers, equilibrium geometries, and vibrational spectra. In order to properly calculate the potential energy surfaces for all these properties a multi-configurational starting point is essential. This means that all methods that will be discussed are based on a multi-reference wavefunction where the reference function is optimised using the multi-configurational Hartree-Fock (MCHF) method.

Beyond the MCHF method there are various methods to account for the correlation energy of a molecule. Not all of these methods are equally suited to calculating potential energy surfaces. In chapter 1 the basic notions are introduced and a list of required qualities for a method is compiled. This list includes size consistency, independence of the orbital representation of the reference wavefunction, efficiency, and others. The methods discussed in this thesis will be checked against these requirements.

The main part of this thesis treats a perturbation method that was first formulated by Møller and Plesset (MP) in 1934. In its original formulation the method is applied starting from a single closed shell determinant. It was known that in this case the method is strictly size consistent. It was known also that the method diverges in cases when two states are close in energy. The basic idea at the start of the work described here was to avoid these divergences by including all nearly degenerate states in the reference function thus generalising the method to the multi-reference case. If this would be possible while retaining the size consistency a very efficient and highly accurate multi-reference Møller-Plesset method (MRMP) would be obtained.

In chapter 2 the implementation of this method for a general reference wavefunction is described. Although the test applications yielded encouraging results, a few results suggested divergences may still show up. In chapter 3 a method to detect divergences is proposed and it was applied to suspicious systems. It is found that the multi-reference perturbation theory may be more strongly divergent than the single-reference approach. Also, the multi-reference results were not exactly size consistent. A detailed study of this problem is

given in chapter 4 were it is concluded from theory that the method should be exactly size consistent. In chapter 5 the practical aspects involved in a size consistent multi-reference perturbation theory are described. The results show that a size consistent approach can be obtained. The crucial aspects are that the projection operators to construct the zeroth-order Hamiltonian should each project onto a subspace of a single excitation level, the orthogonalisation method to generate the orthonormal excited states should be highly accurate, and in open-shell calculations applying the unitary group generators twice is not enough to generate all required spin states.

Perturbation theory is not the only method that yields size consistent results. Already in the sixties it was known that some electron pair approximations yield exactly size consistent correlation energies also. In the single reference case it was shown that the coupled electron-pair approximation (CEPA) could be used to approximate coupled cluster in the singles-doubles configuration space. Ruttink et al. have generalised this approach to the multi-reference case (MRCEPA(0)). Although this approach is not exactly size consistent it is the best alternative to MRMP we have available. For this reason the MRMP results in this thesis are often compared to results obtained with MRCEPA(0).

At the heart of the MRCEPA(0) is the Davidson diagonalisation method that is used to iteratively solve the eigenvalue equations. The efficiency of the MRCEPA depends primarily on the rapid convergence of the Davidson method. Essentially, the Davidson method calculates the best approximation to the wavefunction from a given set of vectors. Through extending this set by one vector (the update vector) in every iteration convergence is guaranteed. The speed of convergence depends on how appropriate the update vectors are. However, Sleijpen and van der Vorst realised that if the method was applied exactly as suggested by Davidson it would never converge. A detailed analysis led to improvements enhancing the speed of convergence. The application of these improvements in quantum chemistry is discussed in chapter 6.

In chapter 7 the results from the main chapters are checked against the requirements list compiled in chapter 1. The conclusion is that although some requirements could be met, none of the methods satisfies all requirements. Because the alternatives employing a determinantal basis are nearly exhausted it is suggested that future developments should go in other directions, e.g. explicitly correlated wavefunctions.

Samenvatting

Dit proefschrift gaat over methoden voor electronenstructuur berekeningen aan moleculen. Het uiteindelijke doel is methoden te construeren die potentiële energie oppervlakken met voldoende nauwkeurigheid opleveren om kwalitatief juiste uitspraken te kunnen doen over de chemie van moleculaire systemen. Dit omvat vormingswarmten, isomerisatie barrières, evenwichtsgeometriën en vibrationele spectra. Om de potentiële energie oppervlakken correct te beschrijven met betrekking tot de genoemde eigenschappen is een meervoudige referentie aanpak vereist. Vandaar dat alle in dit proefschrift beschreven methoden gebaseerd zijn op een meervoudige referentie golffunctie, waarbij deze referentie functie met de meerconfiguratie Hartree-Fock (MCHF) methode geoptimaliseerd is.

Na een MCHF berekening kunnen meerdere methoden gebruikt worden om de electronencorrelatie energie in een molecule te berekenen. Niet al deze methoden zijn even geschikt voor de berekening van potentiële energie oppervlakken. In hoofdstuk 1 worden daarom de basale begrippen uiteengezet en wordt een lijst van eisen voor een geschikte methode opgesteld. Deze lijst bevat eigenschappen als grootte-consistentie, onafhankelijkheid van de orbitaal representatie van de golffunctie en efficiëntie. De methoden die in dit proefschrift aan bod komen worden tegen deze eisen afgezet.

Het grootste deel van dit proefschrift behandelt een storingsrekening methode die voor het eerst door Møller en Plesset (MP) geformuleerd werd in 1934. In de oorspronkelijke formulering werd de methode toegepast uitgaande van één gesloten schil determinant. Het was bekend dat de methode in dit geval strikt grootte consistent is. Ook was bekend dat de methode divergeert wanneer twee toestanden qua energie dicht bij elkaar liggen, in andere woorden bijna ontaard zijn. Het uitgangspunt bij het werk dat hier beschreven wordt was om de divergentie te vermijden door alle bijna ontaarde toestanden als referentie mee te nemen. Indien deze generalisatie met behoud van grootte consistentie uitgevoerd zou kunnen worden, zou een zeer efficiënte en accurate meervoudige referentie Møller-Plesset methode (MRMP) verkregen worden.

In hoofdstuk 2 wordt de implementatie van deze methode voor een algemene referentie functie beschreven. Hoewel de test berekeningen bemoedigende resultaten opleverden werden enkele resultaten verkregen die suggereerden dat ook de gegeneraliseerde methode kan divergeren. In hoofdstuk 3 wordt een methode voorgesteld om deze divergenties te detecteren. Deze methode wordt toegepast op enkele systemen die van divergenties verdacht worden. Gevonden wordt dat de generaliseerde aanpak zelfs sterker divergent kan zijn dan de oorspronkelijke aanpak. Een ander probleem dat ontdekt werd was dat de resultaten in eerste instantie niet grootte consistent bleken te zijn. Een gedetailleerde studie naar dit fenomeen is weergegeven in hoofdstuk 4. Uit de theorie wordt geconcludeerd dat de methode exact grootte consistent zou moeten zijn. In hoofdstuk 5 worden de praktische aspecten beschreven die voor het verkrijgen van grootte consistente resultaten van belang zijn. De resultaten laten zien dat een grootte consistente methode opgesteld kan worden. De essentiële aspecten zijn dat de projectie operatoren die gebruikt worden in de definitie van de nulde-orde Hamiltoniaan elk op een subruimte van slechts één excitatie niveau mogen projecteren, dat de orthogonalisatie methode om de ruimte van aangeslagen toestanden te genereren zeer nauwkeurig moet zijn, en dat in open schil systemen het twee maal toepassen van de unitaire groep generatoren niet voldoende is om alle benodigde spin toestanden te genereren.

Storingsrekening is niet de enige manier om grootte consistente resultaten te verkrijgen. Al in de jaren 60 was bekend dat sommige electronenpaar benaderingen exact grootte consistent zijn. Voor het enkelvoudige referentie geval was aangetoond dat de gekoppelde electronenpaar benadering (CEPA) gebruikt kon worden als benadering van de gekoppelde clusters aanpak in de ruimte van enkele en dubbele aangeslagen configuraties. Ruttink en medewerkers hebben deze methode gegeneraliseerd naar het meervoudige referentie geval (MRCEPA(0)). Hoewel deze methode niet strikt grootte consistent is vormt ze het beste alternatief voor MRMP dat we beschikbaar hebben. Om deze reden worden MRMP resultaten in dit proefschrift vaak vergeleken met MRCEPA(0) resultaten.

Belangrijk bij de MRCEPA(0) methode is de Davidson diagonalisatie methode die gebruikt wordt om iteratief eigenwaarde vergelijkingen op te lossen. De efficiëntie van MRCEPA(0) hangt af van de snelle convergentie van de Davidson methode. In essentie berekent de Davidson methode de beste benadering voor de golffunctie binnen een gegeven verzameling van vectoren. Door deze verzameling elke iteratie met één vector (de storingsvector) uit te breiden is convergentie gegarandeerd. De convergentie snelheid hangt af van de geschiktheid van de storingsvectoren. Sleijpen en van der Vorst wisten dat wanneer de methode zou worden toegepast precies zoals Davidsons voorstel suggereert, ze oneindig langzaam zou convergeren. Een gedetailleerde analyse van de methode leidde tot suggesties voor verbeteringen. De toepassing van deze verbeteringen in de quantumchemie wordt beschreven in hoofdstuk 6.

In hoofdstuk 7 worden de resultaten uit de andere hoofdstukken vergeleken met de lijst van eisen uit hoofdstuk 1. De conclusie is dat hoewel aan sommige eisen voldaan kan worden geen van de gebruikte methoden aan alle eisen voldoet. Omdat de mogelijkheden voor alternatieve methoden, gebruikmakende van een determinant basis, uitgeput beginnen te raken wordt gesuggereerd dat toekomstige ontwikkelingen in de richting van bijvoorbeeld expliciet gecorreleerde golffuncties zouden moeten gaan.

Curriculum vitae

Huub van Dam werd geboren op 20 juli 1968 te Grubbenvorst. In 1985 behaalde hij het HAVO-diploma aan de RSG "Den Hulster" te Venlo. Datzelfde jaar werd begonnen met de studie Chemische Technologie aan de Hogeschool Eindhoven. In 1989 legde hij het examen af, met als specialisatie Materiaalkunde en een afstudeerproject getiteld "The synthesis of polymer bonded acid-base indicators". Aansluitend vervulde hij de militaire dienstplicht bij de geneeskundige troepen. In 1990 werd begonnen met de studie Scheikunde aan de Katholieke Universiteit Nijmegen. In 1993 werd het doctoraal examen afgelegd, met als afstudeer variant informatische chemie en hoofdvak theoretische chemie getiteld "Het ontwikkelen van een *toolbox* voor quantum chemische moleculaire dynamica problemen" onder begeleiding van dr. G. C. Groenenboom.

Van september 1993 tot september 1997 was hij werkzaam als onderzoeker in opleiding bij de vakgroep Theoretische Chemie van de Universiteit Utrecht, onder begeleiding van dr. J. H. van Lenthe.