
Improving the Accuracy of Ab Initio Methods with Summation Approximants and Singularity Analysis

DAVID Z. GOODSON

Department of Chemistry and Biochemistry, University of Massachusetts at Dartmouth, North Dartmouth, MA 02747

Received 21 March 2002; accepted 16 September 2002

DOI 10.1002/qua.10489

ABSTRACT: The accuracy of Møller–Plesset (MP) perturbation theory and coupled-cluster (CC) theory can be significantly improved, at essentially no increase in computational cost, by using summation approximants that model the way in which these theories converge to the full configuration interaction limit. Approximants for MP4 and CCSD(T) are presented, their size scaling is analyzed, and the functional analysis of the MP energy, on which the MP4 approximant is based, is discussed. The MP approximants are shown to have a form that is appropriate for describing resonance energies. © 2003 Wiley Periodicals, Inc. *Int J Quantum Chem* 92: 35–46, 2003

Key words: MP perturbation theory; coupled-cluster theory; ab initio methods; scaling; full configuration interaction

Introduction

Here, an overview is provided of simple mathematical techniques that can significantly improve the accuracy of some standard ab initio methods with essentially no increase in computational cost and without the need to modify existing software packages. The main focus will be on Møller–

Plesset (MP) perturbation theory. The strategy will be to determine the mathematical structure of the MP energy function and then sum the perturbation theory with a summation approximant that models this structure [1, 2]. A recently proposed summation approximant for coupled-cluster (CC) theory [3] will also be discussed.

A perturbation series should not be taken as an end in itself but rather as the starting point for a functional analysis. The MP series is the asymptotic series (a Taylor expansion) of a function $E(z)$, which is the analytic continuation of the energy eigenvalue in terms of a perturbation parameter z . The full configuration interaction (FCI) energy, that is, the variational solution to the Schrödinger equation

This article was presented at the 42nd Annual Sanibel Symposium, February 23–March 1, 2002

This article is dedicated to the memory of Per-Olov Löwdin in honor of his work to elucidate the mathematical structure of quantum chemistry and, in particular, his seminal contributions to perturbation theory of the Schrödinger equation.

within a finite basis set expansion of the wave function, in principle corresponds to $E(1)$, while $E(0)$ is the sum of Hartree–Fock orbital energies. The MP series coefficients are proportional to the derivatives of $E(z)$ at $z = 0$, according to Taylor’s theorem. These derivatives contain information about the functional form of $E(z)$ that can be used to construct summation approximants that model the function over a larger range of z than does the conventional power series expansion. The result is an extrapolation from $z = 0$ to $z = 1$ that is more accurate than the usual method for evaluating the MP series.

Additional information about the mathematical structure of $E(z)$ is available from direct functional analyses of the MP Hamiltonian [4–6]. Using this information in conjunction with the perturbation series proves to be a particularly effective approach. The functional form of the summation approximant can be chosen in advance to model the expected structure of $E(z)$. This strategy greatly improves the accuracy of MP4. It also makes it possible, at least in principle, to apply existing MP and FCI calculation methods to the difficult problem of quasibound resonance states.

CC theory can also be thought of as a series solution for the FCI energy, with the same starting point as the MP series. The mathematical analysis of this series is much less straightforward. However, a comparison with benchmark FCI results in the literature shows that the sequence SCF, CCSD, CCSD(T) can be extrapolated toward the FCI limit with a summation approximant in the form of a continued fraction [3]. A summary of the results, which are remarkably good in many cases, will be presented.

Functional Analysis of the Møller–Plesset Energy

The MP partitioning of the Hamiltonian [7] is

$$H(z) = H_0 + zH_1, \quad H_1 = H_{\text{phys}} - H_0, \quad (1)$$

where H_0 is the sum of one-electron Fock operators and H_{phys} is the true Hamiltonian. This partitioning introduces into the analysis a perturbation parameter z as a continuous variable. The energy eigenvalue is obtained as an asymptotic power series in z of a function $E(z)$. This power series, truncated at some finite order, provides an approximate extrap-

olation from the sum of orbital energies, at $z = 0$, to the physical energy at $z = 1$.

The accuracy of this extrapolation depends on the singularity structure of the function $E(z)$ in the complex z plane [8]. For example, a famous theorem of Cauchy [9] tells us that the power series will diverge at $z = 1$ if $E(z)$ has a singularity within the disk $|z| \leq 1$. Even when the series is convergent, singularities just outside this disk can cause the convergence rate to be slow.

It is known that $E(z)$ contains a large number of branch-point singularities. The following result was proved by Katz [4]:

THEOREM 1. *Let $E^{(0)}(z), E^{(1)}(z), E^{(2)}(z), \dots$ be the spectrum of energy functions corresponding to the power series in z obtained from MP perturbation theory applied to the Hamiltonian of Eq. (1) for some given values of all conserved quantities (angular momentum, for example). Then, for any i and j there exist branch points z_{ij} in the complex z plane at which $E^{(i)}(z)$ is equal to $E^{(j)}(z)$ such that in the neighborhood of z_{ij}*

$$E^{(i)}(z) = b_{ij} \pm c_{ij}(1 - z/z_{ij})^{1/2}, \quad (2)$$

where b_{ij} and c_{ij} are constants.

This remarkable theorem implies that the entire energy spectrum for a given symmetry is described by a single function $E(z)$ with each eigenstate corresponding to a different Riemann sheet of the function. The sheets are connected by square-root branch points. The proof is based on an application of Liouville’s theorem to the function $[E^{(0)}(z) - E^{(i)}(z)]^{-1}$, in a manner closely analogous to Liouville’s proof of the fundamental theorem of algebra.

The implication for us of Theorem 1 is that $E(z)$ will contain complex-conjugate pairs of square-root branch points scattered throughout the complex z plane. However, these are not the only kinds of singularities we need to consider. Consider the following theorem:

THEOREM 2. *There exists a branch point on the negative z axis at which the energy function $E^{(0)}(z)$, which corresponds to the ground state at $z = 0$, crosses the energy function of an ionizing state.*

Baker [5] demonstrated the existence of an analogous phenomenon for many-body perturbation theory of nuclear matter and suggested that it should be a general characteristic of many-fermion sys-

tems. The specific case of an atomic system was treated by Stillinger [6] and nicely illustrated with an analysis of an approximate model for the He atom. The proof is straightforward. The operator H_1 in Eq. (1) contains the interelectron potentials $1/r_{ij}$. If z is negative these potentials are attractive. When z is sufficiently negative the system will dissociate, with the electrons leaving the nuclei to form a bound electron cluster. Baker [5] showed that this dissociation is analogous to a phase transition and noted that the function $E(z)$ should therefore be expected to have an algebraic singularity at the crossing point, as is in general the case for critical phenomena.

The behavior of the perturbation series of $E^{(0)}(z)$, at least in the limit of large order, is determined by the dominant singularity, that is, the singularity closest to the origin of the z plane. A precise statement of this principle is as follows:

THEOREM 3. *If the perturbation series is*

$$E(z) \sim \sum_{n=0}^{\infty} E_n z^n \quad (3)$$

and if the dominant singularity in the neighborhood of the singular point z_1 has the form

$$\sigma_1(z) = c_1(1 - z/z_1)^\nu \quad (4)$$

with Taylor series

$$\sigma_1(z) \sim \sum_{n=0}^{\infty} \sigma_{1,n} z^n, \quad (5)$$

then $E_n = \sigma_{1,n}$ in the limit of large n .

This was proved by Darboux [10].

If z_1 is an algebraic branch point on the negative real axis then the $\sigma_{1,n}$ will alternate in sign. If the dominant singularity structure corresponds to a complex-conjugate pair of square-root branch points,

$$\sigma_1(z) = c_1[(1 - z/z_1)^{1/2} + (1 - z/z_1^*)^{1/2}], \quad (6)$$

then if $\Re z_1 > 0$ there will be ranges of n with $\sigma_{1,n}$ of one sign alternating periodically with ranges with the opposite sign and if $\Re z_1 < 0$ there will be periods with alternating sign separated by two consecutive $\sigma_{1,n}$ with the same sign. The length of the periods is [11]

$$\nu = \pi / \arctan(|\Im z_1 / \Re z_1|). \quad (7)$$

Theorem 3 tells us the behavior for large n , but at smaller n nonsingular contributions to the perturbation series and the effects of other more distant singularities could be more important than the dominant behavior. In practice MP series for atoms and molecules often do show behavior consistent either with that of a dominant Katz branch-point pair close to the positive real axis or with that of a singularity on the negative real axis. The former case has been designated a series of "class A" while the latter has been designated a series of "class B" [12, 13]. Cremer and He [13] developed useful qualitative criteria for predicting in advance whether a given system will have a series of class A or of class B: Class A includes systems with well-separated Hartree-Fock orbitals and radicals while class B systems have orbitals that are crowded or that cluster in regions of space. Alkanes, for example, are typically in class A. Examples of class B are closed-shell systems with double or triple bonds or with highly electronegative atoms. It is possible to have a mixture of class A and class B behavior, due to singularity structure in both half planes at about the same distance from the origin, and the resulting sign patterns in the E_n can be intricate. An example is the C_2 molecule [14], in which the series behavior is apparently dominated by two different Katz branch points, at $-0.97 \pm 0.34i$ and at $1.18 \pm 0.36i$ [1].

Quadratic Approximants

Consider the three polynomials

$$P_L(z) = \sum_{i=0}^L p_i z^i, \quad (8)$$

$$Q_M(z) = 1 + \sum_{i=1}^M q_i z^i, \quad (9)$$

$$R_N(z) = \sum_{i=0}^N r_i z^i, \quad (10)$$

that satisfy the asymptotic relation

$$Q_M E^2 - P_L E + R_N \sim \mathcal{O}(z^{L+M+N+1}), \quad (11)$$

where E represents the power series, Eq. (3). The notation “ $\sim \mathcal{O}(z^j)$ ” means that the power series expansion of the expression on the left side has no terms with degree less than j . Solving for E , we obtain a multiple valued summation approximant

$$S_{[L/M,N]}(z) = \frac{1}{2} \left[\frac{P_L}{Q_M} \pm \frac{1}{Q_M} \sqrt{P_L^2 - 4Q_MR_N} \right]. \quad (12)$$

The two branches (\pm) of $S_{[L/M,N]}$ are connected by square-root branch points at the nondouble roots of the discriminant polynomial,

$$D_{[L/M,N]}(z) = P_L^2 - 4Q_MR_N. \quad (13)$$

The coefficients p_i, q_i, r_i are determined from the set of $L + M + N + 2$ simultaneous linear equations obtained by collecting terms in Eq. (11) according to power of z .

Quadratic approximants are a special case of the algebraic approximants, a general construction in which the left side of Eq. (11) is replaced by a polynomial in E of arbitrary degree m , which was proposed by Padé in his 1892 Ph.D. thesis [15]. The rational approximants, $m = 1$, were investigated in detail by Padé and are well known today. The higher-degree approximants appear to have been left largely unexplored until over 80 years later, when Shafer [16] rediscovered the quadratic approximants and used them to sum Taylor series for some simple multiple valued functions. Since then, the mathematical properties of algebraic approximants have been explored in more detail [17–20] and applications have been found for them in various areas of atomic and molecular physics [21–27].

Quadratic approximants would seem to be well suited for describing the singularity structure of the MP energy function. Conventional MP theory uses the partial summation approximant,

$$S_n(z) = \sum_{i=0}^n E_i z^i. \quad (14)$$

The function $S_n(z)$ is a polynomial, which is non-singular at all finite z . This is the cause of the MP series convergence problems that have attracted so much attention in recent years [14, 28–33]. $S_n(z)$ is unable to model the singularity structure of $E(z)$. Quadratic approximants of MP series converge to the FCI energy, even in cases where the partial summation approximants diverge badly [1].

In principle, Eq. (12) describes two branches of $E(z)$. One branch corresponds to the ground state while the other corresponds to an excited state. It has been demonstrated [19], in a study of anharmonic oscillators, that algebraic approximants can simultaneously converge on different Riemann sheets if the perturbation series is known to sufficiently high order, although the rate of convergence is much faster on the principal sheet. At fourth order, the MP series only provides enough information for the quadratic approximant to accurately describe the Riemann sheet of the ground state. It is usually obvious in practice (by comparison with the value of the rational approximant, for example) which branch of Eq. (12) gives the ground-state energy.

q λ Summation Method

Large-order MP series can be generated by means of FCI methodologies [34, 35], but direct calculation of the series coefficients is at present feasible only through sixth order [36] and calculations beyond fourth order are rare on account of the high computational cost. The fourth-order series (MP4) shows enough large-order behavior for quadratic approximants to significantly improve the accuracy on average, but on occasion the quadratic approximant fails badly and the partial summation approximant S_4 gives the better result [1]. The q λ method [2] is a procedure that improves the dependability of the quadratic approximants.

The basic idea is to shift the position of the dominant singularity away from the origin of the complex plane by repartitioning the Hamiltonian. The λ transformation, originally proposed by Feenberg [37] in the context of nuclear physics, is in particular convenient for this purpose. Let

$$H^{(\lambda)}(z) = \frac{1}{1-\lambda} H_0 + z \left(H_1 - \frac{\lambda}{1-\lambda} H_0 \right), \quad (15)$$

where λ is an arbitrary parameter and H_1 is defined as in MP theory according to Eq. (1). At the physical point $z = 1$ this Hamiltonian is equivalent to H_{phys} , as is the MP Hamiltonian, $H^{(0)}(z)$, but elsewhere in the z plane each value of λ yields a different function $E(z)$. Schmidt et al. [12] found that

$$E_i^{(\lambda)} = \sum_{k=1}^{i-1} \binom{i-2}{k-1} \lambda^{i-k-1} (1-\lambda)^k E_{k+1} \quad (16)$$

for $i \geq 2$. Therefore, an advantage of the Feenberg repartitioning is that the perturbation series coefficients for arbitrary λ can be expressed as a linear combination of the $\lambda = 0$ series coefficients, E_j , given by standard software packages. The Hartree–Fock energy is left unchanged by the λ transformation:

$$E_0^{(\lambda)} + E_1^{(\lambda)} = E_0 + E_1. \quad (17)$$

To choose the optimal value of λ we need to know the position of the dominant singularity, z_d . The $q\lambda$ method takes the roots of the discriminant polynomial, $D_{[L/M,N]}$, as an estimate of the singularity structure. z_d is assumed to be approximately equal to the value of the root nearest the origin.

In practice [1, 2] the quadratic approximant is applied to a series with the zeroth-order term equal to the Hartree–Fock energy,

$$\tilde{E}^{(\lambda)}(z) = \sum_{i=0}^{n-1} \tilde{E}_i^{(\lambda)} z^i, \quad (18)$$

where

$$\tilde{E}_0^{(\lambda)} = E_0^{(\lambda)} + E_1^{(\lambda)}, \quad \tilde{E}_{i>0}^{(\lambda)} = E_{i+1}^{(\lambda)}. \quad (19)$$

This is mainly for the sake of convenience because E_0 and E_1 are not usually reported individually in the literature. It seems to have little effect on the accuracy whether this series or the original series is used. The two functions are related by

$$E^{(\lambda)}(z) = (1 - z)E_0^{(\lambda)} + z\tilde{E}^{(\lambda)}(z). \quad (20)$$

Because the first term on the right side of Eq. (20) is nonsingular, $\tilde{E}^{(\lambda)}(z)$ and $E^{(\lambda)}(z)$ will have the same branch points.

Figure 1 shows the roots of the discriminant as a function of λ for the NH_2 molecule, a clear example of a class A system according to the Cremer–He criteria. At large positive λ the dominant singularity is in the positive half plane while at large negative λ it is in the negative half plane. The MP4 approximant can model the singularity structure in only one of the half planes at a time. In an intermediate range of λ it makes a sharp transition between the class A and class B singularities. In fact, the true $\tilde{E}^{(\lambda)}(z)$ has singularities in both half planes at all λ . Only at higher order is there enough information

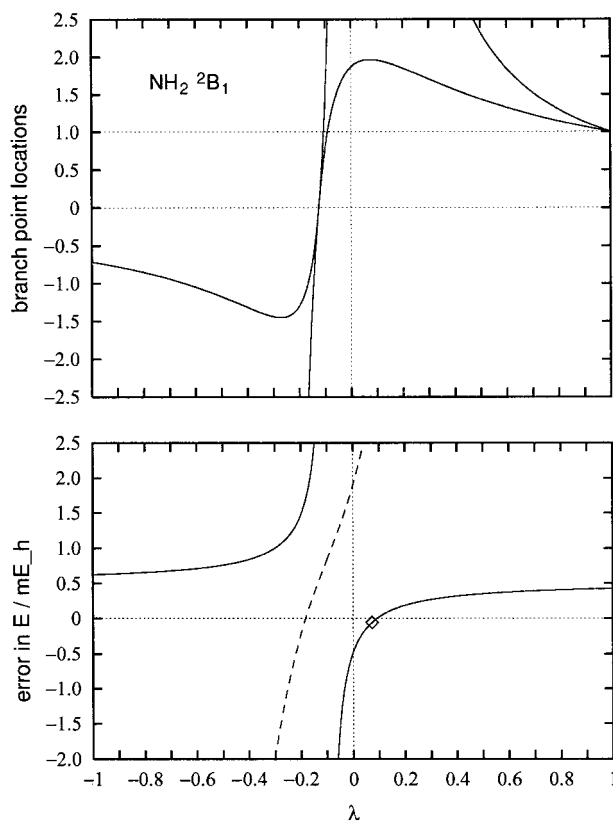


FIGURE 1. Branch points and energy error of quadratic approximant for the 2B_1 state of NH_2 as a function of repartitioning parameter λ . The perturbation series is from Ref. [36], with the DZP basis. The upper panel shows the positions of the branch points of the quadratic approximant (which lie on the real axis). The lower panel shows $E_{\text{approx}} - E_{\text{FCI}}$ vs. λ , where E_{approx} is the quadratic approximant $S_{[1/0,1]}$ (solid curve) or the partial summation approximant S_4 (broken curve). The diamond marks the point corresponding to the maximum in the dominant singularity of the quadratic approximant in the positive half plane.

about the singularities for the approximant to simultaneously model class A and class B structure.

The lower panel of Figure 1 shows the MP4 summation error (i.e., the difference between the summation approximant at $z = 1$ and the FCI energy) as a function of λ . The diamond marks the point corresponding to the maximum z_d . The small increase from the $\lambda = 0$ value of z_d has a significant effect on the summation accuracy. The accuracy is poor in the transition region of λ because the approximant is trying to model the approximately equidistant singularities in the positive and negative half planes by placing a single singularity somewhere in between them. This means that the

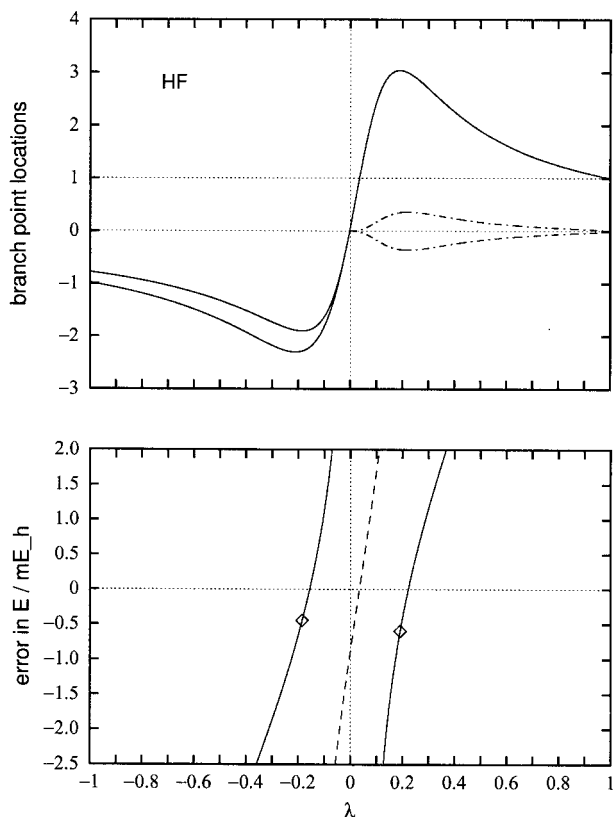


FIGURE 2. Branch point locations and error of the $[1/0, 2]$ constrained quadratic approximant for the hydrogen fluoride molecule as a function of the repartitioning parameter λ . The dash-dot curves in the upper panel show the imaginary parts of the branch points while the solid curves show the real parts. The broken curve shows the energy error from partial summation of MP4. The diamonds mark the points corresponding to the maxima in the dominant singularity of the quadratic approximant. The perturbation series is from Ref. [14] with the aug-cc-pVDZ basis.

approximant has a nonphysical singularity near the origin but has no singularities in either of the regions where the true $E^{(\lambda)}(z)$ is singular. Those cases in which direct quadratic summation performs poorly [1] usually have $\lambda = 0$ lying well within the transition region. The singularity shifting of the $q\lambda$ procedure ensures that the transition region will be avoided. The behavior for NH_2 is typical for class A systems. For a benchmark set of 10 class A molecules at equilibrium geometries it was found [3] that the median MP4 error with partial summation was reduced from $1.890 mE_h$ to $0.109 mE_h$ by the $q\lambda$ approximant.

For class B the situation is somewhat more complicated. For such systems, it is appropriate to in-

roduce the constraint $r_0 = 0$ [1] and use the approximate with index $[1/0, 2]$ [2]. Figure 2 shows results for the hydrogen fluoride molecule with the aug-cc-pVDZ basis [38]. The maximum in $|z_d|$ in the positive half plane and the maximum in the negative half plane both give good results for the energy, although the improvement over partial summation is not as great as for class A systems. The “ $q\lambda-$ ” result, corresponding to the maximum in the negative half plane, gives an error of $-0.447 mE_h$, compared to -0.882 for the partial sum and 0.536 for CCSD(T). However, for class B there is a strong correlation between the accuracy of the quadratic approximant and how far into the negative half plane z_d can be shifted. For hydrogen fluoride, z_d can be shifted as far as -1.90 , but for F^- , with the same basis, z_d is only shifted to -1.36 . The conventional MP4 error for F^- is $-5.502 mE_h$ and MP4- $q\lambda-$ reduces it to $-2.165 mE_h$, a significant improvement but still much worse than the CCSD(T) error of $0.735 mE_h$. The problem here is similar to that with the $\lambda = 0$ approximants for class A systems—singularity structure in both half planes is being modeled by a single branch point in between.

In general, the MP4- $q\lambda$ result tends to be worse than the CCSD(T) result for class B if z_d of the approximant cannot be shifted beyond -1.5 [2, 3]. Figure 3, which shows $-\log_{10}(|E_{\text{MP4-}q\lambda} - E_{\text{FCI}}|/E_{\text{corr}})$ as a function of the maximum $|z_d|$ for 40 different MP4 series,* demonstrates that singularity analysis of the quadratic approximant offers a reasonably reliable diagnostic of the summation accuracy.

Resonance Eigenvalues

A quasibound unstable resonance can be treated as an eigenstate of the time-independent Schrödinger equation with a complex energy eigenvalue

*The systems are, in order of increasing $|z_d|$, H_2O cc-pVDZ with $R = 3R_e$, H_2O cc-pVDZ with $R = 2.5R_e$, CN^+ cc-pVDZ, C_2 cc-pVDZ, C_2 cc-pVDZ(+), H_2O DZP with $R = 2R_e$, H_2O cc-pVDZ with $R = 2R_e$, HF DZP with $R = 2R_e$, CN STO3G, H_2O DZP with $R = 1.5R_e$, H_2O cc-pVDZ with $R = 1.5R_e$, HF cc-pVDZ with $R = 2.0R_e$, N_2 cc-pVDZ, F^- aug-cc-pVDZ, F^- [4s3p1d], F^- [5s4p2d], F^- [4s3p2d], HF DZP with $R = 1.5R_e$, HF cc-pVDZ with $R = 1.5R_e$, HF aug-cc-pVDZ, HF cc-pVTZ-(f/d), H_2O cc-pVTZ(+), F^- cc-pVTZ-(f), H_2O DZP, HF DZP, H_2O cc-pVDZ (core not frozen), Ne aug-cc-pVDZ, H_2O cc-pVDZ, Ne [6s4p1d], Ne [4s2p1d], Ne cc-pVTZ-(f), F [5s4p2d], F [4s3p2d], F [4s3p1d], Ne [5s3p2d], HF cc-pVDZ, HCl aug-cc-pVDZ, Cl^- aug-cc-pVDZ, Ne cc-pVDZ, F^- cc-pVDZ.

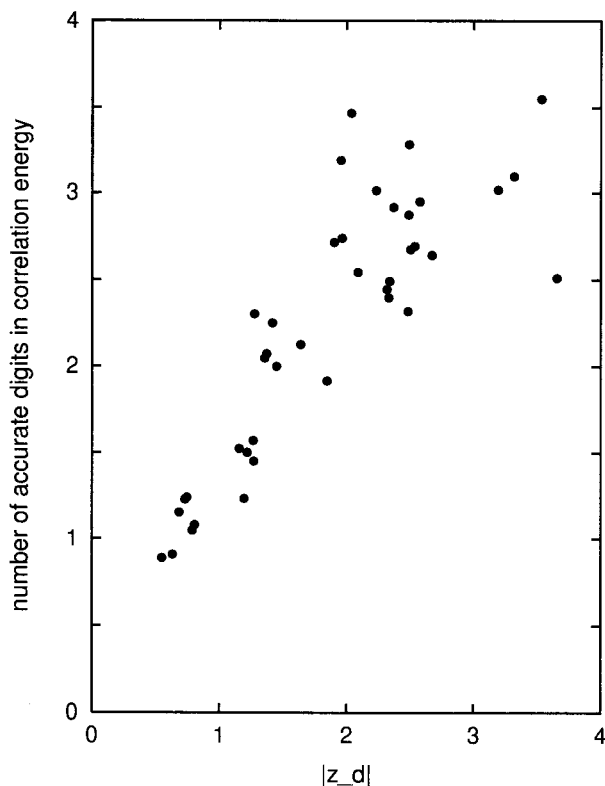


FIGURE 3. Number of accurate digits in correlation energy, defined as $-\log_{10}|\Delta E/E_{\text{corr}}|$, $\Delta E = E_{\text{MP4-q}\lambda} - E_{\text{FCI}}$, as a function of the distance from the origin of the dominant singularity of the MP4-q λ approximant in the negative half plane, for a set of 40 class B systems (see footnote on previous page).

[39–41]. Such states can be observed spectroscopically as broad peaks embedded in an unbound continuum. The real part of the eigenvalue gives the position of the peak while the imaginary part is proportional to the linewidth, Γ , according to the expression (in atomic units) $\Gamma = \Im E/2$.

Perturbation theories of these eigenvalues typically give asymptotic series of functions with branch points. Consider the Eckart potential [42], the classic model problem of scattering theory, with the potential energy

$$V(x) = \frac{Ae^x}{1+e^x} + \frac{Be^x}{(1+e^x)^2}, \quad (21)$$

which is a barrier with maximum value of $V(x_{\text{max}}) = (A+B)^2/4B$ at the point

$$x_{\text{max}} = \ln\left(\frac{A+B}{A-B}\right). \quad (22)$$

The exact expression for the tunneling resonance energies is [43]

$$\epsilon_n^\pm = \frac{\{[\pm\sqrt{4B-1} - (2n+1)i]^2 + 4A\}^2}{16[\pm\sqrt{4B-1} - (2n+1)i]^2}. \quad (23)$$

Consider a perturbation theory with the zeroth-order approximation $B \rightarrow \infty$, corresponding to an infinitely high barrier [41]. The scaled energy $\epsilon_n^\pm/4B$ can be expanded in terms of a power series in the parameter

$$\gamma = 1/\sqrt{4B}. \quad (24)$$

The exact solution, Eq. (23), can be rearranged and expressed as a function of γ in the form

$$\frac{1}{4B} \epsilon_n^\pm(\gamma) = \frac{1}{2} \left[\frac{\alpha(\gamma)}{\kappa(\gamma)} \mp \frac{\sqrt{\beta(\gamma)}}{\kappa(\gamma)} \right], \quad (25)$$

where $\kappa(\gamma)$, $\alpha(\gamma)$, and $\beta(\gamma)$ are polynomials,

$$\kappa(\gamma) = [1 + 4n(n+1)\gamma^2]^2, \quad (26)$$

$$\alpha(\gamma) = \frac{1}{8} [1 + (8A - 4n^2 - 4n - 2)\gamma^2]\kappa(\gamma) + 2A^2\gamma^4[1 - (4n^2 + 4n + 2)\gamma^2], \quad (27)$$

$$\beta(\gamma) = 16\gamma^2(2n+1)^2(\kappa - A^2\gamma^4)^2(\gamma^2 - 1). \quad (28)$$

Equation (25) has the same general form as the expression for the quadratic approximant, Eq. (12), with branch points at $\gamma = \pm 1$. This suggests that quadratic approximants would be well suited for treating problems of this sort.

Some resonance problems lead to perturbation theories with a zero radius of convergence, with a branch point at the origin. An example is the H atom in an external electric field, which has no bound states for any nonzero value of the field strength. The perturbation theory series coefficients, however, are real numbers. The desired complex resonance energies are on a different branch of the energy function than that given by the perturbation series. One approach to this problem is to sum the series with a Padé–Borel approximant with integration over a complex path [43]. Another approach, which is somewhat simpler, is to sum the series with quadratic approximants [21]. Both methods give good results. The true singularity is more complicated than a square-root branch point.

However, the approximants are able to use groupings of square-root branch points to model the singularity, much as rational Padé approximants are able to model branch points with groupings of poles [44].

Quadratic and higher-degree approximants have also been used to calculate resonance energies for anharmonic oscillators [19] and for quasibound molecular rotation eigenstates [23]. It remains to be seen whether this approach could be applied to MP perturbation theory to obtain molecular electronic resonance energies. The other applications required high orders of perturbation theory to obtain reasonable results. One possibility would be to generate a large-order perturbation expansion from a bound-state FCI calculation [34, 35] and then use quadratic or Padé–Borel approximants to obtain the complex energy. For example, one might calculate the (pure real) large-order MP perturbation series for a metastable anion such as N_2^- . The more familiar approaches to this problem are the complex scaling (CS) method [45] and the complex absorbing potential (CAP) method [46]. Both of these methods use a modified non-Hermitian Hamiltonian with an explicit imaginary part. Also, they introduce an arbitrary parameter, and repeated calculations are needed to optimize its value. A large-order MP calculation, which uses the original Hamiltonian and involves no adjustable parameters, would be a more direct approach. Whether its computational cost will be competitive with the CS and CAP methods will depend on the rate of convergence of the summation approximants.

CC Continued-Fraction Approximant

In recent years CCSD(T) has largely supplanted conventional MP4 as the MP convergence problems have become better known and as the relatively greater reliability of CCSD(T) has been demonstrated in various applications [14, 30–33]. The success of summation approximants for perturbation theory suggests that we search for a similar approach for CC theory.

The two theories have much in common. Both use the SCF wave function as the initial approximation and then extrapolate to the FCI solution through a sequence of approximations. Let us express the CCSD(T) energy in the form

$$E_{\text{CC}} = \delta_1 + \delta_2 + \delta_3, \quad (29)$$

where $\delta_1 = E_{\text{SCF}}$ is the Hartree–Fock energy and

$$\delta_2 = E_{\text{CCSD}} - E_{\text{SCF}}, \quad \delta_3 = E_{\text{CCSD(T)}} - E_{\text{CCSD}}. \quad (30)$$

Can Eq. (29) be summed using methods of perturbation theory? The problem is that in this case we do not have an asymptotic series. The MP series is the unique asymptotic expansion of the function $E(z)$ and it is not clear what the relation is between this function and the E_{CC} sequence. Nevertheless, if the E_{CC} sequence has a characteristic convergence pattern then it might be possible to extrapolate it toward the FCI limit with an appropriate summation approximant.

Such an approximant was recently discovered [3]. A comparison with a large number of benchmark FCI energies demonstrated that the continued-fraction approximant

$$E_{\text{CCcf}} = \frac{\delta_1}{1 - \frac{\delta_2/\delta_1}{1 - \delta_3/\delta_2}} \quad (31)$$

can significantly improve the accuracy of the CCSD(T) method. This approximant is in particular successful for class A systems, reducing the median error of CCSD(T) for the 10 benchmark systems from 0.506 mE_h to 0.051 mE_h , smaller than the MP4-q λ error by a factor of 2. For class B systems the accuracy is comparable to that from MP4-q λ , typically slightly less accurate when $|z_d|$ is large but more accurate when it is small.

These accuracy comparisons are for total energies. For practical applications, accuracy of the total energy is less important than accuracy of the Born–Oppenheimer potential energy hypersurface, which requires that the error from the calculation method be reasonably independent of the molecular geometry. Recent comparisons with FCI energies for the BH, CH₃, HF, and H₂O molecules over a range of bond distances show that the higher accuracy of the total energy from CCSD(T)-cf does lead to a more accurate potential energy curve for stretching of the bond distance to at least $1.5R_e$, where R_e is the equilibrium bond distance [3, 47]. Table I compares the total energy error, $\Delta E_{\text{approx}} = E_{\text{approx}} - E_{\text{FCI}}$, from conventional CCSD(T) and CCSD(T)-cf at $R = R_e$ and $1.5R_e$. The improvement from the cf method is clearly significant. For H₂O Table I also presents results from the asymmetrical method, a-CCSD(T), of Crawford and Stanton [48, 49]. This method, based on Löwdin’s partitioning approach to pertur-

TABLE I
Error (mE_h) from CC calculations.

Molecule	Method	$\Delta E(R_e)$	$\Delta E(1.5R_e)$
BH	CCSD(T)	0.521	0.731
	CCSD(T)-cf	0.072	0.148
CH ₃	CCSD(T)	0.527	1.942
	CCSD(T)-cf	-0.138	0.652
HF	CCSD(T)	0.496	0.931
	CCSD(T)-cf	0.041	0.328
H ₂ O	CCSD(T)	0.658	1.631
	a-CCSD(T)	0.698	2.131
	CCSD(T)-cf	0.010	0.394
	a-CCSD(T)-cf	0.032	0.038

ΔE is the difference from the corresponding FCI result. The FCI energies are from Ref. [47] (BH, CH₃, and HF) and Ref. [30] (H₂O). The a-CCSD(T) energies are from Ref. [48]. For the open-shell system CH₃ spin-unrestricted CC calculations are used.

bation theory [50], is formulated as a third-order perturbation series and gives an expression for the energy similar to that from conventional CCSD(T). At R_e , CCSD(T)-cf and a-CCSD(T)-cf are almost the same and both give excellent accuracy. At $1.5R_e$, they differ somewhat and a-CCSD(T)-cf is significantly more accurate than CCSD(T)-cf, even though a-CCSD(T) is less accurate than conventional CCSD(T). This requires further study, but it suggests that the continued fraction is describing the behavior of the a-CCSD(T) perturbation series and the reason it can also sum the CCSD(T) sequence is simply that the numerical value of the CCSD(T) energy is reasonably close to that of the a-CCSD(T) energy.

Size Scaling

An important motivation for the adoption of MP and CC theories by quantum chemists was the fact that these methods scale properly as the size of the system increases, in contrast to truncated CI methods [54]. Size scaling is described in terms of the two related concepts size extensivity and size consistency. The former is analogous to the thermodynamic concept of a size-extensive property while the latter describes the behavior of the theory as a molecule is broken up into noninteracting fragments.

Suppose a system consists of m identical units. A size-extensive theory yields a result for the energy that scales linearly with m [52]. The MP series co-

efficients are individually size extensive [53]. Because the λ transformation, Eq. (16), is linear, the coefficients $E_i^{(\lambda)}$ are also individually size extensive. Let

$$E^{(\lambda,m)} = mE^{(\lambda)}. \quad (32)$$

It follows from Eq. (11) that

$$Q_M = m^2 Q_M^{(m)}, \quad P_L = m P_L^{(m)}, \quad R_N = R_N^{(m)}. \quad (33)$$

Substituting these into Eq. (12) gives

$$S_{[L/M,N]}^{(m)} = m S_{[L/M,N]}. \quad (34)$$

Hence, the $q\lambda$ method is size extensive.

The CCSD and CCSD(T) energies also are individually size extensive [54]. Substituting

$$\delta_i^{(m)} = m \delta_i \quad (35)$$

into Eq. (31) gives

$$E_{\text{CCSD(T)cf}}^{(m)} = m E_{\text{CCSD(T)cf}} \quad (36)$$

Hence, the CCSD(T)-cf method is size extensive.

A theory is size consistent if an energy calculation for a molecule AB in which fragments A and B are sufficiently far apart as to be essentially noninteracting gives a result equal to the sum of the energies given for the isolated fragments [54]. The MP series coefficients and the CC sequence terms are individually size consistent [51], and it follows that the λ transformation is also size consistent. However, the quadratic approximant and the continued fraction are nonlinear transformations, which will strictly preserve size consistency only for dissociation into *identical* fragments.

The consequences of this lack of exact size consistency for nonidentical fragments do not appear to be significant. In contrast to a truncated CI calculation, which can give qualitatively wrong results as a consequence of size inconsistency, the size extensivity of the $q\lambda$ and cf transformations ensures that the results for bond stretching will be qualitatively correct {except, of course, for the RHF symmetry breaking effects [55], which affect the underlying MP4 and CCSD(T) methods as well, and in practice usually prevent any of these methods from correctly describing homolytic molecular dissociation}.

TABLE II
Summation accuracy and singularity analysis of Taylor series of the model functions of Eqs. (37), (38), and (39).

Function	ΔE_{sum}	$\Delta E_{[1/0,1]}$	z_d
f_A	3.831	0.167	2.067
f_B	1.682	0.107	2.536
$f_A + f_B$	5.513	0.274	
f_{AB}	5.513	0.367	2.233

Results are shown for third-order partial summation, the [1/0, 1] quadratic approximant, and the position z_d of the dominant singularity in the quadratic approximant. ΔE is the difference between the summation and the exact value.

This is illustrated by the following simple model problem. Consider the functions

$$f_A(z) = ae^{-\alpha z} + (1 - z/z_A)^{1/2}, \quad (37)$$

$$f_B(z) = be^{-\beta z} + (1 - z/z_B)^{1/2}, \quad (38)$$

which are sums of a function with no singularities at finite z and a function with a square-root branch point, and let

$$f_{AB}(z) = f_A(z) + f_B(z). \quad (39)$$

The parameter values $a_1 = b_1 = -12$, $\alpha = \beta = 0.005$, $z_A = 2.0$, $z_B = 2.5$, lead to Taylor series with behavior comparable to that of typical class A MP \tilde{E} energy series. f_A and f_B differ only in the locations z_A and z_B of their branch points. Table II compares the accuracy of quadratic approximants with that of partial summation of the Taylor series. The quadratic approximant for f_{AB} places its dominant branch point between the true branch points and therefore gives a result that is slightly less accurate than that from separate quadratic summation of f_A and f_B , but is still good. The conventional partial sum is strictly size consistent but not accurate because it makes no attempt to model the branch points.

Thus, the size inconsistency of MP4-q λ is fundamentally different from that of truncated CI. For CI, the “supermolecule” analysis, treating the separate fragments as a single system, includes fewer excited configurations than does the separate treatment of the fragments, which leads to a wave function that has less flexibility for describing the electron correlation. MP4-q λ uses the same wave function as

MP4, with the same level of correlation for the supermolecule and for the separate fragments. The functional analysis for the CC energy is less straightforward. However, CCSD(T) and CCSD(T)-cf are based on the same wave function. Therefore, CCSD(T)-cf can also be expected to be qualitatively size consistent.

Discussion

It is clear that conventional CCSD(T) is more dependable than conventional MP4, but with both methods improved by summation approximants the comparison becomes more complicated. With class A systems CCSD(T)-cf seems to be superior to MP4-q λ , but for class B, for systems with typical class B singularity structure, MP4-q λ seems to be slightly better. “Typical” class B structure is a branch point in the negative half plane that can be shifted beyond approximately -1.5 by the λ transformation. Class B systems for which this is not the case are perhaps best thought of as belonging to a class “AB,” for which an accurate series summation would require the simultaneous modeling of branch points in both half planes. In such cases MP4-q λ fails because the approximant places a single branch point somewhere between the true branch points.

However, the main advantage of the perturbation theory is probably not the higher accuracy for certain cases because CCSD(T)-cf seems to also give consistently reliable results for the typical class B systems, even if the average accuracy is slightly lower [3]. Rather, the advantage is the transparent mathematical structure of the perturbation theory, which makes it possible to use the singularity structure of the quadratic approximant as a diagnostic for predicting the accuracy of the energy result. Without a variational principle to rely on, such a diagnostic is in particular useful to have. The singularity structure can also be used to identify cases in which methods based on single-reference Hartree-Fock determinants are inappropriate. Results in Ref. [3] suggest that if the class B singularity cannot be shifted beyond approximately -1.2 then poor results will be obtained even with single-reference CCSD(T)-cf.

The functional analysis of the MP energy function points the way toward possible further improvements by illuminating the causes for failure of the summation in difficult cases. Consider the problem of homolytic bond cleavage for class A mole-

cules. As the bond distance approaches dissociation the ground state of the physical ($z = 1$) Hamiltonian becomes nearly degenerate with the first excited state. This means that a complex-conjugate pair of branch points in the positive half plane will gradually approach the point $z = 1$ as the bond is stretched. This decreases the accuracy of the summation in two ways. First, it makes the placement of the approximant's singularities less accurate, through the class AB phenomenon. Second, the closer the true singularity structure is to the physical point the more important it becomes to model it precisely because that is the point where the approximant will be evaluated. A small error in the singularity structure in the neighborhood of $z = 1$ will have a much larger effect than it will elsewhere in the complex plane.

A multireference perturbation theory is the obvious choice for dealing with this situation. If the determinants corresponding to the two states connected by the branch points are both included in the reference space, then these branch points should disappear. As an illustration, consider the complete active space (CAS) perturbation series for the HF molecule that Olsen and Fülcher [56] extracted from an FCI calculation. With bond distance at twice the equilibrium value, the single-reference MP4-q λ has maximum z_d in the positive half plane close to 1 while for CAS perturbation theory at fourth order the maximum z_d is almost 5.

The arbitrariness involved in choosing the reference space for a multireference calculation is sometimes cited as a disadvantage. This arbitrariness could be removed by choosing the reference space according to singularity analysis of the perturbation series. For example, the reference space could be chosen so as to shift the dominant branch point as far as possible from the origin of the complex plane. A first step in this direction was taken by Warken [59] in a study of several electronic eigenvalues of H₂O. That study used fourth-order perturbation theory with the Λ transformation [58], a multireference generalization of the λ transformation of Eq. (16). Warken used the closest approach of avoided crossings along the real axis to estimate the real part of the branch point locations. Quadratic approximants offer a more direct and computationally less demanding way to locate the branch points and give values for the imaginary part as well as for the real part.

ACKNOWLEDGMENT

This work was supported in part by a grant from the Welch Foundation.

References

1. Goodson, D. Z. *J Chem Phys* 2000, 112, 4901.
2. Goodson, D. Z. *J Chem Phys* 2000, 113, 6461.
3. Goodson, D. Z. *J Chem Phys* 2002, 116, 6948.
4. Katz, A. *Nucl Phys* 1962, 29, 353.
5. Baker, G. A., Jr. *Rev Mod Phys* 1971, 43, 479.
6. Stillinger, F. H. *J Chem Phys* 2000, 112, 9711.
7. Møller, C.; Plesset, M. S. *Phys Rev* 1934, 46, 618.
8. See, for example, Churchill, R. V. *Complex Variables and Applications*; McGraw-Hill: New York, 1960.
9. Cauchy, A.-L. *Comp Rend* 1837, 4, 216; Kline, M. *Mathematical Thought from Ancient to Modern Times*; Oxford University Press: New York, 1972; p 639.
10. Darboux, M. G. *J Math Pure Appl* 3 1878, 4, 377; Ninham, B. W. *J Math Phys* 1963, 4, 679.
11. Goodson, D. Z.; López-Cabrera, M.; Morgan, J. D., III; Herschbach, D. R. *J Chem Phys* 1992, 97, 8481.
12. Schmidt, C.; Warken, M.; Handy, N. C. *Chem Phys Lett* 1993, 211, 272.
13. Cremer, D.; He, Z. *J Phys Chem* 1996, 100, 6173.
14. Leininger, M. L.; Allen, W. D.; Schaefer, H. F., III; Sherrill, C. D. *J Chem Phys* 2000, 112, 9213.
15. Padé, H. *Annal Ecole Normale Superieure* 1892, 9(Suppl), 1-93.
16. Shafer, R. E. *SIAM J Math Anal* 1975, 11, 447.
17. Baker, G. A., Jr.; Graves-Morris, P. *Padé Approximants*; Cambridge University Press: Cambridge, UK, 1996; 524-569.
18. Sergeev, A. V. *Z Vychisl Mat Mat Fiz* 1986, 26, 348 (*USSR Comput Maths Math Phys* 1986, 26, 17).
19. Sergeev, A. V.; Goodson, D. Z. *J Phys A Math Gen* 1998, 31, 4301.
20. Fernández, F. M.; Diaz, C. G. *Eur Phys J D* 2001, 15, 41.
21. Vaĭnberg, V. M.; Mur, V. D.; Popov, V. S.; Sergeev, A. V. *Pis'ma Z Eksp Teor Fiz* 1986, 44, 9 (*JETP Lett* 1986, 44, 9).
22. López-Cabrera, M.; Goodson, D. Z.; Herschbach, D. R.; Morgan, J. D., III. *Phys Rev Lett* 1992, 68, 1992.
23. Germann, T. C.; Kais, S. *J Chem Phys* 1993, 99, 7739.
24. Sergeev, A. V. *J Phys A Math Gen* 1995, 28, 4157.
25. Goodson, D. Z.; Sergeev, A. V. *J Chem Phys* 1999, 110, 8205.
26. Fernández, F. M.; Tipping, R. H. *J Mol Struct Theochem* 1999, 488, 157.
27. Diaz, C. G.; Fernández, F. M. *J Mol Struct Theochem* 2001, 541, 39.
28. Olsen, J.; Christiansen, O.; Koch, H.; Jørgensen, P. *J Chem Phys* 1996, 105, 5082.
29. Christiansen, O.; Olsen, J.; Jørgensen, P.; Koch, H.; Malmqvist, P.-Å. *Chem Phys Lett* 1996, 261, 369.
30. Olsen, J.; Jørgensen, P.; Koch, H.; Balkova, A.; Bartlett, R. J. *J Chem Phys* 1996, 104, 8007.

GOODSON

31. Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J Chem Phys* 1997, 106, 6430.
32. Dunning, T. H., Jr.; Peterson, K. A. *J Chem Phys* 1998, 108, 4761.
33. Larsen, H.; Halkier, A.; Olsen, J.; Jørgensen, P. *J Chem Phys* 2000, 112, 1107.
34. Handy, N. C.; Knowles, P. J.; Somasundram, K. *Theor Chim Acta* 1985, 68, 87.
35. Laidig, W. D.; Fitzgerald, G.; Bartlett, R. J. *Chem Phys Lett* 1985, 113, 151.
36. He, Z.; Cremer, D. *Int J Quantum Chem* 1996, 59, 71.
37. Feenberg, E. *Phys Rev* 1956, 103, 1116.
38. Kendall, R. A.; Dunning, T. H., Jr. *J Chem Phys* 1992, 96, 6796.
39. Siegert, A. J. F. *Phys Rev* 1939, 56, 750.
40. Connor, J. N. L.; Smith, A. D. *Mol Phys* 1981, 43, 397.
41. Fernández, F. M. *J Chem Phys* 1996, 105, 10444.
42. Eckart, C. *Phys Rev* 1930, 35, 1303.
43. Franceschini, V.; Grecchi, V.; Silverstone, H. J. *Phys Rev A* 1985, 32, 1338.
44. Baker, G. A., Jr. *The Essentials of Padé Approximants*; Academic Press: New York, 1975.
45. McCurdy, C. W. In: Tomkin, A., ed. *Autoionization: Recent Developments and Applications*; Plenum: New York, 1985; p 153–170.
46. Sommerfeld, T.; Riss, U. V.; Meyer, H.-D.; Cederbaum, L. S.; Engels, B.; Suter, H. U. *J Phys B* 1998, 31, 4107.
47. Goodson, D. Z.; Zheng, M. *Chem Phys Lett* 2002, 365, 396.
48. Crawford, T. D.; Stanton, J. F. *Int J Quantum Chem* 1998, 70, 601.
49. Stanton, J. *Chem Phys Lett* 1997, 281, 130.
50. Löwdin, P.-O. *J Math Phys* 1962, 3, 969.
51. Bartlett, R. J. *Annu Rev Phys Chem* 1981, 32, 359, and references therein.
52. Bartlett, R. J.; Purvis, G. D. *Int J Quantum Chem* 1978, 14, 561.
53. Goldstone, J. *Proc Roy Soc Lond A* 1957, 239, 267.
54. Pople, J. A.; Binkley, J. S.; Seeger, R. *Int J Quantum Chem Symp* 1976, 10, 1.
55. Löwdin, P.-O. *Rev Mod Phys* 1963, 35, 496.
56. Olsen, J.; Fülscher, M. P. *Chem Phys Lett* 2000, 326, 225.
57. Warken, M. *J Chem Phys* 1995, 103, 5554.
58. Dietz, K.; Schmidt, C.; Warken, M.; Hess, B. A. *J Phys B Atom Mol Opt Phys* 1993, 26, 1885; Dietz, K.; Schmidt, C.; Warken, M.; Hess, B. A. *J Phys B Atom Mol Opt Phys* 1993, 1897; Dietz, K.; Schmidt, C.; Warken, M.; Hess, B. A. *Chem Phys Lett* 1994, 220, 397.