

INVITED TOPICAL REVIEW

Quantitative quantum chemistry

Trygve Helgaker^{a*}, Wim Klopper^b and David P. Tew^b

^aDepartment of Chemistry, Centre for Theoretical and Computational Chemistry, University of Oslo, Oslo, Norway; ^bLehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie, Universität Karlsruhe (TH), Karlsruhe, Germany

(Received 3 April 2008; final version received 1 June 2008)

We review the current status of quantum chemistry as a predictive tool of chemistry and molecular physics, capable of providing highly accurate, quantitative data about molecular systems. We begin by reviewing wave-function based electronic-structure theory, emphasizing the N -electron hierarchy of coupled-cluster theory and the one-electron hierarchy of correlation-consistent basis sets. Following a discussion of the slow basis-set convergence of dynamical correlation and basis-set extrapolations, we consider the methods of explicit correlation, from the early work of Hylleraas in the 1920s to the latest developments in such methods, capable of yielding high-accuracy results in medium-sized basis sets. Next, we consider the small corrections to the electronic energy (high-order virtual excitations, vibrational, relativistic, and diagonal Born–Oppenheimer corrections) needed for high accuracy and conclude with a review of the composite methods and computational protocols of electronic-structure theory.

Keywords: electronic-structure theory; basis-set extrapolation; explicit correlation; coupled-cluster theory

1. Introduction

Over the last 50 years, quantum chemistry has developed into an indispensable part of modern chemistry and molecular physics, providing us with concepts and ideas that are central to our modern understanding of chemical systems and processes. Equally important, the development of quantum chemistry has provided us with a versatile tool, whose application in many areas of chemistry and physics provides us with quantitative data about chemistry, helping us predict, confirm, or reject experimental observations and measurements. It is our purpose here to review quantum chemistry as such a quantitative tool, capable of producing numerical results that agree quantitatively with the true or exact results, to within some estimated uncertainty.

The evolution of quantum chemistry into a quantitative tool over the last 50 years has been made possible thanks to several important methodological advances [1]. Particularly important have been the development of coupled-cluster theory since the early 1980s by Bartlett and others [2] and the development of the correlation-consistent basis sets since the late 1980s by Dunning and others [3]. With additional developments of coupled-cluster theory to all levels of excitation, explicit correlation and basis-set extrapolation, it has become possible to achieve the

goal of a quantitative theory, for a wide range of problems in chemistry and molecular physics. Indeed, any serious attempt at an accurate, quantitative study of molecular systems requires a combination of all these techniques, as reflected in the development of several prescriptions or recipes for achieving ‘chemical accuracy’. These prescriptions have developed from the realisation that, except in very rare cases, high accuracy cannot be achieved in a single calculation: we are still not able to construct a single approximate electronic wave function from which we may extract all relevant information about the system under study, to sufficient accuracy. Instead, the information about the system must be extracted by a combination of different techniques, using various additivity schemes. The success of such additivity schemes is perhaps one of the more surprising developments in quantum chemistry over the last 10–15 years.

In our overview of quantitative chemistry, we begin in Section 2 by reviewing the fundamentals of coupled-cluster theory, the general framework for constructing approximate wave functions for high-accuracy quantum-chemical calculations. Next, in Section 3, we discuss the construction of basis sets for such calculations, paying special attention to the slow convergence of the electronic energy with respect to the size of the basis set, caused by the singularity in the two-electron

*Corresponding author. Email: trygve.helgaker@kjemi.uio.no

Coulomb operator. Although the slow convergence may be accelerated by means of basis-set extrapolation, we consider, in Section 4, a much more satisfactory solution – namely, the use of explicitly correlated wave functions, which depend explicitly on the interelectronic distances. Next, in Section 5, we discuss various (usually) small corrections to our calculations, necessary for high-precision work: vibrational corrections, high-order correlation corrections, core-correlation corrections, relativistic and diagonal Born–Oppenheimer corrections. Finally, in Section 6, we review the various composite methods that have been developed for high-precision quantum-chemical calculations.

2. Coupled-cluster theory

The standard quantum-chemistry approach to high accuracy is based on the concept of mean-field theory, where the electronic state is represented by a set of occupied molecular orbitals (MOs). This description is next refined by introducing virtual electronic excitations: double excitations, triple excitations, quadruple excitations, and so on, into a set of virtual MOs. In the limit of arbitrary high virtual excitations and complete virtual orbital spaces, the exact solution to the Schrödinger equation is obtained. In practice, however, this limit is never reached and we must instead be content with an approximate description of the electronic system. Specifically, we must be content with a truncation of the excitation level (omitting, for example, all triple and higher excitations) and the use of a finite set of virtual orbitals. How these truncations are performed will dictate the accuracy of the calculations.

2.1. Standard coupled-cluster theory

In traditional ground-state coupled-cluster theory, we begin our description of the electronic system by generating the Hartree–Fock wave function, obtained by minimizing the expectation value of the electronic energy of a determinantal wave function (i.e. an antisymmetrised product of spin orbitals) with respect to orthonormal transformations of the orbitals (assuming normalisation):

$$\langle \text{HF} | \hat{H} | \text{HF} \rangle = \min_{\Psi_{\text{det}}} \langle \Psi_{\text{det}} | \hat{H} | \Psi_{\text{det}} \rangle. \quad (1)$$

The Hartree–Fock wave function $|\text{HF}\rangle$ is thus the variationally best one-determinantal approximation to the electronic ground state of the system.

In a given orbital basis, the Hartree–Fock description divides the orbital space into a set of occupied spin orbitals (denoted by indices I, J, K and L) and a set of virtual spin orbitals (denoted by indices A, B, C , and D). From the Hartree–Fock determinant, we may generate any other determinant in this basis by the application of an excitation operator

$$\begin{aligned} \hat{\tau}_I^A &= a_A^\dagger a_I, & \hat{\tau}_{IJ}^{AB} &= a_A^\dagger a_I a_B^\dagger a_J, \\ \hat{\tau}_{IJK}^{ABC} &= a_A^\dagger a_I a_B^\dagger a_J a_C^\dagger a_K, \dots, \end{aligned} \quad (2)$$

where a_I annihilates an electron in spin orbital I , while a_A^\dagger creates an electron in spin orbital A . The creation and annihilation operators satisfy the usual algebra $[a_P, a_Q^\dagger]_+ = \delta_{PQ}$. Introducing the generic notation $\hat{\tau}_\mu$ for excitation operators and noting that $[\hat{\tau}_\mu, \hat{\tau}_\nu] = 0$, we may express the coupled-cluster wave function in the standard exponential form

$$|\text{CC}\rangle = \exp(\hat{T})|\text{HF}\rangle, \quad (3)$$

where the cluster operator

$$\hat{T} = \sum_{\mu} t_{\mu} \hat{\tau}_{\mu} \quad (4)$$

is a linear combination of excitation operators $\hat{\tau}_\mu$, each multiplied by an associated cluster amplitude t_μ . To determine these amplitudes and calculate the electronic energy, we insert the wave function Equation (3) into the Schrödinger equation $\hat{H}|\text{CC}\rangle = E|\text{CC}\rangle$ and multiply from the left with $\exp(-\hat{T})$, yielding the similarity-transformed Schrödinger equation

$$\exp(-\hat{T})\hat{H}\exp(\hat{T})|\text{HF}\rangle = E|\text{HF}\rangle. \quad (5)$$

Finally, multiplying this equation from the left by the Hartree–Fock state $\langle \text{HF} |$ and by the excited states

$$\langle \mu | = \langle \text{HF} | \tau_{\mu}^{\dagger}, \quad (6)$$

we obtain the equations

$$\langle \text{HF} | \exp(-\hat{T})\hat{H}\exp(\hat{T})|\text{HF}\rangle = E, \quad (7)$$

$$\langle \mu | \exp(-\hat{T})\hat{H}\exp(\hat{T})|\text{HF}\rangle = 0. \quad (8)$$

By projecting against all possible states $\langle \mu |$ that may be generated by the excitation operators $\hat{\tau}_\mu$ present in the cluster operator Equation (4) according to Equation (6), we may uniquely determine the coupled-cluster state from the non-linear Equation (8) and then obtain the energy from Equation (7). Because of the exponential parameterisation of the coupled-cluster state, Equation (3), the calculated energy E is size-extensive. For a general discussion of coupled-cluster theory and the coupled-cluster equations, see [1].

With all possible excitations included in the cluster operator of Equation (4), coupled-cluster theory provides an exact solution to the Schrödinger equation in a given orbital basis, equivalent to a full configuration-interaction (FCI) treatment in the same basis. However, except in a very small orbital basis, this approach is prohibitively expensive. Useful approximate solutions are instead obtained by truncating the cluster operator of Equation (4). Introducing the notation

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots, \quad (9)$$

where \hat{T}_n contains all possible n -excitation operators, we arrive at the following levels of coupled-cluster theory:

$$|\text{CCS}\rangle = \exp(\hat{T}_1)|\text{HF}\rangle, \quad (10)$$

$$|\text{CCSD}\rangle = \exp(\hat{T}_1 + \hat{T}_2)|\text{HF}\rangle, \quad (11)$$

$$|\text{CCSDT}\rangle = \exp(\hat{T}_1 + \hat{T}_2 + \hat{T}_3)|\text{HF}\rangle, \quad (12)$$

and so on. Thus, at the lowest level of coupled-cluster singles (CCS) theory, we include all possible single excitations in the cluster operator; in coupled-cluster singles-and-doubles (CCSD) theory [4], we include also all double excitations; whereas, in coupled-cluster singles–doubles–triples (CCSDT) theory [5,6], all triple excitations are also included. In general, higher-order excitations are less important than the lower-order ones, except that the singles play a relatively minor role, effectively performing small orbital adjustments to the doubles and higher excitations. In Figure 1, we have plotted the error in the energy at various levels of coupled-cluster theory relative to the FCI energy, illustrating how the error is reduced by about the same factor at each new level of theory.

Note that, at each level of coupled-cluster theory, we include through the exponential parameterisation of Equation (3) all possible determinants that can be generated within a given orbital basis – that is, all determinants that enter the FCI wave function in the same orbital basis. Thus, the improvement in the sequence CCSD, CCSDT, and so on does not occur because more determinants are included in the description but from an improved representation of their expansion coefficients. For example, in CCS theory, the doubly-excited determinants are represented by $\frac{1}{2}\hat{T}_1^2|\text{HF}\rangle$, whereas the same determinants are represented by $(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2)|\text{HF}\rangle$ in CCSD theory. Thus, in CCSD theory, the weight of each doubly-excited determinant is obtained as the sum of

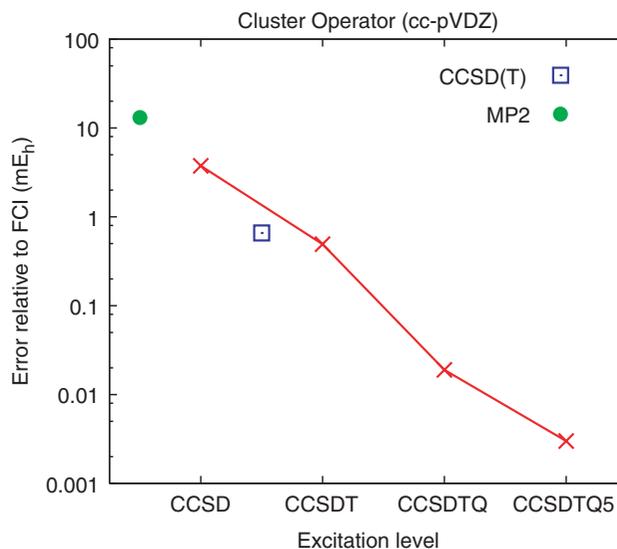


Figure 1. Convergence of the coupled-cluster hierarchy for the water molecule in the cc-pVDZ basis. At each level of theory, the error in the energy has been plotted relative to the FCI energy in the same basis [1].

a connected doubles contribution from \hat{T}_2 and a disconnected singles contribution from $\hat{T}_1^2/2$. This parameterisation of the wave function is not only more compact than the linear parameterisation of configuration-interaction (CI) theory $|\text{CI}\rangle = \sum_{\mu} c_{\mu}|\mu\rangle$, it also ensures size-extensivity of the calculated electronic energy.

2.2. Modified and approximate coupled-cluster theories

Over the years, a number of modifications and alternative formulations of coupled-cluster theory have been proposed. One popular such modification of standard coupled-cluster theory is quadratic CI (QCI) theory and – in particular, the popular QCI singles-and-doubles (QCISD) model [7]. The QCISD model was originally conceived as a size-extensive revision of the CI singles-and-doubles (CISD) model but it is closely related to the CCSD model, from which it may be obtained by the omission of a few terms in the CCSD equations. Since the QCISD model neither represents an improvement on the standard CCSD model nor a computational simplification, its use should be discouraged.

More important are several approximate treatments of coupled-cluster theory based on perturbation theory. Thus, although the coupled-cluster hierarchy converges rapidly with increasing excitation level as illustrated in Figure 1, its cost also increases rapidly with increasing size of the system and increasing size of

the basis set. Specifically, the cost of CCSD theory (in a traditional implementation) scales as $N_o^2 N_v^4$, where N_o is the number of occupied orbitals and N_v is the number of virtual orbitals. Clearly, CCSD calculations become expensive in the large basis sets that are needed for accurate work (as discussed in Section 3). Moreover, in CCSDT theory, the corresponding cost is $N_o^2 N_v^6$, clearly prohibitive for routine work, even in rather small basis sets.

To reduce cost while retaining the essential elements of coupled-cluster theory (the exponential parameterisation of the wave function in terms of a cluster operator), a useful strategy is to decrease the cost of the optimisation of the cluster amplitudes by invoking perturbation theory. Thus, in second-order Møller–Plesset theory (MP2) [8,9], the doubles amplitudes of CCSD theory are not determined from the projected equations [Equation (8)] but from perturbation theory. This approach not only reduces the cost of the time-consuming step from $N_o^2 N_v^4$ to $N_o^2 N_v^3$, it also makes the calculation noniterative. The resulting MP2 model is less robust than the CCSD model and typically overestimates the effect of the doubles by 10%, but represents an attractive alternative to full CCSD theory, useful in many situations.

Similarly, we may simplify CCSDT theory by estimating the contribution from the triples by perturbation theory, while the dominant (and less expensive) doubles amplitudes are optimised in the usual manner of coupled-cluster theory. This can be achieved in several different ways, the most popular approach being the CCSD(T) model, often referred to as the ‘gold standard’ of quantum chemistry. The CCSD(T) model [10], whose time-consuming step scales as $N_o^2 N_v^5$ (noniteratively) rather than as $N_o^2 N_v^6$ (iteratively) as in full CCSDT theory, is indeed a highly successful method, often yielding results in good agreement with experimental measurements, provided the calculations are carried out in a sufficiently large orbital basis. Perturbation-based approximations similar to CCSD(T) have also been derived for excited-state energies in the context of equation-of-motion coupled-cluster theory [11–14]. As in MP2 theory, the perturbative treatment of the cluster amplitudes overestimates their true contribution to the total energy, typically by 10%. This overestimation of the triples contribution is the key to success of the CCSD(T) theory, as it serves to cancel the error incurred by neglect of quadruples and higher excitations. No such fortuitous cancellation occurs in full CCSDT theory, making this more rigorous (and more robust) model often less accurate than the CCSD(T) theory.

A great advantage of the perturbative treatments of MP2 and CCSD(T) theories is their noniterative nature, greatly reducing the cost of the calculations. Unfortunately, this noniterative treatment is unsuitable for the calculation of frequency-dependent molecular properties, using the general framework of response theory. In such cases, the coupled-cluster rank- n models CC_n are more appropriate. In these models, the highest-order cluster amplitudes are approximated perturbatively, but in an iterative rather than noniterative manner. Particularly important are the CC2 model [15] and the CC3 model [16], as approximations to the CCSD and CCSDT models, respectively.

For a review of the application of orbital-based coupled-cluster theory to the calculation of molecular energies and properties, see [17].

3. One-electron basis sets

In coupled-cluster theory, the atomic basis functions should not only provide an accurate description of the occupied MOs but also provide a flexible set of virtual MOs in terms of which the cluster operator and hence the wave function may be expanded. To ensure a rapid convergence of the results (with respect to the number of MOs included in the virtual space), these orbitals should be chosen in an optimal manner. To see how this may be accomplished, it is useful to consider the simplest of all two-electron systems: the ground-state helium atom.

3.1. The partial-wave expansion

The ground-state helium wave function may be written as a product of a symmetric spatial part and an antisymmetric spin part [1]

$$\Psi(x_1, x_2) = \Psi(r_1, r_2) \frac{1}{2^{1/2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)], \quad (13)$$

where r_i is the position vector of electron i and x_i the combined spatial and spin coordinates of this electron. In terms of a complete set of AOs $\phi_p(r)$, the spatial part of the two-electron wave function may be expanded as

$$\Psi(r_1, r_2) = \sum_{p \geq q} c_{pq} [\phi_p(r_1)\phi_q(r_2) + \phi_q(r_1)\phi_p(r_2)]. \quad (14)$$

We shall here consider orthonormal AOs of the general form

$$\psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) Y_{lm}(\theta, \varphi), \quad n > l \geq |m| \geq 0, \quad (15)$$

where the $Y_{lm}(\theta, \varphi)$ are the spherical-harmonic functions and where, for each angular momentum l , the $R_{nl}(r)$ constitute a complete orthonormal radial set

$$\int_0^\infty R_{nl}^*(r)R_{n'l}(r)r^2 dr = \delta_{nn'}. \quad (16)$$

In terms of such AOs, the spatial part of the ground-state helium wave function in Equation (13) may be decomposed into partial waves

$$\Psi(r_1, r_2, \theta_{12}) = \sum_{l=0}^{\infty} \Psi_l(r_1, r_2)P_l(\cos \theta_{12}), \quad (17)$$

where the two-electron angular part $P_l(\cos \theta_{12})$ is the Legendre polynomial of degree l in the angle θ_{12} between the position vectors \mathbf{r}_i of the two electrons, while the two-electron radial part $\Psi_l(r_1, r_2)$ is expanded in symmetric products of the one-electron radial forms of Equation (15)

$$\begin{aligned} \Psi_l(r_1, r_2) \\ = \sum_{n_1 \geq n_2 \geq l} c_{n_1, n_2, l} [R_{n_1, l}(r_1)R_{n_2, l}(r_2) + R_{n_2, l}(r_1)R_{n_1, l}(r_2)]. \end{aligned} \quad (18)$$

An explicit orthonormal set of radial forms is provided by the Laguerre functions (also known as Sturmians), whose radial part is proportional to $(2\zeta r)^l L_{n-l-1}^{2l+2} \times (2\zeta r) \exp(-\zeta r)$, where $L_n^\alpha(x)$ is an associated Laguerre polynomial [1]. Unlike the hydrogenic functions, the Laguerre functions form a complete set for a fixed exponent ζ . Noting that the Laguerre functions share the nodal structure of the hydrogenic functions, we shall refer to n as the principal quantum number.

The convergence of the partial-wave expansion of the helium atom has been studied in great detail. In particular, Schwartz considered the rate of convergence of the second-order helium energy in a perturbative treatment with the bare-nucleus Hamiltonian (in atomic units) $H_0 = -\frac{1}{2}\nabla^2 - \frac{1}{2}\nabla^2 - Z/r_1 - Z/r_2$ as the zero-order operator with nuclear charge Z and with $V = r_{12}^{-1}$ as the perturbation operator [18]. Since the simple coordinate rescaling $\mathbf{r}_i \rightarrow Z^{-1}\mathbf{r}_i$ yields the expansion $E = \sum_{n=0}^{\infty} Z^{2-n} E^{(n)}$ valid for all two-electron atoms, this perturbation theory is known as the $1/Z$ expansion [19]. Decomposing the second-order energy in partial-wave contributions $E^{(2)} = \sum_{l=0}^{\infty} E_l^{(2)}$, Schwartz obtained [18]

$$E_l^{(2)} = -\frac{45}{256} \left(l + \frac{1}{2}\right)^{-4} + \frac{225}{1024} \left(l + \frac{1}{2}\right)^{-6} + \dots, \quad (19)$$

demonstrating that the contributions from each partial wave converge very slowly with increasing l . Moreover, the slow convergence arises from the

singularity in the perturbation operator r_{12}^{-1} at coalescence.

The analysis of Schwartz was restricted to the $1/Z$ perturbative expansion of the ground-state helium wave function. Following the work of Klahn and Morgan [20], the slightly more relevant case of a variationally optimised helium ground-state wave function was studied by Hill [21], who found that the contributions from each partial wave to the total (variationally determined) energy $E^{\text{CI}} = \sum_l E_l^{\text{CI}}$ are given as

$$E_l^{\text{CI}} = -3C_1 \left(l + \frac{1}{2}\right)^{-4} - 4C_2 \left(l + \frac{1}{2}\right)^{-5} + \dots, \quad (20)$$

where the coefficients depend on the value of the electronic wave function at the points of coalescence $C_1 = 2\pi^2 \int_0^\infty |\Psi(r, r, 0)|^2 r^5 dr \approx 0.024742$ and $C_2 = (12\pi/5) \int_0^\infty |\Psi(r, r, 0)|^2 r^6 dr \approx 0.007747$. Note that the CI expansion Equation (20) differs from the perturbation-theory expansion Equation (19) in that also odd-order terms contribute to the energy. Otherwise, the main result is the same – namely, that the partial-wave contributions to the ground-state electronic energy converge as $(l + 1/2)^{-4}$ with increasing l .

Subsequently, Kutzelnigg and Morgan studied the partial-wave convergence in greater detail, considering more general two-electron atomic states [22]. More specifically, for states of orbital-angular momentum L , these authors found that the energy converges as $(l + 1/2)^{-4}$ for singlet states of parity $(-1)^L$ (natural-parity states), as $(l + 1/2)^{-6}$ for triplet states, and as $(l + 1/2)^{-8}$ for singlet states of parity $(-1)^{L+1}$ (unnatural-parity states). They also established that, although only even-order terms appear in the second-order energy studied by Schwartz, odd-order terms appear to higher orders in the perturbation, in agreement with the result of Hill for CI wave functions [21].

Let us assume that we have calculated the electronic energy of the ground-state helium atom by truncating the partial-wave expansion, omitting all terms $l > \mathcal{L}$. We may then estimate the error in the resulting electronic energy $\Delta E_{\mathcal{L}} = E - \sum_{l=0}^{\mathcal{L}} E_l$ by summing over the leading term of the contribution from each omitted partial wave:

$$\begin{aligned} \Delta E_{\mathcal{L}} &= 3C_1 \sum_{l=\mathcal{L}+1}^{\infty} \left(l + \frac{1}{2}\right)^{-4} \\ &\approx 3C_1 \int_{\mathcal{L}+1/2}^{\infty} \left(l + \frac{1}{2}\right)^{-4} dl = C_1 (\mathcal{L} + 1)^{-3}. \end{aligned} \quad (21)$$

Clearly, the error in the energy of the partial-wave expansion converges slowly with increasing \mathcal{L} . Similarly, we find that the error is proportional to $(\mathcal{L} + 1)^5$ for triplet states and to $(\mathcal{L} + 1)^7$ for singlet

states of unnatural parity. Comparing the C_1 coefficient of the CI expansion (≈ 0.024742) with the corresponding coefficient in the $1/Z$ perturbation expansion in Equation (19) ($45/768 = 0.05859375$), we see that the convergence with l is different for the two methods and that the CI error is smaller than the error in perturbation theory. This interference between basis-set and correlation treatment is also present when comparing Møller–Plesset, coupled-cluster and CI theories.

The results of Schwartz, Hill, Kutzelnigg, and others demonstrate that the electronic energy of the helium atom – in particular, the ground-state helium atom – converges very slowly with increasing angular momentum of the partial-wave expansion. Clearly, the brute-force evaluation of the electronic energy to a given precision (e.g. to within $1 mE_h$) by the inclusion of sufficiently many partial waves in the expansion in Equation (17) is a daunting task. On the other hand, we may use our knowledge of the convergence rate of the expansion to extrapolate to the limit of a complete expansion, from results obtained with low maximum angular momentum \mathcal{L} . This approach has been attempted by a number of authors since the asymptotic form was established by Schwartz in the 1960s. A more fundamental solution is to employ a different ansatz for our wave function, which contains r_{12} terms explicitly, as discussed later in this paper.

3.2. The principal expansion

The analysis of the partial-wave expansion of the helium atom is relevant only to the extent that each partial wave $\Psi_l(r_1, r_2)$ in Equation (17) has been calculated accurately. In practice, the expansion of $\Psi_l(r_1, r_2)$ in a finite set of one-electron functions introduces an additional error into our description, beyond the truncation of the partial-wave expansion at $l = \mathcal{L}$. It then becomes relevant to study how the truncation of the one-electron basis of Equation (15) affects the calculated energy, rather than to focus on the partial-wave expansion. To discuss the convergence of the helium energy with respect to the size of the one-electron basis, we must specify the form of the one-electron basis functions in Equation (15). A useful basis are the natural orbitals (NOs) of the system [23], whose radial form is determined by diagonalisation of the one-electron density matrix in a large set of AOs (e.g. Laguerre functions). For helium, the NOs have the same nodal structure as the hydrogenic and Laguerre functions – in particular, the total number of nodes (radial and angular) is given by $n - 1$, where n is the principal quantum number of the NO.

In the basis of NO functions, Carroll *et al.* [24] found that, to a good approximation, each NO provides an amount of energy to the ground-state helium atom proportional to

$$\varepsilon_{nlm} \approx -a_l \left(n - \frac{1}{2}\right)^{-6}, \quad (22)$$

where $a_0 \approx 0.29$ for $l = 0$ and $a_l \approx 0.22$ for $l > 0$. Clearly, for $l > 0$ and large n , the energy contribution of each orbital is determined solely by the principal quantum number of the NO – that is, by the number of nodes in the NO. In expanding the wave function, therefore, the NOs should be included in order of increasing n . Moreover, we should always include all n^2 orbitals belonging to the same shell simultaneously, noting that these all make the same contribution to the energy. In this principal expansion of the helium wave function [1,25], each new shell of NOs contribute an amount of energy proportional to n^{-4} to the ground-state system. From the numerical work of Carroll *et al.* [24] we may then postulate the following principal expansion of the energy

$$\varepsilon_n = \sum_{l=0}^{n-1} \sum_{m=-l}^l \varepsilon_{nlm} = A_1 n^{-4} + A_2 n^{-5} + \dots, \quad (23)$$

where A_i are constants independent of n . Very recently, Kutzelnigg has re-examined the second-order energy of the $1/Z$ expansion theoretically, in a semi-rigorous manner, confirming that each shell of NOs contribute an amount of energy proportional to n^{-4} in agreement with Equation (23) [26]. We here add that the empirical expression in Equation (22) yields the following asymptotic energy contribution from each partial wave:

$$E_l = \sum_{n=l+1}^{\infty} \sum_{m=-l}^l \varepsilon_{nlm} = -\frac{2a_l}{5} \left(l + \frac{1}{2}\right)^{-4} + \dots \quad (24)$$

in agreement with the rigorous expression for the partial-wave expansion of Equation (20).

We now assume that we have calculated the helium energy by including all NOs of principal quantum number $n \leq \mathcal{N}$. As for the partial-wave expansion, we may estimate the error $\Delta\varepsilon_{\mathcal{N}} = E - \sum_{n=1}^{\mathcal{N}} \varepsilon_n$ by adding up the contributions from all omitted shells in the principal expansion:

$$\Delta\varepsilon_{\mathcal{N}} = \sum_{n=\mathcal{N}+1}^{\infty} A_1 n^{-4} \approx \frac{A_1}{3} \mathcal{N}^{-3}. \quad (25)$$

For a given maximum principal quantum number \mathcal{N} , the remaining error is thus proportional to the

inverse cube of \mathcal{N} . We also note that the total number of orbitals in the calculation is given by $N = \sum_{n=1}^{\mathcal{N}} n^2 \approx \mathcal{N}^3$. Combining these two observations, we find that, for the principal expansion, the error in the energy is inversely proportional to the number of NOs in the calculation [27,28].

The analysis of the convergence of the helium energy in terms of NOs may seem of rather limited interest. It turns out, however, that the results for the principal expansion of the helium energy are valid, in some generalised sense, also for molecular systems, when the MOs are expanded in Gaussians. Indeed, we shall see that, with sufficiently many Gaussians N included in our calculations, the error in the energy converges in the same manner as in the helium atom – namely, as N^{-1} . Moreover, this convergence is sufficiently smooth to form the basis for a successful extrapolation of the total electronic energy.

3.3. Molecular Gaussian basis sets

Having discussed the convergence of the correlation energy of the helium atom with respect to increasing the one-electron basis set, let us now consider how such sets should be constructed for general electronic systems. In practically all molecular wave-function calculations, basis sets of Gaussian orbitals are used [1]. In practice, these functions take the general form

$$\chi_{alm}^{\text{GTO}}(r, \theta, \varphi) = N_{al}^{\text{GTO}} r^l \exp(-\alpha r^2) Y_{lm}(\theta, \varphi), \quad (26)$$

where α is the orbital exponent and N_{al}^{GTO} is the normalisation constant. In these basis sets, no explicit use is made of the principal quantum number. Instead, flexibility in the radial part of the one-electron basis functions is generated by the linear combination of Gaussian-type orbitals (GTOs) with the same l and m quantum numbers but different exponents. In general, the one-electron basis-functions from which the MOs are constructed are not simple primitive Gaussians of the form given in Equation (26) but fixed linear combinations of primitive Gaussians

$$\chi_{alm}^{\text{GTO}}(r, \theta, \varphi) = \sum_i c_{ilm} \chi_{\alpha_i, l, m}^{\text{GTO}}(r, \theta, \varphi). \quad (27)$$

In the following, we shall discuss basis sets of this general form, with emphasis on the requirements for the accurate calculation of the correlation energy.

In most calculations, the GTOs are atom-fixed functions, obtained from atomic calculations. For calculations at the uncorrelated Hartree–Fock level of theory or for density-functional theory (DFT), the basis set should only be sufficiently flexible to represent accurately the occupied MOs and the one-electron

density matrix. We shall not consider the construction of such basis sets in detail, noting that these are usually obtained from accurate atomic calculations [1]. The resulting basis functions of atomic orbitals (AOs) are then augmented with polarisation functions – that is, by functions of angular momentum higher than those of the occupied AOs. These polarisation functions give the basis set the flexibility needed to describe the distortion of the atomic density in a molecular environment.

3.4. Atomic natural orbitals

For the accurate calculation of dynamical correlation, there is an additional requirement on the basis set – namely, it must also provide a flexible set of virtual MOs for the accurate expansion of the cluster operators of Equation (4). Originally, such basis sets were generated in a somewhat *ad hoc* manner, typically by adding a few high angular-momentum Gaussians to basis sets generated for uncorrelated calculations. In 1987, Almlöf and Taylor proposed to generate basis sets for correlated calculations from atomic natural orbitals (ANOs), obtained by diagonalising the one-electron density matrix from an atomic calculation [29,30]. In general, natural orbitals are the eigenfunctions of the one-electron density matrix, and the associated eigenvalues (the occupation numbers) represent a measure of the contributions that the different natural orbitals make to the density matrix [23]. Natural orbitals with small occupation numbers can thus be presumed to be less important than those with large occupations. The occupation numbers of the ANOs thus provide a natural criterion for which AOs to include in a given molecular calculation: we simply include all ANOs with occupation numbers greater than a given threshold.

Consider the construction of ANOs for the carbon atom. We begin by carrying out a correlated calculation on the 3P ground state of carbon in a large basis of primitive Gaussians, for example $13s8p6d4f2g$. Next, we diagonalise the density matrix and obtain a set of ANOs, which are then sorted in order of decreasing occupation number. These ANOs are then the AOs for the subsequent molecular calculations. By including all orbitals of occupation numbers greater than 10^{-2} , 10^{-3} , 10^{-4} , and 10^{-5} , respectively, we arrive at the contracted hierarchical ANO basis sets $[2s1p] \subset [3s2p1d] \subset [4s3p2d1f] \subset [5s4p3d2f1g]$. Of these ANO basis sets, the $[2s1p]$ basis is a minimal one, containing only orbitals from the occupied shells in Hartree–Fock theory. Such basis sets are unsuitable for the calculation of correlation energies. In the larger ANO basis

sets, more and more *correlating orbitals* are introduced, providing an increasingly flexible representation of the cluster operator \hat{T} .

The ANO basis sets may be characterised in a simple manner in terms of their nodal structure, noting that the $[2s1p]$, $[3s2p1d]$, $[4s3p2d1f]$, and $[5s4p3d2f1g]$ basis sets consist of all ANOs containing up to one, two, three, and four nodal planes, respectively. Recalling that $n-1$ (the principal quantum number minus one) is equal to the number of nodal planes in the hydrogenic orbitals and Laguerre functions, we see how the ANOs provide a hierarchical structure very similar to that observed in the principal expansion of the helium energy. The analogy with the principal expansion is complete in the sense that, with each new ANO basis, the new functions introduced all make similar contributions to the correlation energy, which are smaller than those of the old functions. The ANOs thus provide a highly systematic scheme for the construction of basis sets for correlated calculations.

3.5. Correlation-consistent basis sets

In the ANOs, basis functions are selected based on occupation numbers. However, it is observed that ANOs grouped according to occupation numbers also provide similar contributions to the correlation energy [30]. This latter observation suggests that it may be possible to build basis sets for electron correlation by relying on the energy criterion directly. This is the approach taken with the *correlation-consistent* basis sets, where each correlating orbital is represented by a single primitive Gaussian chosen to maximize its contribution to the correlation energy, and where all orbitals that make similar contributions to the correlation energy are added simultaneously [31]. A hierarchy of basis sets may then be set up that is correlation consistent in the sense that each basis set contains all correlating orbitals that lower the energy by similar amounts as well as all orbitals that lower the energy by larger amounts. Not surprisingly, this approach leads to basis sets whose structure is similar to that of the ANOs. The main advantage of the correlation-consistent scheme over the ANO scheme is that it allows us to construct hierarchical basis sets from smaller primitive sets, since we no longer need to expand the different correlating functions from the same primitive set. On the other hand, unlike for the ANOs, the correlation-consistent orbitals do not provide a subset structure.

Let us consider the structure of the correlation-consistent basis sets in more detail. The suite of

correlation-consistent polarised valence basis sets is denoted by cc-pVXZ [31,32], where X is the *cardinal number*. The cardinal number is two for basis sets of double-zeta quality, three for sets of triple-zeta quality, and so on. For the carbon atom, for example, the double-zeta basis cc-pVDZ with $X=2$ contains the functions $[3s2p1d]$, the triple-zeta basis cc-pVTZ with $X=3$ contains the functions $[4s3p2d1f]$, the quadruple-zeta set cc-pVQZ contains $[5s4p3d2s1g]$, and so on. In these basis sets, the occupied orbitals are contracted combinations of primitive Gaussians, whereas the correlating orbitals are uncontracted primitive Gaussians. Moreover, the primitive Gaussians used to construct the correlation-consistent basis sets cc-pVXZ are different for different cardinal numbers X .

The correlation-consistent polarised valence basis sets cc-pVXZ are designed for the calculation of the valence-electron correlation energy only. To recover core correlation, the more flexible correlation-consistent polarised core-valence basis sets cc-pCVXZ basis sets must be used [33]. These basis sets have been obtained from the cc-pVXZ valence basis sets by the addition of Gaussians with large exponents, as required for calculation of all-electron correlation energies. Moreover, for applications where flexibility is required in the outer valence region, the augmented correlation-consistent basis sets aug-cc-pVXZ [34], which have been obtained from the standard sets cc-pVXZ by the addition of diffuse functions, are appropriate.

3.6. Basis-set extrapolation

With the introduction of the ANO and correlation-consistent basis sets, it became possible to carry out systematic calculations of correlation energies, where the exact result is approached in a systematic manner. Nevertheless, in view of the slow FCI basis-set convergence observed for the helium atom in the NO basis, it is perhaps not surprising that we often cannot converge our results sufficiently well, at least if traditional orbital-based coupled-cluster methods are used. However, in such cases, we can usually improve upon our results by carrying out extrapolations to the basis-set limit.

To carry out extrapolations, some simple analytical model for the convergence of the energy is required. In our discussion of the principal expansion of the helium atom in Section 3.2, we found that, if we include in our wave function all NOs of principal quantum number $n \leq \mathcal{N}$, then the error in the energy $\Delta E_{\mathcal{N}}$ is asymptotically proportional to \mathcal{N}^{-3} , see Equation (25). Moreover, this convergence rate

should be the fastest possible, in an orbital-based calculation. For helium, the correlation-consistent basis set cc-pVXZ has the same composition and the same size as the NO basis with $n \leq \mathcal{N} = X$. Given that the cc-pVXZ basis set has been designed to recover as much as possible of the correlation energy for a given basis-set size, we would expect the truncation error to behave in the same manner as for the principal expansion, leading us to the following simple relation between the exact energy and the energy obtained in an energy-optimised basis of cardinal number X proposed by Helgaker and co-workers X [35,36]:

$$E_\infty = E_X + AX^{-3} + \mathcal{O}(X^{-4}), \quad (28)$$

which contains unknown terms of order X^{-3} and higher. By carrying out an additional calculation of cardinal number $X-1$, we may eliminate the third-order term, obtaining

$$E_\infty = \frac{X^3 E_X - (X-1)^3 E_{X-1}}{X^3 - (X-1)^3} + \mathcal{O}(X^{-4}), \quad (29)$$

which contains unknown terms of order X^{-4} and higher only. Such an extrapolation may significantly improve the accuracy of an orbital-based calculation of the electronic energy, reducing the error in the calculation significantly, often by an order of magnitude or more. In Figure 2, we have plotted normal distributions of errors in atomisation energies for the standard models

of quantum chemistry, calculated with and without extrapolation according to Equation (29).

Although the above extrapolation scheme appears to be a reasonable one, it should be kept in mind that our arguments have been of a somewhat heuristic nature, based on the assumed equivalence of the NO and energy-optimised correlation-consistent basis sets for the ground-state helium atom. For other atoms and for molecular systems, the electronic system is more complicated. First, for such systems, the identification of the principal quantum number with the cardinal number is less obvious – in particular, for systems containing atoms from various rows of the periodic table. Second, such systems contain triplet- as well as singlet-coupled electron pairs. We may then generalize the scheme slightly, extrapolating the singlet- and triplet-coupled pairs separately, using the formula [37]

$$E_\infty = E_X + AX^{-s} + \mathcal{O}(X^{-(s+1)}) \quad (30)$$

with $s=3$ for singlet pairs and $s=5$ for triplet pairs. However, since the convergence of the total energy will be dominated by the slower convergence of the singlet pairs, reasonably accurate extrapolations are obtained by applying the singlet formula to the total energy.

Third, many-electron systems contain also contributions from triple and higher excitations. Little is known about the convergence of the energy contributions from higher virtual excitations except

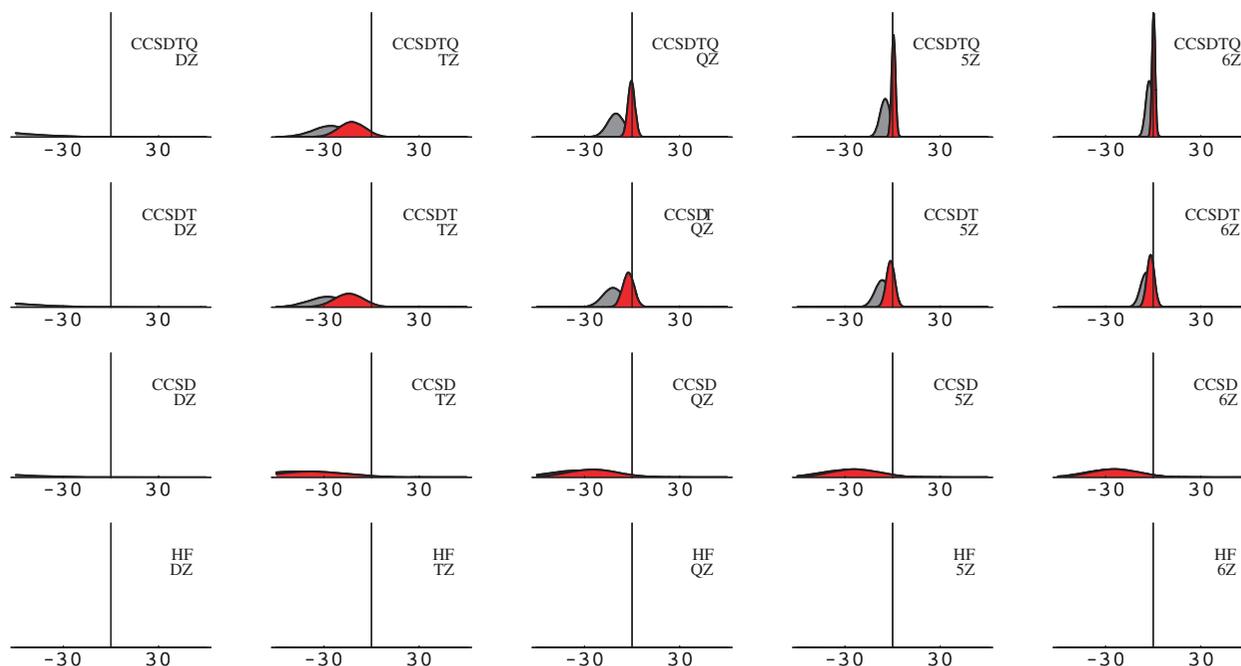


Figure 2. Normal distributions of errors in atomisation energies (kJ mol^{-1}), calculated with all electrons correlated, without (grey) and with (red) X^{-3} -extrapolation of Equation (29). For details on systems, see [1].

that these are usually much smaller than the double contributions. In general, therefore, the above formula Equation (29) can be applied directly to the total correlation energy of the system.

Finally, in a molecular system, basis-set convergence may be affected by other aspects of the electronic-structure, such as the polarisation of the atomic charge distributions in the presence of other atoms and external electromagnetic fields. Usually, such polarisations of the charge distribution is easier to describe in an orbital basis than the correlation problem. Provided the basis set has been properly extended to cope with polarisations, the formula in Equation (29) can still be applied, although it is advisable to consider the Hartree–Fock and correlation energies separately, noting that the Hartree–Fock energy converges in a different manner.

Many other extrapolation formulas have appeared over the years, some of which are reviewed in Section 6, in our discussion of composite methods. For new developments and further references to the literature, see the recent work of Bakowies [38,39].

4. Explicit correlation

The fact that dynamical correlation presents a special problem to quantum chemistry has been known since the late 1920s. In 1928, Hylleraas attempted to apply what we would call CI theory to the helium atom and was unable to reproduce the experimental ionisation potential. Hylleraas was thus one of the first to encounter the slow convergence of the orbital expansion with the size of the orbital basis. Convinced that higher accuracy was possible, he introduced arbitrary powers of r_{12} into the wave-function form and, with only a six-term expansion, obtained agreement with experiment to within 0.01 eV, which he knew was the order of magnitude of the contributions from relativity and nuclear motion. This was the very first explicitly correlated calculation. Indeed, this was also one of the first confirmations that quantum-mechanical predictions are quantitatively correct for chemical properties.

With the introduction of automated computation, the interest in methods general to many-electron systems grew and most researchers directed their efforts towards the CI expansion, based on an orbital description and Slater determinants. The overriding advantage of these methods is the reduction of n -electron Hamiltonian matrix elements to one- and two-electron integrals, as stated in the Slater–Condon rules. This compares very favourably to the daunting many-electron integration required for generalisations of Hylleraas' approach. However, as more experience

was gained, it became evident that high accuracy could not easily be attained in this manner. Schwartz published his analysis of the rate of convergence of a partial wave CI expansion for helium in 1962, showing it to be extremely slow – see Section 3.1. It was clear that the standard brute-force approach could never by itself become a practical quantitative theory – at least not for any non-trivial molecule with more than two atoms and a few electrons. Nevertheless, some early successes using pseudo-natural orbitals were published in this journal [40].

Dissatisfaction with the level of accuracy obtainable with orbital-based approaches has led to renewed interest in methods that include an explicit r_{12} dependence into the wave function. Any such explicitly correlated method must overcome the formidable integration over n coupled electronic degrees of freedom, required for the evaluation of Hamiltonian expectation values. Even if the n -particle basis functions are restricted to correlate explicitly only one electron pair at a time, numerous three- and four-electron integrals must be evaluated. Several strategies have been proposed:

- Use physically motivated basis functions and rely on mathematical ingenuity to perform the integration:
 - generalisations of the Hylleraas method.
- Use functions with simple integral formulae and introduce a high degree of flexibility:
 - exponentially correlated Gaussian (ECGs) and strong-orthogonality Gaussian-type geminal (SO-GTGs) methods.
- Reformulate the equations to require low-dimensional (three-electron) integration:
 - trans-correlated and weak-orthogonality GTG (WO-GTG) methods.
- Employ efficient approximate methods to evaluate the high-dimensional integrals:
 - R12 methods and variational Monte Carlo methods.

In the following, we give an historical account of the development of explicitly correlated theories. We next review the main aspects of the approaches listed above, indicating the successes and limitations of each method with respect to quantitative chemical predictions and highlighting recent developments. The variational quantum Monte Carlo method is used almost exclusively as a precursor to the diffusion quantum Monte Carlo method and the stochastic approach is of

a rather different nature to the rest of the approaches. Although we mention it here for completeness, we will not discuss it further in this article.

In this section on explicitly correlated approaches, we discuss accuracy primarily with respect to the basis-set limit. For variational calculations, this is the FCI basis-set limit, and for MP2 and CCSD methods, we refer to errors from their respective basis-set limits. For quantitative agreement with experiment, corrections beyond the Born–Oppenheimer and nonrelativistic approximations are often also required. These will be discussed in detail in Section 5. The following discussion of the explicitly correlated methods contains sufficient details for an appreciation of the key components of the various methods, but for a detailed description we refer the reader to the monograph *Explicitly Correlated Wave Functions in Chemistry and Physics in Progress in Theoretical Chemistry and Physics*, edited by Jacek Rychlewski (2002) [41]. The methods described therein are up to date, with the exception of the recent progress in R12 and GGn methods.

4.1. History

Explicitly correlated methods have been and continue to be at the forefront of truly quantitative computational treatments of chemical problems. The historical account given here is intended to give the reader a flavour of the difficulties that have faced researchers in developing these high-accuracy methods into practical tools for routine use and to put the sub-sequent more detailed description of these methods, their successes and their limitations, into context.

4.1.1. Early history

In the 1930s, shortly after the success of the Hylleraas approach for helium, James and Coolidge used a similar functional form including all powers of r_{12} for calculations on the ground state of molecular hydrogen, achieving millihartree accuracy [42]. Elliptic coordinates were used and the difficult Coulomb integrals were evaluated by means of a Neumann expansion. They observed a rapid convergence with the number of terms in the wave function and together with Present, similar success was met for excited states [43–45]. The James–Coolidge wave function does not have the correct asymptotic form to describe properly dissociation and, in 1965, Kołos and Wolniewicz (KW) generalised their basis to rectify this deficiency [46]. It is possible to generalise the Hylleraas approach to systems with more electrons or more nuclear centres, but the expense and complexity of the many-electron

integration required has severely limited the range of application, although some success was met by Clary and Handy [47,48].

In 1960, exponentially correlated Gaussians (ECGs) were introduced to molecular physics by Boys and Singer independently in their landmark papers in the *Proceedings of the Royal Society* [49–51]. The formulae required for integral evaluation using Gaussian orbitals as basis functions for molecular wave functions had been derived previously by Boys in 1950 [52] and these functions were shown to be highly suitable for automated computation. In his paper in 1960, Boys generalised his analysis to include explicit $\exp(\alpha_{ij}r_{ij}^2)$ dependence for all electron pairs ij . Lester and Krauss applied Gaussian geminals in calculations on two-electron systems such as He, H₂ and H₃⁺ [53,54]. They examined the Coulomb hole in these systems and observed that Gaussian geminals are well suited to describe its overall shape despite the fact that a Gaussian function has no cusp at the point of coalescence. Similar variational two-electron calculations were reported a few years later by Handy in this journal [55] and, in 1975, Karunakaran and Christoffersen used ECGs in a variational calculation on LiH, evaluating four-electron integrals [56]. In the 1960s and early 1970s, the consensus of opinion was that the convergence of ECGs was much slower than that of functions containing powers of r_{12} and it was not until later, when non-linear optimisation techniques were sufficiently advanced, that the power of ECGs was fully appreciated [57].

A different approach, based on a similarity-transformation of the Hamiltonian was proposed by Hirschfelder in 1963 [58]. In his approach, the wave function takes the form $\psi = C\phi$ where C depends on the interelectronic distances. He showed that ϕ may be determined using a transformed Hamiltonian $H' = C^{-1}\hat{H}C$ where, if C is properly chosen, H' is free from Coulomb singularities, thereby simplifying the construction of ϕ . This method was pursued by Jankowski in [59,60]. Boys and Handy developed the ideas further in a series of papers in 1969, where they established a practical method for the determination of parameters in both C and ϕ , which required the evaluation of at most three-electron integrals [61–64]. The trans-correlated method of Boys and Handy was further refined by Handy, who reduced the problems associated with the lack of a variational upper bound to the energy for the similarity-transformed Hamiltonian. Handy published his work in this journal, with applications to the ground-state energies of He, H₂, LiH and H₂O [65,66].

In the early 1960s, Sinanoglu presented his pair theories for treating many-body correlation effects [67].

The formal completeness of a Gaussian-geminal basis set for pair functions in many-electron wave functions was proved by King in 1967 [68]. Subsequently, Pan and King promoted the use of Gaussian geminals to compute Sinanoglu's first-order pair functions – that is, to compute the MP2 correlation energy [69,70]. They obtained remarkably good results for their calculations on He, Be, B⁺ and Ne with a relatively small number of geminal basis functions and expressed optimism that suitable initial exponents for the geminals could be generated from Gaussian-type orbitals with correlation factors. Adamowicz and Sadlej were probably the first to compute the MP2 energy for molecules such as H₂, LiH and BH using Gaussian-type geminals (GTGs) [71,72].

4.1.2. Modern history

To overcome the bottleneck of the four-electron integral evaluation, required in GTG theory for the second-order pair energies, Szalewicz *et al.* introduced in 1982 the weak-orthogonality (WO) functional [73]. Here the formal condition that the geminal basis be strongly orthogonal to the set of occupied Hartree–Fock orbitals is relaxed and a penalty function is inserted to ensure that the computed energy is an upper bound to the true MP2 value, for a given set of Hartree–Fock orbitals. Application of the variational principle then produces strongly orthogonal pair functions in the basis-set limit. This breakthrough enabled a much wider range of application of the GTG method, which had the stated goal of achieving chemically accurate correlation energies for small molecules. In a subsequent series of papers, stable and efficient procedures for computing MP2, MP3 and approximate CCD energies were established that require evaluation of at most three-electron integrals [74–78]. Thereafter, research focused on reducing the number of non-linear parameters through random tempering [79–81], before turning in the late 1990s to MP2 and MP3 first-order properties [82] and the inclusion of single excitations in the CCSD method [83]. Applications of the GTG method have focused mainly on four-electron systems, although calculations on neon [78] and water [84] have also been reported. This method is still actively pursued today and has recently been used to compute the CCSD contribution to the helium-dimer interaction energy to an accuracy of 0.003 K [85,86].

In 1985, a new explicitly correlated approach was proposed by Kutzelnigg: the R12 approach [87]. Rather than expanding the first-order pair functions in a basis of Gaussian geminals, Kutzelnigg proposed to use an orbital basis set and to augment it with

geminals that depend on a single, universal, linear r_{12} term. This term was to describe the electronic cusp that is poorly represented in conventional orbital expansions. Kutzelnigg avoided direct evaluation of the numerous three- and four-electron integrals by an approximate resolution of the identity, reducing integral evaluation to at most two-electron integrals.

The first MP2-R12 calculations on molecules were reported by Klopper and Kutzelnigg in 1987 [88] and by 1993 calculations on larger systems were feasible. Applications included MP2 energies of Be₄ and Mn₄ [89], benzene–Ne and benzene–Ar complexes [90] and ferrocene [91]. The first extension of the R12 approach to CI wave functions was made in 1991 [92], and in 1992 Noga *et al.* published pilot CCSD[T]-R12 calculations on Be₂ [93]. In 1993, Gdanitz combined the R12 approach with multireference CI calculations [94]. An integral-direct coupled-cluster implementation was completed in 1995 [95], with an open-shell extension in 2000 [96]. Many benchmark applications for small molecules followed, usually in combination with higher-order corrections beyond that of CCSD(T) theory. These included Born–Oppenheimer (BO) conformational barriers in CH₅⁺ and H₅⁺ to an accuracy of 0.02 kcal mol⁻¹ [97,98], static polarisabilities of Be, LiH, Ne and HF (better than 0.05 a.u.) [99,100], atomisation energies (1–2 kJ mol⁻¹) [101], geometries (0.01 pm) [102] and frequencies (1–3 cm⁻¹) [103] of small molecules of first-row atoms, heats of formation of C_xH_y up to C₃H₄ [104], as well as basis-set convergence studies to develop extrapolation methods [35].

In the pursuit of spectroscopic accuracy for molecular energies, Rychlewski and co-workers returned to the highly flexible ECGs of Boys and Singer. In 1993, Cencek and Rychlewski published a general method for the variational calculation of many-electron systems using a basis of floating spherical ECGs, using Powell's conjugate-direction method for the challenging non-linear optimisation [105]. Together with Komasa, they subsequently produced a series of highly accurate calculations on small molecular systems, obtaining nanohartree accuracy for potential energy curves of two-electron systems such as H₃⁺ and HeH⁺ [106] and sub-microhartree accuracy for three- and four-electron systems such as He₂⁺, He₂, LiH and Be like ions [107–110]. Similar success was met for excited states [106,111] and, with a focus on comparison with experiment, highly accurate nonadiabatic, relativistic and quantum-electrodynamics (QED) corrections have also been reported [112–114]. The ECG method continues to be extremely successful in treating small systems and remains an area of active development.

In 1996, Persson and Taylor combined features of the R12 and GTG approaches to produce a new Gaussian-geminal method, applicable to larger molecules [115]. They adopted the philosophy of Kutzelnigg and retained the orbital expansion, but augmented it with a basis of Gaussians with fixed exponents, demonstrating that their basis was able to describe accurately the correlation hole and the linear r_{12} behaviour at the cusp. In this way, they avoided the non-linear optimisation of the GTG method and also the approximate many-electron integral evaluation necessary in the R12 method, together with the requirement that the orbital basis set be sufficiently large to resolve the identity. To avoid four-electron integral evaluation, they used the WO functional. This GTG model was later generalized to the full GTO–GTG (GGn) model in [116,117]. In 2006, this method was combined with localised orbitals by Polly and co-workers, who published MP2 calculations on molecules including C_2H_4 , CH_2O and N_2H_4 [118]. A comprehensive study of the GGn method was recently presented by Dahle *et al.* [117], with a further application to the water molecule in [119].

The trans-correlated method has received a moderate level of renewed interest over the past decade. In 1998, Nooijen and Bartlett suggested using a frozen similarity transformation instead of optimising the factor C [120]. Ten-no adopted this approach in his examination of the trans-correlated method in 2000, giving efficient formulae for the three-electron integration [121,122]. However, he found that a more flexible one-electron basis is required in comparison with other explicitly correlated methods. The idea of trans-correlated methods was also employed by Imamura and Scuseria in 2003 to develop a new correlation functional for DFT [123]. It has recently been pursued within the framework of the variational Monte Carlo method by Umezawa and Chikyow [124].

Over the past six years the development within R12 theory has greatly accelerated. In 2002, Klopper and Samson introduced an auxiliary basis set for the resolution of the identity (RI), making it possible to use the R12 method with small orbital basis sets [125]. Up until this time, the orbital basis set had been used for the resolution of the identity and large sets were necessary. Subsequently, Manby introduced density fitting into R12 theory, greatly reducing the computation time for evaluating the various kinds of two-electron integrals [126]. In 2004, Ten-no suggested the use of Slater-type geminals in place of the linear r_{12} dependence, and to evaluate the necessary integrals either by numerical quadrature or by representing the

exponential with a linear combination of Gaussians [127]. The performance of the Slater-type geminals is markedly superior to that of linear r_{12} and this approach has now been adopted by the groups of Manby [128], Klopper [129], Valeev [130] and Werner [131]. These advances for MP2 theory were followed closely by extensions to the CC2 [132,133], CCSD [134,135] and CCSD(T) [136,137] levels of theory within the groups of Klopper and Hättig. In combination with the Slater-type geminal, reaction energies were computed to chemical accuracy using only triple-zeta orbital basis sets [137,138]. Simplified CCSD(T)-F12 methods with a computational cost only slightly higher than a conventional CCSD(T) calculation have very recently been proposed by Valeev [139] and Adler *et al.* [140]. These methods also return similarly high-quality results and open the way to obtaining chemical accuracy for medium-size systems.

4.2. Hylleraas wave function

The original expansion for the helium wave function, as used by Hylleraas for his calculations in 1929, is of the form

$$\sum_i c_i s^i t^m u^n \exp(-\zeta s) \Theta(\sigma_1, \sigma_2), \quad (31)$$

where $\Theta(\sigma_1, \sigma_2)$ is a singlet or triplet spin function and $s = r_1 + r_2$, $t = r_2 - r_1$ and $u = r_{12}$. The wave function is thus constructed from an orbital component combined with all powers of the distances r_1 , r_2 and r_{12} , which efficiently describe the radial and angular correlation and have the correct form to satisfy the two-particle coalescence conditions. The Hamiltonian matrix elements simply involve factorials and the coefficients c_i and exponent ζ may be determined by variationally minimising the energy. In the pursuit of ever more digits, researchers have used modified expansions, including perimetric coordinates [141], negative [142] and half-integer [143] powers and an $\ln(s)$ dependence [144,145]. It is not the purpose of this review to survey these methods, since we are primarily concerned with molecular physics, but the interested reader will find a useful historical table in the most recent work of Nakashima and Nakatsuji [146], who obtain 40-digit accuracy. The reader should keep in mind, however, that this accuracy is with respect to the FCI basis-set limit and not with respect to experiment. We note that the most successful expansions are those that contain a logarithmic dependence on s , which is required to satisfy the three-particle coalescence point at the nucleus [147].

The Hylleraas approach has been extended to the hydrogen molecule [42,46], using elliptic coordinates ξ_i , η_i and φ_i in the KW expansion

$$\sum_i c_i \xi_1^{j_i} \xi_2^{k_i} \eta_1^{l_i} \eta_2^{m_i} r_{12}^{n_i} \exp(-\alpha_1 \eta_1 - \alpha_2 \eta_2) \times \cosh(\beta_1 \eta_1 - \beta_2 \eta_2) \Theta(\sigma_1, \sigma_2) \quad (32)$$

and to the lithium atom [148], using the basis

$$\sum_i c_i r_1^{j_i} r_2^{k_i} r_3^{l_i} r_{12}^{m_i} r_{13}^{n_i} r_{23}^{o_i} \exp(-\alpha r_1 - \beta r_2 - \gamma r_3) \Theta_i(\sigma_1, \sigma_2, \sigma_3), \quad (33)$$

and total energies of ground and excited states, accurate to nE_h , have been computed [149,150]. In 1977, Clary proposed a generalisation of this approach for application to diatomic molecules, using nonorthogonal orbitals in elliptic coordinates [47]

$$\varphi_i^k(\mathbf{r}_i) = \xi_i^l \eta_i^m \exp(-\alpha_k \xi_i + \beta_k \eta_i). \quad (34)$$

He wrote the wave function as

$$\Psi(\mathbf{r}, \boldsymbol{\sigma}) = \hat{A}[\Phi(\mathbf{r})\Theta_{S,M_S}(\boldsymbol{\sigma})], \quad (35)$$

where \hat{A} is the antisymmetrising operator, $\Theta_{S,M_S}(\boldsymbol{\sigma})$ is a proper spin eigenfunction and

$$\Phi(\mathbf{r}) = \sum_k c_k r_{ij}^{v_k} \prod_{i=1}^n \varphi_i^k(\mathbf{r}_i). \quad (36)$$

Clary gave explicit formulae for the one- to four-electron integrals required for the secular equations, but the storage requirements and complexity of the three- and four-electron integrals limited the number of expansion terms and the accuracy obtained [47,48]. A simpler generalisation is given by the Hyl-CI approach, where the configuration state functions of the conventional orbital-based CI expansion are combined with a single r_{ij} or $\exp(-\gamma r_{ij})$ correlation factor. This approach has been used successfully for two-electron systems, but the complexity of the numerous three- and four-electron integrals again limited the applicability of this approach to systems with more than two electrons [151,152].

4.3. Exponentially correlated Gaussians (ECG)

In the ECG method of Rychlewski and co-workers [105], the wave function is an antisymmetrised product of a spatial function and a proper spin eigenfunction

$$\Psi(\mathbf{r}, \boldsymbol{\sigma}) = \hat{A}[\hat{S}\Phi(\mathbf{r})\Theta_{S,M_S}(\boldsymbol{\sigma})], \quad (37)$$

where the operator \hat{S} is a spatial symmetry operator that ensures that $\Phi(\mathbf{r})$ transforms according to the

point group of the molecule. The spatial function is expanded in a basis of spherical ECGs

$$\Phi(\mathbf{r}) = \sum_l c_l \exp[(\mathbf{r} - \mathbf{s}_l)^T \mathbf{A}_l (\mathbf{r} - \mathbf{s}_l)], \quad (38)$$

$$(\mathbf{r} - \mathbf{s}_l)^T \mathbf{A}_l (\mathbf{r} - \mathbf{s}_l) = -\sum_i \alpha_i^l (\mathbf{r}_i - \mathbf{C}_i)^2 - \sum_{i < j} \beta_{ij}^l (\mathbf{r}_i - \mathbf{r}_j)^2. \quad (39)$$

The linear parameters, c_l and non-linear parameters \mathbf{A}_l , \mathbf{s}_l are determined by variationally minimising the Rayleigh–Ritz quotient for the energy. The motivation for this choice of expansion is twofold. First, Boys [49] and Singer [50] have shown that all of the necessary Hamiltonian matrix elements are easily evaluated and are not much more complicated than two-electron GTO Coulomb integrals. Second, the ECGs are highly flexible and, with careful optimisation of the non-linear parameters, the rate of convergence with the number of ECGs is as fast as that of the Hylleraas approach [57]. This is illustrated in Figure 3 for molecular hydrogen, where we plot the error of the BO nonrelativistic energy against the number of terms in the ECG and KW wave-function expansions [41]. The success of the ECG method depends crucially on the ability to optimise a large number of non-linear parameters. The optimisation must be reliable and efficient since repeated re-evaluation of the n -electron integrals is expensive. The method of choice is an iterative use of Powell’s conjugate derivative method, where subsets of $n(n+7)/2$ parameters belonging to individual ECGs

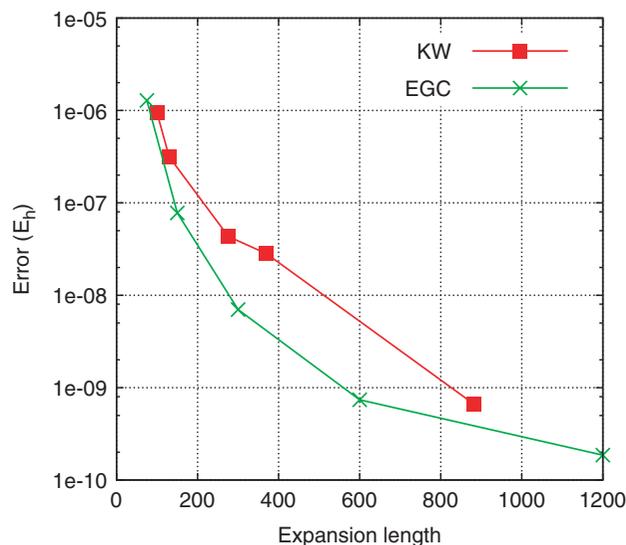


Figure 3. Errors in the total energy of molecular hydrogen against the number of basis functions for KW ■ and ECG × expansions.

are optimised, keeping all others fixed. It is then only necessary to recompute the matrix elements in the rows and columns effected by the modified basis functions. For S and Σ states, it is both practical and sufficient to employ purely spherical ECGs. For excited states, this restriction can lead to a loss of rotational invariance and the spherical ECGs should be augmented with Cartesian functions representing the angular momentum of the state being computed.

The ECG method has been applied, with much success, to compute total energies of ground and excited states of three- and four-electron atoms and molecules, as well as several properties. For example, a 3700-term ECG expansion has been used to compute an upper bound to the ground-state energy of Be with an error of less than 1 cm^{-1} , which is the most accurate estimate currently available [153]. An ECG basis of over 4000 terms was used to compute the electron affinity of Li to within 0.06 cm^{-1} of experiment, which required accurate evaluations of nonadiabatic, relativistic and QED corrections [114]. For properties such as these, which depend on the interparticle coalescence points, the convergence with the size of ECG basis is rather slow. However, accurate values may be obtained by using elegant expectation-value identities to convert the singular operators into linear combinations of operators with weaker singularities. The largest molecule treated with ECGs is LiH [109], where a 2400-term expansion has been used to obtain an accuracy of about $15 \mu E_h$. The helium dimer has also been the subject of repeated investigation and the ECG method provides the best available potential curve for this challenging system [154]. No applications to molecules with more than four electrons have yet been reported, the principal problem being in the large computational expense of the repeated integral evaluation.

4.4. Explicitly correlated MP2 theories

The high computational demand of variational methods such as the Hylleraas and ECG approaches has so far prevented these methods from evolving into useful tools for general chemical applications. The combination of explicitly correlated techniques with Sinanoglu's pair theories, and in particular with coupled-cluster wave functions, is one way to extend the range of applicability. Although a quantitative treatment of chemical properties often requires a correlation treatment beyond that of second-order perturbation theory, MP2 theory is the natural first step in the development of new explicitly correlated pair theories and serves as an important testing ground.

Furthermore, highly accurate MP2 energies may be combined with coupled-cluster energies in various focal-point additivity schemes – see Section 6. With these considerations in mind, we review the explicitly correlated MP2 theories prevalent in the current literature.

The equation for the Møller–Plesset first-order correction to the Hartree–Fock wave function is well known

$$(\hat{H}^0 - E^0)\psi^1 + (\hat{V} - E^1)\psi^0 = 0. \quad (40)$$

In 1964, Sinanoglu proved that this many-body equation is separable into pair equations [67]:

$$(\hat{f}_1 + \hat{f}_2 - \varepsilon_i - \varepsilon_j)u_{ij}^s(1, 2) + \hat{Q}_{12}r_{12}^{-1}\phi_{ij}^s(1, 2) = 0. \quad (41)$$

We have here used the spin-free formulation, where the spatial functions $u_{ij}^s(1, 2)$ and $\phi_{ij}^s(1, 2)$ are to be combined with singlet ($s=0$) or triplet ($s=1$) spin factors. As usual, $i, j \dots$ refer to occupied orbitals, $a, b \dots$ to virtual and $p, q \dots$ to arbitrary orbitals. The spin-free Hartree–Fock pair functions are

$$\phi_{ij}^s(1, 2) = [\phi_i(1)\phi_j(2) + (1 - 2s)\phi_j(1)\phi_i(2)](2 + 2\delta_{ij})^{-1/2} \quad (42)$$

and ϕ_i are canonical orbitals, $\hat{f}_i\phi_i = \varepsilon_i\phi_i$. \hat{Q}_{12} is the strong orthogonality (SO) projector:

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2), \quad (43)$$

$$\hat{O}_1 = \sum_i |\phi_i\rangle\langle\phi_i|. \quad (44)$$

Sinanoglu's first-order pair functions $u_{ij}^s(1, 2)$ are thus defined to be strongly orthogonal to the Hartree–Fock orbitals:

$$\int u_{ij}^s(1, 2)\phi_k(1) d1 = 0 \quad \forall k = 1, n. \quad (45)$$

This condition is automatically satisfied when u_{ij}^s is expanded in a basis of virtual orbitals, but if geminal functions with explicit r_{12} dependence are used, an SO projector must be applied, $u_{ij}^s = \hat{Q}_{12}w_{ij}^s(r_{12})$. The MP2 pair energies may be computed directly from the unprojected functions w_{ij}^s via the Hylleraas functional, which is a Lagrangian formulation of the MP2 pair energies [155],

$$\begin{aligned} {}^{\text{SO}}F[w_{ij}^s] &= \langle w_{ij}^s | \hat{Q}_{12}(\hat{f}_1 + \hat{f}_2 - \varepsilon_i - \varepsilon_j)\hat{Q}_{12} | w_{ij}^s \rangle \\ &+ 2\langle w_{ij}^s | \hat{Q}_{12}r_{12}^{-1} | \phi_{ij}^s \rangle. \end{aligned} \quad (46)$$

Minimisation of ${}^{\text{SO}}F[w_{ij}^s]$ with respect to the variational parameters in w_{ij}^s gives an upper bound to each MP2 pair energy, for a given Hartree–Fock reference determinant. The explicitly correlated methods developed to compute the MP2 energy are characterised both by the basis functions used to expand w_{ij}^s and by the strategy for handling the three- and four-electron integrals which arise in Equation (46) due to the presence of the SO projector. The three dominant methods in use today are the GTG method, the GGn method and the R12 method.

4.4.1. The GTG method

Following the initial ideas of Pan and King, the approach adopted in the GTG method is to expand the first-order pair functions in a basis of highly flexible Gaussian geminals. The general form of the expansion is

$$w_{ij}^s = \sum_{vk} c_{vk}^{ij} G_{vk} \exp\left[-\alpha_v(\mathbf{r}_1 - \mathbf{P}_v)^2 - \beta_v(\mathbf{r}_2 - \mathbf{Q}_v)^2 - \gamma_v r_{12}^2\right], \quad (47)$$

where the set of exponents α_v , β_v and γ_v and centres \mathbf{P}_v and \mathbf{Q}_v are to be optimised for each pair function together with the linear coefficients c_{vk}^{ij} . The function G_{vk} is a polynomial

$$G_{vk} = x_{1P_v}^{k_{1x}} y_{1P_v}^{k_{1y}} z_{1P_v}^{k_{1z}} x_{2Q_v}^{k_{2x}} y_{2Q_v}^{k_{2y}} z_{2Q_v}^{k_{2z}}, \quad (48)$$

with $x_{1P_v} = x_1 - P_{vx}$. The set of non-negative integers \mathbf{k} is usually determined by symmetry and several completeness theorems now exist proving that in many cases only certain combinations of low powers are necessary [68,84,156]. For example, for pairs with an angular quantum number L , it is only necessary to include $2L + 1$ particular sets of \mathbf{k} in the geminal basis. The matrix elements in Equation (46) are evaluated analytically using the Boys function. However, the enormous number of four-electron integrals that arise restricts practical applications to small systems or short geminal expansions. It should be noted that non-linear optimisation of the exponents and centres requires repeated re-evaluation of the matrix elements. In all modern applications, the weak-orthogonality (WO) functional is used, where at most three-electron integrals occur:

$${}^{\text{WO}}F[w_{ij}^s] = \left\langle w_{ij}^s \left| \left(\hat{f}_1 + \hat{f}_2 - \varepsilon_i - \varepsilon_j + \hat{W}_{ij} \right) \right| w_{ij}^s \right\rangle + 2 \left\langle w_{ij}^s \left| \hat{Q}_{12} r_{12}^{-1} \right| \phi_{ij}^s \right\rangle. \quad (49)$$

The SO projector has been removed from the Fock matrix elements and a penalty function has been

added, which vanishes if w_{ij}^s is strongly orthogonal to the occupied space. The penalty operator \hat{W}_{ij} is constructed such that the SO violating terms that enter the Fock matrix elements never lead to an artificial overestimation of the correlation energy and the WO functional is a strict upper bound to the MP2 energy, for a given set of Hartree–Fock orbitals. The original form for \hat{W}_{ij} was [73]

$$\hat{W}_{ij} = \left[\frac{1}{2}(\varepsilon_i + \varepsilon_j - 2\varepsilon_1) + \eta \right] (\hat{O}_1 + \hat{O}_2), \quad (50)$$

where ε_1 is the lowest orbital eigenvalue and η is a positive number. However, this penalty was found to be more severe than necessary and gave an unbalanced treatment of core and valence electron pairs. The modified penalty operator is therefore recommended [78]:

$$\hat{W}_{ij} = [A_{ij} + \varepsilon_{\text{HO}} + \eta] (\hat{O}_1 + \hat{O}_2) - \hat{O}_1^\varepsilon - \hat{O}_2^\varepsilon - A_{ij} \hat{O}_1 \hat{O}_2, \quad (51)$$

$$A_{ij} = \varepsilon_i + \varepsilon_j - 2\varepsilon_{\text{HO}}, \quad (52)$$

$$\hat{O}_i^\varepsilon = \sum_k \varepsilon_k |\phi_k\rangle \langle \phi_k|, \quad (53)$$

where ε_{HO} is the highest occupied orbital eigenvalue. Through variational minimisation of the WO functional, the highly flexible Gaussian geminal expansion adapts to produce accurate pair functions w_{ij}^s , which, to a good approximation, satisfy the SO constraint, and very accurate pair energies are obtained. The requirement that the geminal expansion simultaneously describes correlation and satisfies SO has frustrated attempts to find recurrent patterns in the optimised exponents and centres. Consequently, the success of the GTG method relies on the ability to optimise a large set of non-linear parameters and sophisticated techniques such as conjugate derivatives and random tempering are necessary. For MP2 energies of four-electron systems such as Be and LiH [83], the unrivalled accuracy of the GTG method is to within a couple of microhartrees. First-order properties such as MP2 dipole and quadrupole moments of LiH are converged to 0.001 and 0.01 a.u., respectively [82]. However, for ten-electron systems such as neon and water, it is has so far only been possible to use much shorter GTG expansions and accuracy for the MP2 correlation energy is limited to a few mE_h [78,84].

4.4.2. The GGn method

The GGn method is a hybrid method, which combines aspects of the GTG and the R12 approaches [117]. In particular, the non-linear optimisation of the GTG

method is avoided, enabling a treatment of more general systems. In the GGn method, the many-electron integration is performed analytically and the WO functional is used to avoid four-electron integrals. The form of the pair-function expansion is similar to R12 theory and a dual basis is used, comprised of the conventional set of virtual orbitals and a set of GTGs constructed from orbital pairs with an attached correlation factor. The hierarchy of methods GG0, GG1, GG2, respectively, are defined by the set of orbital pairs used in the geminal basis [117,118]:

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s + \sum_{k \leq l} \sum_v c_{ij,s}^{kl,v} \exp(-\gamma_v r_{12}^2) \phi_{kl}^s, \quad (54)$$

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s + \sum_{k \leq q} \sum_v c_{ij,s}^{kq,v} \exp(-\gamma_v r_{12}^2) \phi_{kq}^s, \quad (55)$$

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s + \sum_{p \leq q} \sum_v c_{ij,s}^{pq,v} \exp(-\gamma_v r_{12}^2) \phi_{pq}^s. \quad (56)$$

These expansions may equally be viewed as a specific choice of centres and exponents in the GTG expansion of Equation (47): the centres \mathbf{P}_v and \mathbf{Q}_v are fixed at the nuclei and the exponents α_v and β_v are those of a standard Gaussian orbital basis set. The set of exponents γ_v are usually taken to be an even-tempered series of six or nine terms. Since the exponents and centres are fixed, the convergence is much slower with respect to basis size than for the original GTG method, and polynomials representing much larger angular-momentum quantum numbers occur. However, because the conventional virtual orbital pairs and standard basis sets are employed, one ensures that the expansion is capable of yielding a good overall description of w_{ij}^s and the additional GTGs act to improve the correlation description. In particular, they are capable of describing the short-range r_{12} dependence of the wave function near the cusp.

The GG2 method has been used to compute the MP2 pair energies of small systems such as Ne and HF to within 0.1 mE_h of the basis-set limit, requiring only triple-zeta orbital basis sets [117]. MP2-GG1 calculations have also been performed on the series of molecules CO, CH₄, C₂H₂, C₂H₄, H₂CO and N₂H₄ using double-zeta orbital basis sets and between 93 and 96% of the MP2 correlation energy was recovered [118]. This corresponds to an error of 10 mE_h in total correlation energies and less than 4 kJ mol⁻¹ in the correlation contribution to reaction energies and activation barriers for these molecules. In [117,119], Dahle *et al.* presented the most accurate all-electron MP2 correlation energies so far published for Ne, H₂, HF, and H₂O.

Unfortunately, the size of system the GG1 method and, in particular, the GG2 method can treat is limited by the memory requirements of the computation. The performance of the cheaper GG0 pair function was disappointing, being inferior to the R12 approach even though the flexibility in the GG0 basis is greater than that of the R12 basis. This anomaly arises since the GG0 basis has insufficient flexibility to form strongly orthogonal pair functions, which is required when the WO functional is used [116,157]. Tew *et al.* proposed a new WO functional that makes the GG0 method more competitive. Their method is equivalent to replacing w_{ij}^s with $\hat{P}_{12} w_{ij}^s$ in the WO functional, where

$$\hat{P}_{12} = 1 - \hat{O}_1 \hat{O}_2 - \hat{O}_1 \hat{V}_2 - \hat{V}_1 \hat{O}_2, \quad (57)$$

$$\hat{V} = \sum_a |\phi_a\rangle \langle \phi_a|. \quad (58)$$

The key feature that only three-electron integrals arise is retained and the flexibility required for a strongly orthogonal w_{ij}^s is greatly reduced. Use of this new WO functional with the GG0 method will permit larger systems to be treated, but applicability will still be limited unless the RI approximation is invoked to avoid the explicit calculation of three-electron integrals.

4.4.3. The R12 method

The original R12 ansatz for the MP2 pair function, as proposed by Kutzelnigg, is given by

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s + \sum_{k \leq l} c_{ij,s}^{kl} \hat{V}_{12} r_{12} \phi_{kl}^s, \quad (59)$$

where the projector $\hat{V}_{12} = 1 - \hat{V}_1 \hat{V}_2$ serves only to decouple the two contributions. The choice of r_{12} dependence was motivated by the fact that, if the coefficients $c_{ij,s}^{kl} = \delta_{ij,kl} / (2s + 2)$, then the R12 geminal functions exactly satisfy the cusp condition for Sinanoglu's pair equations. The final details of the correlation hole, which cannot be captured by the orbital basis, are therefore very well described by the R12 geminals and the slow convergence of the CI expansion is removed. In the R12 method, the coefficients $C_{ij,s}^{ab}$ and $c_{ij,s}^{kl}$ are determined by variational minimisation of the SO functional and the three- and four-electron integrals that arise are reduced to sums of products of two-electron integrals by successive insertions of an approximate resolution of the identity (RI).

In all modern implementations, the identity is resolved using the union of the orbital basis and a dedicated complementary auxiliary basis (CABS) [158] and the error associated with the approximate RI

can be reduced to the order of $0.01 mE_h$ in the correlation energy for systems the size of Ne [159]. Listed in increasing accuracy, and computational cost, A, A', A'', C and B, are a series of approximations for evaluating the matrix elements over $(\hat{f}_1 + \hat{f}_2 - \varepsilon_i - \varepsilon_j)$ in the SO functional. B and C are the most complete, the only neglected terms are those associated with Fock matrix elements connecting occupied and CABS orbitals, which would be zero if the generalised Brillouin condition held. The error incurred by neglecting these terms is of the order of $0.5 mE_h$ when triple-zeta basis sets are used, but this error has been observed to always be smaller than the remaining basis-set incompleteness error and is of little concern [128]. Detailed discussions of the relative accuracy of the various approximations may be found in [125,131]. Approximation C, introduced by Noga and co-workers [160] is recommended, since only five kinds of two-electron integrals are then required.

Although the linear r_{12} term works well when large orbital basis sets are used, the R12 geminals are not optimal for use with double-zeta or triple-zeta sets. If the linear r_{12} correlation factor is replaced with an exponential function $\exp(-\gamma r_{12})$, as proposed by Ten-no, the performance for small basis sets is greatly improved and more than 98% of the MP2 correlation energy is recovered with a double-zeta basis set [127]. This corresponds to an error of $4 mE_h$ for water and less than 2 kJ mol^{-1} in the MP2 reaction energies of small molecules. For the triple-zeta basis the error is further reduced by more than a factor of two. A similar choice of correlation factor, known as a Jastrow factor, is used in quantum Monte Carlo methods, although typically the Jastrow factor depends on all interelectron distances and acts on n -electron basis functions. Valeev has shown that the choice of $\exp(-\gamma r_{12})$ as the correlation factor for the geminal basis is very close to the optimal factor [130] and this F12 approach has now replaced the original R12 method. The computed correlation energy is rather insensitive to the exponent γ in the F12 basis functions and a value close to $1.5 a_0^{-1}$ is a good choice for the valence electrons of light atoms [129].

Tew and Klopper observed that, for weak interactions, the geminal basis-set superposition error (BSSE) becomes problematic [161]. For example, when the orbitals i and j are on one fragment, the pair function for these orbitals uses geminals with $f(r_{12})\phi_{kl}^s$ with k and l localized on another fragment. Geminal BSSE can be removed if the pair function expansion is simplified to

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s + c_{ij,s} \hat{V}_{12} \exp(-\gamma r_{12}) \phi_{ij}^s. \quad (60)$$

This expansion is not invariant to rotations between occupied orbitals, but that problem is solved by constraining the $c_{ij,s}$ to be the same for every pair. Indeed, appropriate values for the coefficients may be predetermined from the cusp conditions as one half for singlet pairs and one quarter for triplet pairs. This is the rational generator method of Ten-no [162] and the pair function expansion then takes the simple form

$$w_{ij}^s = \sum_{a \leq b} C_{ij,s}^{ab} \phi_{ab}^s - \frac{1}{(2s+2)\gamma} \hat{V}_{12} \exp(-\gamma r_{12}) \phi_{ij}^s. \quad (61)$$

This is now the recommended approach and, as we shall see below, this correlation function may also be used in CCSD theory.

In Table 1, we present a comparison of the performance of the R12, F12 (rational generator), GGn and GTG methods for the MP2 correlation energy of water. The GTG value is poor because it was only possible to use a short expansion for the pair functions. For a given orbital basis, F12 and GG1 perform similarly and are superior to R12 and GG0. GG2 is the best of all. However, the computational cost of F12 is far smaller than that of GG1 and GG2 because the three-electron integrals are avoided. Moreover, the rational-generator F12 approach scales as n^5 with system size, compared to n^6 for GG2. In fact, if 98% of the MP2 correlation energy is desired, the F12 calculation using a double-zeta basis is less expensive than the conventional calculation using a quintuple-zeta basis. The error in the double-zeta Hartree-Fock energy is often larger than that of the MP2-F12 correlation energy and Noga has introduced a perturbative correction to the Hartree-Fock energy to reduce this error, using the CABS orbitals as an additional singles basis [165].

The MP2-F12 method has been combined with density fitting [126,166] and local approximations [167,168] and computations on large systems have become feasible. For example, the MP2 correlation energy of the 49 atom (Gly)₄ has been computed with an error of 1% using an aug-cc-pVTZ basis, around 1600 basis functions [168]. The MP2-R12 and MP2-F12 methods have also been extended to treat open-shell systems using unrestricted [96,130,169], restricted [170,171] and multi-reference [170] formalisms. Kordel *et al.* have implemented analytic first derivatives for the MP2-R12 method and have computed basis-set limit MP2 dipole and quadrupole moments of small molecules [172] and geometries of water dimers [173]. However, this was only the first step towards analytic derivatives for the MP2-F12 method with the CABS approach and the gradients for this method are likely to appear in the future.

Table 1. Water frozen-core and all-electron MP2 correlation energies at a fixed geometry with $\angle(\text{HOH}) = 104.52^\circ$ and $r(\text{OH}) = 95.720$ pm.

Authors	$E(\text{fc-MP2})/mE_h$	$E(\text{ae-MP2})/mE_h$
Dahle <i>et al.</i> [119]		
MP2-GG0 (aug-cc-pCVTZ)	-292.05	-351.36
MP2-GG1 (aug-cc-pCVTZ)	-299.82	-361.26
MP2-GG2 (aug-cc-pCVTZ(spdf, sp))	-300.44	-361.95
Yamaki <i>et al.</i> [163]		
MP2-F12/B (aug-cc-pVDZ)	-290.18	
(aug-cc-pVTZ)	-297.66	
(aug-cc-pVQZ)	-299.48	
(aug-cc-pV5Z)	-300.09	
(aug-cc-pV6Z)	-300.32	
Estimated limit	-300.4(3)	
Klopper [37]		
MP2-R12/B (19s14p8d6f4g3h/9s6p4d3f) ^b	-300.35	-361.92
Klopper [164]		
MP2-R12/A (15s9p7d5f3g1h/9s7p5d3f1g)	-300.53	-362.01
Bukowski <i>et al.</i> [84]		
Nonlinearly optimised GTGs	-295.04	-356.43

^aUsing a Slater-type geminal with exponent $1.5 a_0^{-1}$.

^bGeometry: $\angle(\text{HOH}) = 104.225^\circ$ and $r(\text{OH}) = 95.7119$ pm.

4.5. CCSD theory

As discussed in Sections 2 and 3, the coupled-cluster hierarchy provides a clear route towards quantitative quantum chemistry, but the slow basis-set convergence limits the accuracy obtainable even for small molecules, making meaningful calculations on large systems practically impossible. Much effort has therefore been spent on combining explicitly correlated techniques with the coupled-cluster ansatz to alleviate this problem of basis-set convergence. The central idea is to generalise the \hat{T}_2 operator in the cluster function to include excitations where a pair of occupied orbitals are replaced by an explicitly correlated geminal function. Coupled-cluster methods are usually discussed in the language of second quantisation, where occupation-number vectors refer to Slater determinants of orbital products. We may discuss explicitly correlated geminal functions $g_\nu(1, 2)$ in second quantisation if we introduce a formal, complete orbital basis, such that

$$g_\nu(1, 2) = \sum_{xy} r_{xy}^{\nu} \phi_x(1) \phi_y(2). \quad (62)$$

Indeed, we see that geminal functions are successful because they correspond to the particular linear combinations of orbital products that are required to reproduce the correlation hole. The complete orbital basis does not exist in practical calculations and each contraction over an index x in the resulting equations

increases the dimension of the required n -electron integration by one electron. As for MP2, the coupled-cluster formalism requires that the pair functions satisfy the SO constraint. This is equivalent to requiring that $r_{ij}^{\nu} = 0$ if x or y are occupied indices and may be enforced by using the projected geminals $\hat{Q}_{12} g_\nu(1, 2)$. The doubles excitation operator, where occupied orbital pairs ij are excited into geminal functions ν with amplitudes c_{ij}^{ν} , is

$$\hat{T}_2^\nu = \frac{1}{2} \sum_{ij\nu} c_{ij}^{\nu} \sum_{xy} r_{xy}^{\nu} E_{xi} E_{yj}, \quad (63)$$

where $E_{xi} = a_{x\alpha}^\dagger a_{i\alpha} + a_{x\beta}^\dagger a_{i\beta}$. The explicitly correlated coupled-cluster equations are derived by inserting \hat{T}_2^ν into the usual expressions for the coupled-cluster wave function and energy. The amplitudes c_{ij}^{ν} are determined in a manner analogous to conventional amplitudes and the similarity-transformed Schrödinger equation is projected onto the space of geminal double replacements, $\langle \mu_2 |$:

$$\sum_{xy} r_{xy}^{\nu} \langle \mu_2 | \hat{H} e^{\hat{T}} | \text{HF} \rangle = 0. \quad (64)$$

The presence of the SO projector in the geminal basis leads to three-, four- and five-electron integrals in this amplitude equation. As we have seen for the MP2 methods, explicit evaluation of many-electron integrals severely restricts the range of application and the successful approaches are those that involve two- or

three-electron integration at most. Of the MP2 methods reviewed in the previous section, only the GTG and R12 (F12) approaches have so far been generalised to the CCSD wave function.

4.5.1. The GTG method

In the above discussion of MP2-GTG theory, we have seen that the general strategy is to use a highly flexible basis of floating GTGs for the pair excitations, Equation (47), to evaluate the three-electron integrals explicitly and to avoid four-electron integrals by the WO formulation. The extension of this approach to the CCSD wave function requires an additional singles basis, the neglect of certain small, non-factorisable terms and the use of the SWOP (super weak orthogonality with projection) approximation to avoid the four- and five-electron integrals in the amplitude equations [76,77,83]. For the singles excitations, a basis of floating GTOs is employed [83],

$$\phi_v(1) = x'_{1P_v} y''_{1P_v} z''_{1P_v} \exp[-\alpha_v(r_1 - P_v)^2]. \quad (65)$$

Indeed, the Hartree–Fock wave function is usually already expanded using these functions, since a much smaller basis with lower powers of l, m, n is required than for atom-centred GTOs. The singles excitation operator is

$$\hat{T}_1 = \sum_{iv} c_i^v \sum_x s_v^x E_{xi}. \quad (66)$$

The integrals $s_v^x = 0$ if x is an occupied index, which is enforced by using the projected GTOs $(1 - \hat{O})\phi_v$ as basis functions. The non-linear parameters in the GTO and GTG basis sets are not optimised during the CCSD iterations, since this is an unnecessary expense. The GTG basis that was optimised for the MP2 calculations is usually appropriate for the CCSD method and the singles exponents and centres are pre-optimised by making use of the variational nature of the singles contribution to the MP4 energy. In the CCSD method, the singles c_i^v and doubles c_{ij}^v amplitudes that define the wave function $\exp(\hat{T}_1 + \hat{T}_2)|\text{HF}\rangle$ are determined iteratively using the equations

$$\langle \mu_1 | [\hat{f}, \hat{T}_1] | \text{HF} \rangle = -\langle \mu_1 | e^{-\hat{T}} \hat{V} e^{\hat{T}} | \text{HF} \rangle, \quad (67)$$

$$\langle \mu_2 | [\hat{f}, \hat{T}_2] | \text{HF} \rangle = -\langle \mu_2 | e^{-\hat{T}} \hat{V} e^{\hat{T}} | \text{HF} \rangle. \quad (68)$$

The amplitudes from the previous iteration are used to evaluate the right-hand side and new amplitudes are determined from the resulting linear equations. In the WO formulation of the CCSD amplitude equations,

as proposed by Jeziorski *et al.* [76], the amplitudes for the next iteration are determined by minimising the functionals

$$\langle w_i | \hat{f}_1 - \varepsilon_i + \hat{W}_i | w_i \rangle + 2 \langle w_i | e^{-\hat{T}} \hat{V} e^{\hat{T}} | \text{HF} \rangle, \quad (69)$$

$$\langle w_{ij} | \hat{f}_1 + \hat{f}_2 - \varepsilon_i - \varepsilon_j + \hat{W}_{ij} | w_{ij} \rangle + 2 \langle w_{ij} | e^{-\hat{T}} \hat{V} e^{\hat{T}} | \text{HF} \rangle, \quad (70)$$

where

$$|w_i\rangle = \sum_v c_i^v \sum_x r_v^x E_{xi} | \text{HF} \rangle, \quad (71)$$

$$|w_{ij}\rangle = \frac{1}{2} \sum_v c_{ij}^v \sum_{xy} r_v^{xy} E_{xi} E_{yj} | \text{HF} \rangle, \quad (72)$$

$$\hat{W}_i = [\varepsilon_i - \varepsilon_1 + \eta] \hat{O}, \quad (73)$$

and \hat{W}_{ij} has been given in Equation (51). In this way, the projectors $1 - \hat{O}$ and \hat{Q}_{12} may be dropped from the definition of the cluster operator and from the projection manifold, in the same way as for MP2 theory. If the cluster operator from the previous iteration satisfies SO, then the penalty ensures that the new amplitudes will also satisfy SO in the limit of a complete basis. In practice, a degree of non-SO enters the cluster function and a cleansing projector is applied to reduce this contamination before evaluating the right-hand side for the next iteration,

$$c_i^v = \sum_{\mu} \langle \phi_v | \phi_{\mu} \rangle^{-1/2} \langle \phi_{\mu} | (1 - \hat{O}) | w_i \rangle, \quad (74)$$

$$c_{ij}^v = \sum_{\mu} \langle g_v | g_{\mu} \rangle^{-1/2} \langle g_{\mu} | \hat{Q}_{12} | w_{ij} \rangle. \quad (75)$$

This is the WOP method and the parameter η may be set to zero when the projector is used. The Fock-matrix elements require only three-electron integration, but several terms on the right-hand side require four-electron integration. Some of these terms may be reduced to three-electron integrals by applying the super-weak-orthogonality (SWO) approximation, where particular $(1 - \hat{O})$ projectors are strategically dropped from some terms of the commutators $[\hat{V}, \hat{T}_1 + \hat{T}_2]$ and $[[\hat{V}, \hat{T}_1], \hat{T}_1 + \hat{T}_2]$ in the doubles vector function. The remaining four-electron terms are the non-factorisable contributions to the commutators $[[\hat{V}, \hat{T}_2], \hat{T}_1 + \hat{T}_2]$ and $[[[\hat{V}, \hat{T}_1], \hat{T}_1], \hat{T}_2]$. These terms are generally small and are simply neglected, to give the FCCSD-SWOP method, which only requires three-electron integral evaluation. Detailed expressions for the above formulae may be found in [76,83].

The SWOP method has been used to compute the CCSD energies of Li^+ , H_2 , Be , Li^- and LiH , with an

accuracy of a few microhartrees in all cases but Li^- , where an error of the order of $100 \mu E_h$ was reported [83]. To obtain accurate total correlation energies and quantitative agreement with experiment, triples and higher corrections should be added, but good estimates for these contributions can be obtained using conventional orbital-based calculations. Patkowski *et al.* have recently used the SWOP method to compute the CCSD contributions to the He_2 potential energy curve. Combining the results with post-CCSD corrections, they achieve an accuracy of 3 mK for the interaction energy, which is smaller than the contributions from nonadiabatic and relativistic corrections [85,86]. The primary limitation of the CCSD-GTG method is the evaluation and storage of the large number of three-electron integrals. It is currently feasible to treat systems with ten electrons at most.

4.5.2. The R12 method

The generalisation of the R12 method to the coupled-cluster wave function is much more straightforward than for the GTG method, since no additional approximations are required beyond those applied at the MP2 level. In the CCSD-R12 method, the wave function is

$$\exp(\hat{T}_1 + \hat{T}_2 + \hat{T}_2)|\text{HF}\rangle. \quad (76)$$

The singles and doubles cluster operators of an orbital only calculation, which excite into the Hartree-Fock virtual orbitals, are combined with an additional doubles operator for the geminal excitations. The basis functions for the R12 geminals are the same as those used in MP2 theory,

$$g_\nu(1, 2) = \hat{Q}_{12} \hat{V}_{12} f(r_{12}) \phi_k \phi_l, \quad (77)$$

where $f(r_{12}) = r_{12}$ in the original CCSD-R12 method and $f(r_{12}) = \exp(-\gamma r_{12})$ in the recent CCSD-F12 method. The index ν may be replaced with the pair index kl . The coupled-cluster energy and amplitude equations in the CCSD-R12 and CCSD-F12 methods are

$$E = E_{\text{HF}} + \langle \text{HF} | \hat{V} (\frac{1}{2} \hat{T}_1^2 + \hat{T}_2 + \hat{T}_2) | \text{HF} \rangle, \quad (78)$$

$$\mathbf{0} = \langle \mu_1 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \text{HF} \rangle, \quad (79)$$

$$\mathbf{0} = \langle \mu_2 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \text{HF} \rangle, \quad (80)$$

$$\mathbf{0} = \langle \mu_2 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \text{HF} \rangle. \quad (81)$$

By successive insertions of an approximate RI, the three-, four- and five-electron integrals in the amplitude equations are expressed in terms of two-electron integrals. Some terms in the geminal amplitude

equations require up to four RI insertions and a large auxiliary basis would be needed to ensure numerical stability. Since these terms are both small and expensive, they are neglected by making an approximation analogous to CC2 for the R12 doubles vector function

$$\mathbf{0} = \langle \mu_2 | e^{-\hat{T}} \hat{f} e^{\hat{T}} + e^{-\hat{T}_1 - \hat{T}_2} \hat{V} e^{\hat{T}_1 + \hat{T}_2} | \text{HF} \rangle. \quad (82)$$

When the amplitudes are determined in this way, the method is referred to as CCSD(R12) or CCSD(F12). Detailed expressions for the CCSD-R12 method without an auxiliary basis set have been presented by Noga and co-workers [93,174] and the CCSD(R12) formula with the inclusion of a CABS basis have been derived by Hättig and co-workers [132,134,135,137]. The basis-set errors in the correlation energy of Ne for the CCSD, CCSD(R12) and CCSD(F12) methods are plotted in Figure 4 for the orbital basis sets aug-cc-pVXZ with $X=2, 3, 4, 5$ and 6. As a general rule, the accuracy of a CCSD(F12) calculation using an orbital basis with cardinal number X is better than that of a conventional calculation using a basis with $X+2$. For a given orbital basis, a CCSD(F12) calculation is only a factor of three times more expensive than the CCSD method and is therefore a practical tool for performing meaningful coupled-cluster calculations on large systems in small orbital basis sets. A perturbative triples correction may be computed from the \hat{T}_1 and \hat{T}_2 amplitudes in the same way as for conventional CCSD(T) and Tew *et al.* have demonstrated that

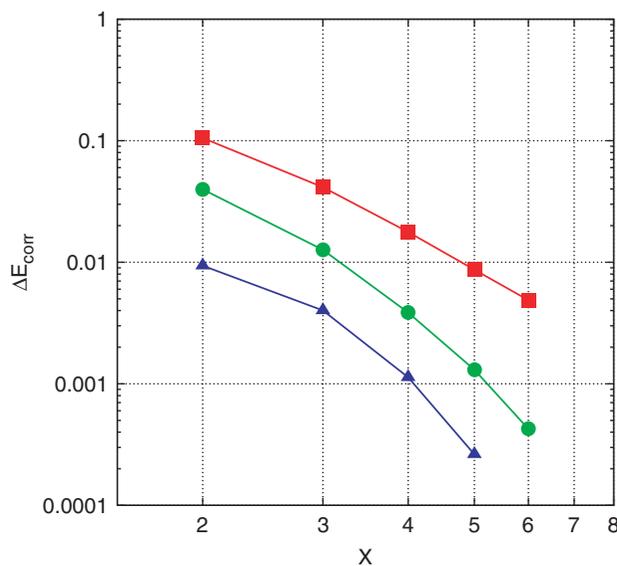


Figure 4. Basis-set error of CCSD ■, CCSD(R12) ● and CCSD(F12) ▲ correlation energies of Ne using aug-cc-pVXZ basis sets.

reaction energies can be computed to within a few kJ mol^{-1} of the basis-set limit using the CCSD(T)(F12) method and a triple-zeta orbital basis [137].

The CCSD(F12) method may also be applied to compute response properties such as excitation energies and polarisabilities, although so far calculations have only been performed using the CC2-R12 and CCSD(R12) methods [132,133,135]. To treat ground and excited states on the same footing, Hättig and co-workers found that it is necessary to extend the R12 basis such that the correlation hole of the excited state is also represented [132,133]

$$g_v(1,2) = \hat{Q}_{12} \hat{V}_{12} f(r_{12}) \phi_p \phi_q. \quad (83)$$

Tew *et al.* have observed that geminal BSSE is problematic for CCSD(F12) calculations of weak interactions, just as in MP2-F12 theory, and have extended Ten-no's rational-generator approach to the CCSD method [138]. The \hat{T}_2 amplitudes are chosen such that the first-order coalescence conditions are fulfilled, Equation (61), and the \hat{T}_1 and \hat{T}_2 are optimised in the presence of the fixed F12 amplitudes. The energy is evaluated from the Lagrangian

$$E = E_{\text{HF}} + \langle \text{HF} | \hat{V} (\frac{1}{2} \hat{T}_1^2 + \hat{T}_2 + \hat{T}_2) | \text{HF} \rangle + \bar{c} \langle \mu_2 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \text{HF} \rangle, \quad (84)$$

with multipliers $\bar{c} = c$. This simplified approach performs as well as the CCSD(F12) method for reaction energies. This is illustrated in Figure 5, where we have plotted idealised error distributions of the fixed amplitude CCSD(F12) and CCSD(T)(F12) correlation energies for a set of 15 reactions involving molecules of H, C, N, O, S and Cl atoms, using an aug-cc-pVTZ basis. The corresponding errors for the conventional CCSD and CCSD(T) methods with $X=3, 4$ and 5 are presented for comparison and the basis-set limit for the

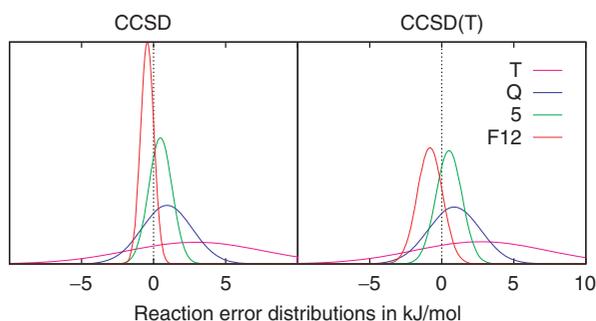


Figure 5. Basis-set errors for CCSD and CCSD(T) correlation contributions to 15 reaction energies using the aug-cc-pVXZ basis sets and the fixed-amplitude F12 method with the aug-cc-pVTZ basis.

correlation energies was obtained from extrapolated 56 orbital-only values [138]. The fixed-amplitude method returns quintuple-zeta quality results with a triple-zeta orbital basis, which for CCSD(T) is often sufficient for an agreement with experiment to within chemical accuracy (4 kJ mol^{-1}).

Even simpler approaches have been proposed by Valeev [139] and Adler *et al.* [140]. In Valeev's method, the R12 contribution to the energy is computed in a perturbative manner, using the converged CCSD similarity-transformed Hamiltonian as the zeroth order operator. In Adler's method, only the explicitly correlated contributions that are linear in the amplitudes are retained in the coupled-cluster equations. The computational cost of both of these approaches is only marginally greater than a conventional CCSD calculation and the intermediates required are all available from the preceding MP2-F12 calculation. These efficient methods both return quintuple-zeta quality correlation energies using only a triple-zeta orbital basis and many applications are expected to appear in the near future.

5. Small terms

For small molecules, it is possible to compute valence-shell CCSD(T) correlation energies to within a few tenths of a millihartree, provided that large correlation-consistent basis sets (cc-pV5Z, preferably with extrapolation) or explicitly correlated wave functions are used. The remaining basis-set error corresponds to a few kJ mol^{-1} per atom (not counting H) in the total electronic energy and the accuracy for relative energies (binding energies, heats of formation, atomisation energies) is typically an order of magnitude smaller due to consistent error cancellation. Typical experimental error bars vary from 0.1 kJ mol^{-1} up to a few kJ mol^{-1} . For example, the atomisation energy of methane in the Active Thermochemical Tables (ATcT) amounts to $1642.24 \pm 0.12 \text{ kJ mol}^{-1}$ (concerning the ATcT data, see for example [175] and [176]). It should be noted that error bars on experimentally derived values usually refer to a 95% confidence limit, implying that the mean absolute deviation of computed results should be 2–3 times smaller than the experimental error bar for a direct and fair comparison between theory and experiment. Nonetheless, for many systems it is possible to perform frozen-core (fc) CCSD(T) calculations with basis-set truncation errors smaller than the experimental error bars on the quantity of interest.

However, many small contributions to the energy are neglected in a fc-CCSD(T) calculation and, for

complete agreement to within the experimental error bar, a number of other contributions to the total molecular energy must be computed. In particular, we would have to take into account the effect of the full CCSDT model, the effect of quadruple excitations (CCSDTQ) and higher, as well as core orbitals, to approach the all-electron (ae) FCI limit as closely as possible. Furthermore, zero-point vibrational energies (ZPVE) must be considered, and, if thermochemical data are to be computed with sub-kJ mol⁻¹ accuracy, not only harmonic ZPVEs but also the corresponding anharmonic corrections. Another source of error is related to the fact that we have so far considered the nonrelativistic Schrödinger equation in the Born–Oppenheimer approximation. These errors may also be significant and need be addressed by computing relativistic corrections and diagonal BO corrections (DBOCs). In Table 2, we give a few examples of the magnitudes of these corrections for atomisation

energies of small hydrocarbons, whereas, in Table 3, we give post-CCSD(T) contributions to harmonic vibrational frequencies of diatomic molecules. Of course, when comparing computed atomisation energies with experimentally derived values, it is important to have access to very accurate and reliable experimental data. For example, the atomisation energies of alkanes depend critically on the standard heats of formation of the constituent atoms, particularly on $\Delta_f H^0(0\text{ K})$ of gaseous carbon atoms [181]. In the following sections, we discuss the various small terms in detail, providing further examples for the magnitudes of each term and the ways in which they are computed.

5.1. Anharmonic ZPVEs

Whenever computed molecular heats of formation or atomisation energies are compared directly with

Table 2. Small corrections to CCSD(T)-R12 atomisation energies ($\sum D_0$ in kJ mol⁻¹) of selected hydrocarbons. Data taken from [104].

Molecule	CCSD(T) ^a	CV ^b	ZPVE ^c	Anh. ^d	SO ^e	D1 ^f	MV ^g	Total	JANAF ^h
Methane	1753.1	4.9	-116.3	1.4	-0.4	2.1	-2.9	1642.0	1642.2 ± 0.6
Acetylene	1684.9	9.6	-69.2	0.8	-0.7	3.5	-4.6	1624.3	1619 ± 1
Ethylene	2349.0	9.2	-132.6	1.5	-0.7	3.8	-5.2	2225.1	2225.5 ± 0.7
Allene	2929.0	14.3	-143.2	1.7	-1.1	5.6	-7.5	2798.7	2799.2 ± 1.3

^aFrozen-core fc-CCSD(T)-R12 value.

^bContribution due to including core orbitals (ae-CCSD(T) treatment).

^cHarmonic zero-point vibrational energy.

^dAnharmonic correction to the harmonic zero-point vibrational energy.

^eCorrection due to the spin-orbit splitting of the ³P state of atomic carbon.

^fRelativistic one-electron Darwin correction.

^gRelativistic one-electron mass-velocity correction.

^hNIST-JANAF thermochemical tables [177].

Table 3. Best estimates of the harmonic vibrational frequencies (in cm⁻¹) of the diatomic closed-shell molecules HF, N₂, F₂ and CO.

Contribution	HF	N ₂	F ₂	CO
CCSD(T)-R12 ^a	4142.5	2363.4	931.9	2167.3
CCSDTQ-CCSD(T) ^b	-4.5	-9.1	-12.2	-6.5
CCSDTQ5-CCSDTQ ^c	-0.1	-3.9	-0.8	0.0
Core-correlation correction ^d	4.0	9.8	1.6	9.9
Relativistic DPT correction ^e	-2.8	-0.8	-0.4	-0.9
Breit correction ^e	-0.6	-0.5	-0.3	-0.4
Best estimate	4138.5	2358.9	919.8	2169.4
Experiment ^f	4138.4	2358.6	916.6	2169.8

^aFrom [103].

^bFrozen-core cc-pVTZ level, taken from [178].

^cFrozen-core cc-pVDZ level, taken from [178].

^dFrom ae- and fc-CCSD(T)/aug-cc-pCV5Z calculations, taken from [178].

^eFrom [103].

^fFrom [179]. See [180] for HF.

experimentally derived data, or used to predict thermochemical data, the vibrational and rotational energies must be computed. The standard procedure for evaluating the energy levels is through the rigid-rotor and harmonic-oscillator approximations, but if the target accuracy is one kJ mol^{-1} or less, anharmonic corrections must be taken into account, even for the zero-point vibrational energies (ZPVEs). Using the rigid-rotor and harmonic-oscillator as the zeroth order Hamiltonian, second-order ro-vibrational perturbation theory can be used to evaluate the energy levels. The *ab initio* calculation of anharmonic corrections is very time consuming, however, since it involves the calculation of the full cubic force field together with the semidiagonal part of the quartic force field (of the form ϕ_{ijj} only). Usually, the full cubic and partial quartic force fields are computed from finite differences of analytic second derivatives. For example, full quartic force fields have been computed for small hydrides such as AlH_3 [182], BH_3 and BH_5 [183,184] as well as SiH_3^- [185]. Table 2 shows that the anharmonic correction to the atomisation energy of a small hydrocarbon is of the order of $1\text{--}2\text{ kJ mol}^{-1}$. These anharmonic corrections were obtained by DFT (B97-1 functional) as discussed in [186] and [187]. Recently, Wheeler and co-workers published accurate calculations of the enthalpies of formation of propyne, allene and C_3H_3 isomers. In their work, these authors computed the anharmonic corrections from MP2 quartic force fields [188]. The effect of the anharmonicities was of the order of $\pm(0.3\text{--}1.3)\text{ kJ mol}^{-1}$ on the heats of formation at 0 K.

Bomble *et al.* have pointed out that it is important to consider the constant contribution G_0 to the zero-point energy, which is often absent from VPT2 calculations [189]. The correct equation for the vibrational zero-point energy is

$$E_{\text{ZPVE}} = G_0 + \sum_i \frac{\omega_i}{2} + \frac{1}{4} \sum_{i \leq j} x_{ij}, \quad (85)$$

where ω_i are the harmonic frequencies and x_{ij} the anharmonicity constants. The full expression for the G_0 contribution is given in Reference 30 in the work of Bomble *et al.* [189]; see also [183,190]. The effect of neglecting the constant G_0 term amounted to a few tenths of a kJ mol^{-1} for the enthalpies of formation (at 0 K) for molecules/radicals such as C_2H_2 , CH_3 , HCO and HO_2 [189].

For a simple small ‘organic’, semi-rigid molecule such as ethane, with six CH stretches, the anharmonic ZPVE amounts to 194.1 kJ mol^{-1} [191], while its harmonic counterpart is 196.3 kJ mol^{-1} [191,192]. In [191], in a first attempt to provide a crude estimate

of the anharmonic ZPVE, the mean value of the harmonic ZPVE and half the sum of the experimentally observed fundamentals was used. This approach yielded a ZPVE of 192.9 kJ mol^{-1} , which is in error by about 1.2 kJ mol^{-1} . This error consists of one quarter of the sum of the diagonal anharmonicity constants and the G_0 term. It thus appears that this simple approach of combining harmonic frequencies and experimental data, is not accurate enough for sub- kJ mol^{-1} accuracy.

In summarising, we conclude that the anharmonic corrections including the G_0 term must be computed for high-accuracy work, but it seems sufficient to compute the full cubic and the semidiagonal quartic force fields at the level of MP2 or DFT.

5.2. Post-CCSD(T) theory

Although the ever-popular CCSD(T) method goes a long way towards chemical accuracy, it is not always sufficient and is almost always insufficient for sub- kJ mol^{-1} accuracy. For example, for the (slightly endothermic) isomerisation reaction propyne \rightarrow allene, the effect of full triples, CCSDT, is to decrease the reaction energy by 0.7 kJ mol^{-1} with respect to the CCSD(T) value, while the step from CCSDT to CCSDT(Q) increases the reaction energy by 0.4 kJ mol^{-1} . The CCSDT and CCSDT(Q) corrections were computed using cc-pVTZ and cc-pVDZ basis sets, respectively [188]. These energy differences are typical for molecules of this size, and must be evaluated when sub- kJ mol^{-1} accuracy is desired. Using the CCSDTQ model, Ruden and co-workers [193] have examined the contributions from connected quadruple excitations to the atomisation energies of CH_2 , H_2O , HF , N_2 , F_2 and CO . For N_2 and F_2 , the T_4 contribution was larger than 3 kJ mol^{-1} . On average, for these six molecules, the quadruples correction to the CCSDT atomisation energies reduced the mean errors from 2.3 to 0.5 kJ mol^{-1} .

In comparison with CCSD(T), it has been observed that the full CCSDT model almost universally decreases molecular binding energies. At the next level, connected quadruple excitations (T_4) systematically increase molecular binding energies. Because of this error cancellation, superior performance with respect to the CCSD(T) model is only obtained after including T_4 , either in the CCSDTQ or in the CCSDT(Q) model. Indeed, the CCSDT(Q) model seems to have become a promising method for high-accuracy work. Its perturbative correction for connected quadruple excitations has been derived from a non-Hermitian perturbation theory analogous to

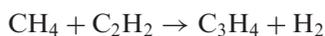
that used to justify the perturbative correction in CCSD(T) theory [194]. The CCSDT(Q) model has been shown to perform very well in comparison with the full CCSDTQ model [195].

The effect of quintuple excitations (T_5) was studied for a set of molecules (eight-electron systems) by Boese and co-workers in the framework of W3 theory [196]. Except for molecules with a multireference character, the T_5 contributions seemed to be largely negligible. It should be noted, however, that for multireference cases such as C_2 , N_2 , NO_2 or ozone, the connected quintuple excitations account for as much as 0.5–1.7 kJ mol^{-1} [193,197] to their atomisation energy. For C_2 , even the contribution due to sextuple excitations (T_6) still amounts to about 0.3 kJ mol^{-1} [197].

In conclusion, we note that the above high-level coupled-cluster treatments have only recently become possible through the advancement of general coupled-cluster schemes [198–200]. Furthermore, an extensive study of post-CCSD contributions to molecular atomisation energies has been published recently by Karton and co-workers [201], including fully iterative coupled-cluster theory up to the level of connected quintuple excitations (CCSDTQ5).

5.3. Core and core–valence correlation effects

Since coupled-cluster calculations of medium-sized molecules in large basis sets are often very time-consuming, it is tempting to include only the occupied orbitals of the valence shell in the correlation treatment. This approach, known as the frozen-core (fc) approximation, is usually regarded as a good approximation when energy differences are concerned because the core orbitals (e.g. $1s$ of C, $1s2s2p$ of Si, etc.) do not change much from one chemical system to another. Nevertheless, Table 2 shows that the energy contribution to the atomisation energy of methane, acetylene, ethylene and allene from core–valence (CV) correlation amounts to 4.9, 9.6, 9.2 and 14.3 kJ mol^{-1} , respectively. Admittedly, the effect on a reaction energy such as



is much smaller, with $\Delta_{\text{CV}} = 0.2 \text{ kJ mol}^{-1}$. In any case, the error due to the frozen-core approximation cannot be neglected when very high accuracy is aimed at. Usually, CCSD(T) calculations are performed in a large basis that is able to describe core and core–valence correlation effects (aug-cc-pwCVTZ and/or aug-cc-pwCVQZ) both with and without freezing the core. For example, the values in Table 2 have been obtained from (all-electron) ae-CCSD(T) and fc-CCSD(T)

calculations in a cc-pCVQZ-like basis. In the W4 theory of Karton *et al.* [197], the CV contribution to the atomisation energy is obtained from CCSD(T) calculations in the basis sets aug-cc-pwCVTZ and aug-cc-pwCVQZ followed by basis-set extrapolation.

One might wonder whether the CCSD(T) model is sufficient to describe core and core–valence correlation corrections. In 1999, Martin and de Oliveira noted in their work on the Weizmann-1 theory (*vide infra*) that the perturbative (T) triples correction alone is responsible for about one half of the total CCSD(T) CV correction [202]. Fortunately, as shown by Karton *et al.* [197], post-CCSD(T) CV contributions from the full CCSDT model or even the CCSDT(Q) model are basically negligible, except perhaps for molecules such as C_2 and BN, where the change from CCSD(T) to CCSDT has an effect of about 0.4–0.7 kJ mol^{-1} on the atomisation energy. The corresponding effect on the reaction energy of the reaction



was of the order of one tenth of a kJ mol^{-1} . In summary, post-CCSD(T) CV corrections appeared to be negligibly small in most cases, at least for first-row systems. Of course, not much is known about CCSDT and CCSDT(Q) CV corrections to atomisation energies of second-row systems due to the immense computational costs (amplified by the need for large basis sets), except for the smallest molecules. Reference [197] provides a good overview of the data available.

5.4. Relativistic corrections

Relativistic corrections to electronic energies of light (elements H–Ar) molecular systems can be computed at little or no cost by first-order perturbation theory, using the one-electron Pauli Hamiltonian – that is, the one-electron mass–velocity and Darwin operators,

$$\hat{H}^{\text{MV}} = -\frac{1}{8c^2} \sum_i p_i^4, \quad (86)$$

$$\hat{H}^{\text{D1}} = -\frac{\pi}{2c^2} \sum_{i,K} Z_K \delta(r_{iK}). \quad (87)$$

This approach is often referred to as the Cowan–Griffin method [203]. Applied at the Hartree–Fock level, this method is usually not in error by more than 10% with respect to the full relativistic correction at the FCI level including all terms to second order in the fine-structure constant $\alpha = e^2/(\hbar c 4\pi\epsilon_0) = 1/137.03599907$ ($=1/c$ in atomic units). Here, and in the other perturbative treatments of relativistic effects

discussed below, α^2 is the perturbational parameter. Hence, first-order perturbation theory yields corrections of order $\alpha^2 = \mathcal{O}(1/c^2)$, second-order perturbation theory yields corrections of $\mathcal{O}(\alpha^4) = \mathcal{O}(1/c^4)$, and so on. We shall restrict our discussion to closed-shell systems, where, to first order in perturbation theory, only scalar relativistic terms need to be considered.

Direct perturbation theory (DPT) was developed by Rutkowski and Kutzelnigg in the late 1980s [204–209]. It is an almost equivalent alternative to the Cowan–Griffin method. Scalar relativistic, first-order DPT can be understood as a perturbation that modifies the overlap integrals as well as the one-electron and two-electron integrals of the nonrelativistic Hamiltonian [210,211],

$$S'_{pq} = (4c^2)^{-1} \langle \chi_p | \mathbf{p} \cdot \mathbf{p} | \chi_q \rangle = (2c^2)^{-1} T_{pq}, \quad (88)$$

$$H'_{pq} = (4c^2)^{-1} \langle \chi_p | \mathbf{p} V \cdot \mathbf{p} | \chi_q \rangle, \quad (89)$$

$$(pq|rs)' = (4c^2)^{-1} (\chi_p \chi_q | \mathbf{p} \chi_r \cdot \mathbf{p} \chi_s) + (4c^2)^{-1} (\mathbf{p} \chi_p \cdot \mathbf{p} \chi_q | \chi_r \chi_s), \quad (90)$$

where T_{pq} is a matrix element of the kinetic energy operator and V is the electron–nucleus Coulomb potential. Since the DPT involves perturbed two-electron integrals, it seems more complicated than the Cowan–Griffin approach, where only the perturbed one-electron Hamiltonian $H'_{pq} = H_{pq}^{\text{MV}} + H_{pq}^{\text{D1}}$ is taken into account. However, if all terms of order $1/c^2$ are retained when expanding the Dirac–Coulomb equation, then the two-electron Darwin operator also arises in addition to the one-electron Darwin operator in the Cowan–Griffin method,

$$\hat{H}^{\text{D2}} = -\frac{\pi}{2c^2} \sum_{i \neq j} \delta(\mathbf{r}_{ij}), \quad (91)$$

and the full Darwin operator is $\hat{H}^{\text{D}} = \hat{H}^{\text{D1}} + \hat{H}^{\text{D2}}$. A Cowan–Griffin-like method including the full Darwin operator yields the one- and two-electron perturbed operators, in a similar way to DPT,

$$S'_{pq} = 0, \quad (92)$$

$$H'_{pq} = H_{pq}^{\text{MV}} + H_{pq}^{\text{D1}}, \quad (93)$$

$$(pq|rs)' = -\frac{\pi}{c^2} (\chi_p \chi_q | \delta(\mathbf{r}_{12}) | \chi_r \chi_s). \quad (94)$$

In the limit of a complete one-electron basis, the perturbation theories (88)–(90) and (92)–(94) become fully equivalent. Minor differences are found using finite basis sets, as shown in [212].

The scalar-relativistic, first-order operators of the Cowan–Griffin and direct perturbation theories can be interpreted as one of the perturbations in

a higher-order perturbation theory for the analytic calculation of the relativistic effect on molecular properties [211,213,214]. Similarly, in the spirit of perturbation theory, Helgaker *et al.* [215] have constructed a second-quantisation electronic Hamiltonian, valid for all values of the fine-structure constant α . The resulting α -dependent Hamiltonian is then used to set up a perturbation theory in orders of α^2 , using the general framework of time-independent response theory, in the same manner as for geometrical and magnetic perturbations. In an invited article in this journal, Tarczay *et al.* give a comprehensive account of various relativistic corrections to the total energies of molecules such as SiC_2 , H_2S , C_2H_6 (and others) as well as to the activation barriers of various isomerisation reactions [212].

In pursuit of very accurate electronic-structure calculations, one may wonder whether one should not start from the Dirac–Coulomb–Breit equation in place of its Dirac–Coulomb counterpart to include all terms of $\mathcal{O}(\alpha^2)$. Again restricting the discussion to first-order corrections to the energies of closed-shell systems, the perturbation expansion of the Dirac–Coulomb–Breit equation gives rise to two new operators, the spin–spin (\hat{H}^{SS}) and orbit–orbit (\hat{H}^{OO}) two-electron operators,

$$\begin{aligned} \hat{H}^{\text{SS}} &= \frac{1}{8c^2} \sum_{i \neq j} \left(\boldsymbol{\sigma}_i \cdot \mathbf{r}_{ij}^{-3} \boldsymbol{\sigma}_j - 3 \boldsymbol{\sigma}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij}^{-5} \mathbf{r}_{ij} \cdot \boldsymbol{\sigma}_j - \frac{8\pi}{3} \delta(\mathbf{r}_{ij}) \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \right), \end{aligned} \quad (95)$$

$$\hat{H}^{\text{OO}} = -\frac{1}{4c^2} \sum_{i \neq j} \left(\mathbf{p}_i \cdot \mathbf{r}_{ij}^{-1} \mathbf{p}_j + \mathbf{p}_i \cdot \mathbf{r}_{ij} \mathbf{r}_{ij}^{-3} \mathbf{r}_{ij} \cdot \mathbf{p}_j \right). \quad (96)$$

The expectation value of \hat{H}^{SS} for a closed-shell system is equal to minus two times the two-electron Darwin term,

$$\langle \text{cs} | \hat{H}^{\text{SS}} | \text{cs} \rangle = -2 \langle \text{cs} | \hat{H}^{\text{D2}} | \text{cs} \rangle. \quad (97)$$

Hence, adding all terms up to the second order in the fine-structure constant α , one obtains for a closed-shell system the following expectation value of the Breit–Pauli Hamiltonian:

$$\langle \text{cs} | \hat{H}^{\text{BP}} | \text{cs} \rangle = \langle \text{cs} | \hat{H}^{\text{NR}} + \hat{H}^{\text{MV}} + \hat{H}^{\text{D1}} - \hat{H}^{\text{D2}} + H^{\text{OO}} | \text{cs} \rangle, \quad (98)$$

where \hat{H}^{NR} is the nonrelativistic Hamiltonian. All of the terms in Equation (98) have been implemented at the closed-shell CCSD(T) level by Coriani *et al.* [216].

Although the two-electron Darwin energy correction is not substantial in absolute terms, its magnitude is approximately equal to the *correlation contribution*

to the mass–velocity and one-electron Darwin energies [212]. Therefore, when these energies are calculated at the correlated level, the two-electron Darwin term should also be computed. But then the correction due to the Breit interaction ($\hat{H}^{SS} + \hat{H}^{OO}$) should be computed as well, since, in turn, this correction is of the same magnitude as the two-electron Darwin energy correction. Hence, consistent treatments of scalar relativistic effects are obtained either by applying the Cowan–Griffin method at the Hartree–Fock level, with an error of 10%, or by including all operators of Equation (98) at the correlated coupled-cluster level for sub-kJ mol⁻¹ accuracy.

The Cowan–Griffin and DPT approaches are not the only methods for computing (scalar) relativistic corrections to molecular energies. Other examples include the Douglas–Kroll–Hess (DKH) method [217–220], the zero-order regular approximation (ZORA) method of Chang *et al.* [221] and van Lenthe *et al.* [222,223] (see also [224] for using the ZORA in Hartree–Fock-based theories) and Dyall’s method of elimination of small components [225–228] within the matrix representation of the modified Dirac equation. See also [229–231] for discussions of the transformations of the Dirac operator and its matrix representation and the relations between the DPT, DKH and ZORA methods. All of these methods could be used as an alternative to the Cowan–Griffin and DPT approaches, but Kutzelnigg and Liu [232] conclude that ‘DPT has always been a good approximation in those cases, where *only low orders matter*, that is, mainly for molecules built from light atoms. It will continue to be competitive in this domain even with exact quasi-relativistic treatments’. Indeed, especially for light systems, it is important to treat the two-electron terms properly [224,233] and the DPT approach seems the method of choice if highly accurate energies for light and small molecules are to be computed. For heavy elements, relativistic corrections of high order in the fine-structure constant α may become important. For example, effects due to second-order spin–orbit coupling were found to increase the binding energies of Br₂ and I₂ by 1.7 and 8.4 kJ mol⁻¹, respectively [234]. For Cl₂, the effect may be of the order of a few tenths of a kJ mol⁻¹ and thus significant for high-accuracy work [196].

The Lamb shift is another relativistic correction of higher order in α , which, for example, has been included in highly accurate calculations of a potential-energy hypersurface for the water molecule and its rotation–vibration transitions [235]. For the results presented in Tables 2 and 3, neither the second-order spin–orbit coupling nor the Lamb shift were calculated, but the Breit correction was included in Table 3.

The Lamb shift can be estimated (presumably accurately to within 10%) from the D1 term with an appropriate scaling factor [236], which is about 0.05 for hydrogen, 0.025 for the atoms Li–Ne and 0.015 for second-row atoms (Na–Ar). These are very small scaling factors, but the reader should note that the one-electron Darwin term \hat{H}^{D1} is much larger than the total first-order relativistic correction due to the cancellation between the \hat{H}^{D1} and \hat{H}^{MV} terms (Table 2). Dyall *et al.* have investigated whether the Lamb shift is chemically significant for atomisation energies of the group 13 fluorides BF₃, AlF₃ and GaF₃ [237]. They found that it increased the atomisation energies of BF₃, AlF₃ and GaF₃ by about 0.2, 0.3, and 0.8 kJ mol⁻¹, respectively. This correction is of the same order of magnitude as the Breit interaction for the valence shells of atoms [238]. We conclude that in highly accurate calculations with a target accuracy of 0.1 kJ mol⁻¹, one cannot neglect the second-order spin–orbit coupling and the Lamb-shift contributions.

Finally, first-order spin–orbit corrections must be considered in calculations of open-shell systems. Usually, experimental values for the spin–orbit splittings are used to correct the atomic energies when computing atomisation energies as well as to correct ²Π states of diatomics. For example, in Table 2, the atomisation energies are corrected for the (experimentally observed) spin–orbit splitting of the ³P state of the carbon atom (the ³P₀ state lies about 0.354 kJ mol⁻¹ below the uncoupled nonrelativistic, ninefold degenerate ³P level).

5.5. Diagonal Born–Oppenheimer corrections

The leading correction to the BO approximation is given by [239]

$$\Delta E_{\text{DBOC}} = \langle \Psi(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi(\mathbf{r}; \mathbf{R}) \rangle, \quad (99)$$

where $\Psi(\mathbf{r}; \mathbf{R})$ is the electronic wave function, $\hat{T}_N(\mathbf{R})$ the nuclear kinetic-energy operator, and \mathbf{r} and \mathbf{R} the Cartesian electronic and nuclear coordinates, respectively. This diagonal Born–Oppenheimer correction (DBOC) was used in conjunction with explicitly correlated methods as early as 1964 by Kołos and Wolniewicz in their accurate calculations on the H₂ molecule [240]. The first researchers to derive general formula for the evaluation of the DBOC for general polyatomic molecules appear to have been Sellers and Pulay, at the Hartree–Fock level of theory [241,242]. Handy *et al.* [243,244] showed how the usual analytic-derivative techniques can be utilised to compute the DBOC, and Kutzelnigg [245] has discussed the issue of

the separation of the molecular centre-of-mass motion. Many authors have presented calculations of the DBOC at the CI level, mostly by numerical differentiation [246–251], and recently, Gauss *et al.* have implemented the general analytic calculation of the DBOC at the coupled-cluster level of theory [252]. In a pilot application, Gauss *et al.* computed the DBOC to the atomisation energy of *trans*-butadiene. At the Hartree–Fock level, the DBOC increases the atomisation energy by about 0.8 kJ mol^{-1} , but at the CCSD level, the DBOC is only about 0.45 kJ mol^{-1} . Hence, to obtain high accuracy, it seems important to compute the DBOC at the correlated level. In the same work, Gauss *et al.* show that the effect of the DBOC on the reaction energy for the reaction $\text{F} + \text{H}_2 \rightarrow \text{HF} + \text{H}$ is to make it 0.2 kJ mol^{-1} more exothermic. The effect on the activation barrier was to raise it by about 0.2 kJ mol^{-1} [252,253].

6. Composite methods

From all the examples and discussion in the preceding sections, it is clear that currently no single method or single quantum-chemical calculation can provide a practical quantitative quantum chemistry. Even the CCSD(T)(F12) method, which returns the near-basis-set limit CCSD(T) energy in a single calculation, lacks the many small terms of the previous section, and to compute all of these terms at this level of theory is prohibitively expensive. Fortunately, it is also unnecessary, since the many contributions may be computed at different levels of theory, appropriate to the accuracy desired. This observation formed the basis of Pople's model chemistry, Gaussian-1 (G1) theory, proposed in 1989 [254]. Pople and co-workers designed a general procedure for the computation of the total energies of molecules at their equilibrium geometries. They optimised the geometries at the MP2/6-31G* level and computed the energy using the MP4/6-311G** method, and subsequently added individually computed corrections for the effect of diffuse *sp* functions, higher polarisation functions *df* and higher-level correlation effects through QCISD(T) corrections. Zero-point vibrational energy corrections were also computed, and in the original 1989 work, atomisation energies for a set of 31 molecules were found to agree with experimental thermochemical data to an accuracy better than about 8 kJ mol^{-1} in most cases. Similar agreement was achieved for ionisation energies and electron and proton affinities [254]. The Gaussian-1 procedure was extended and tested on compounds containing second-row atoms (Na–Cl) in 1990 [255] and other composite methods related to

G1 have since been developed, including the following approaches:

- the Gaussian-*n* ($n = 2, 3, 4$) theories of Curtiss and co-workers [256–258]
- the correlation-consistent composite approach (ccCA) of DeYonker *et al.* [259,260]
- the complete basis-set (CBS) methods of Petersson and co-workers [261–267]
- the focal-point analysis (FPA) approach of Allen and co-workers [268,269]
- the multicoefficient correlation methods (MCCMs) of Truhlar and co-workers [270–272]
- the HEAT approach of Tajti, Bomble and others [189,190]
- the Weizmann-*n* ($n = 1 - 4$) theories of Martin and co-workers [196,197,201,202]

In all of these approaches, various methods and basis-set expansions are combined to determine equilibrium structures, vibrational frequencies (and ZPVEs) and electronic energy contributions. Below, we shall briefly present their most important ingredients, restricting our discussion to the most recent variants of the respective methods.

We note that the above list is necessarily incomplete because there exist a huge number of 'composite schemes' in the literature, in which workers have combined results from various calculations or added empirical corrections or scaling factors. Examples are the BAC-MP4 method of Melius and co-workers [273], in which empirical bond-additivity corrections are added to the MP4 energy, or the HL1 and HL2 methods of Miller and Klippenstein [274], in which QCISD(T) energies are improved by adding large basis-set MP2 energies (HL1) or where TQ-extrapolated MP2 energies are added to DT-extrapolated QCISD(T) energies (HL2). See also [275] and [276], where MP2-limit corrections are added to CCSD(T) energies. Other examples include the 'scaling-all-correlation' (SAC) approach of Gordon and Truhlar [271,272,277] and the similar PCI-X method of Siegbahn *et al.* [278], which is a parameterised correlation method with one adjustable scaling parameter *X*.

Dixon and co-workers (cf. [279–282] and references therein) have been developing an approach to calculate reliably molecular thermodynamic properties (e.g. heats of formation) based on fc-CCSD(T) coupled-cluster theory using correlation-consistent basis sets and extrapolations to the basis-set limit. The fc-CCSD(T) results are supplemented with a number of smaller additive corrections such as core–valence interactions, scalar and spin–orbit relativistic effects and ZPVEs, obtained from experiment or theory,

or some combination. Also Feller and Peterson employed a composite approach when re-examining the atomisation energies of the Gaussian-2 set of molecules [283,284], and Bytautas and Ruedenberg have recently introduced a method based on an additive resolution of the wave function and the energy in terms of contributions from successive excitations in the CI framework [285,286].

6.1. Gaussian-4 theory

The Gaussian-4 theory (G4 theory) has been developed for the calculation of energies of molecules containing first- (Li–Ne), second- (Na–Ar) and third-row main-group atoms (K, Ca, Ga–Kr) [258]. Its key ingredients are B3LYP/6-31G(2df,p) geometries and scaled (by a factor of 0.9854) harmonic vibrational frequencies for the ZPVE, Hartree–Fock energies extrapolated to the limit of a complete basis, and MP4/6-31G(d) single-point calculations. When we consider the MP4/6-31G(d) energy E_{MP4} as the base line of G4 theory, then the following energy corrections are added in the composite scheme:

$$\begin{aligned} E_{\text{G4}} = & E_{\text{MP4}} + \Delta E_{\text{HF}} + \Delta E_{\text{CCSD(T)}} \\ & + \Delta E_{\text{diffuse}} + \Delta E_{\text{polarisation}} \\ & + \Delta E_{\text{G3LargeXP}} + \Delta E_{\text{HLC}} \\ & + \Delta E_{\text{SO}} + \Delta E_{\text{ZPVE}}. \end{aligned} \quad (100)$$

ΔE_{HF} is a correction for the Hartree–Fock basis-set limit,

$$\begin{aligned} \Delta E_{\text{CCSD(T)}} = & E(\text{CCSD(T)}/6\text{-}31\text{G(d)}) \\ & - E(\text{MP4}/6\text{-}31\text{G(d)}), \end{aligned} \quad (101)$$

$\Delta E_{\text{diffuse}}$ and $\Delta E_{\text{polarisation}}$ are corrections for diffuse and polarisation functions, respectively (calculated at the MP4 level), and $\Delta E_{\text{G3LargeXP}}$ is a correction for larger basis-set effects. ΔE_{HLC} is an empirical energy correction that depends on the number of occupied α and β orbitals, ΔE_{SO} is a spin–orbit term (taken from experiment or accurate calculations) and ΔE_{ZPVE} is the zero-point vibrational energy. Core orbitals are taken into account in the ae-MP2/G3LargeXP calculation that gives the $\Delta E_{\text{G3LargeXP}}$ basis-set correction.

The Gaussian-4 theory was tested rigorously on the G3/05 test set [287] including 454 energies. The overall average absolute deviation for these energies was reported to be about 3.5 kJ mol^{-1} . Although it does not provide sub- kJ mol^{-1} accuracy, the G4 theory is inexpensive compared with the more accurate methods such as W4 and HEAT (*vide infra*).

6.2. Correlation-consistent composite approach

DeYonker and co-workers [259,260,288–292] have developed a series of methods denoted ccCA-DZ, ccCA-TZ, ccCA-aTZ and ccCA-CBS. These composite approaches have been developed in the spirit of the Gaussian- n model chemistries of Pople, Curtiss and co-workers, but they do not use empirical corrections such as the energy term ΔE_{HLC} of Gaussian- n theory. The ccCA methods start with B3LYP geometries and scaled frequencies (by a factor of 0.9854) for the ZPVE. Then, the reference energy is either an fc-MP4 energy in the basis sets cc-pVDZ, cc-pVTZ or aug-cc-pVTZ (in the approaches ccCA-DZ, ccCA-TZ or ccCA-aTZ, respectively) or a frozen-core MP2 energy extrapolated to the limit of a complete basis (ccCA-CBS).

In their most recent work [292], DeYonker *et al.* optimise the B3LYP geometries in the cc-pVTZ basis and employ two extrapolation formulae to estimate the basis-set limit of frozen-core MP2 theory. The first formula is due to Peterson *et al.* [293],

$$E(X) = A + Be^{-(X-1)} + Ce^{-(X-1)^2}, \quad (102)$$

and the second formula is the Schwartz4(mn) form of Martin [294],

$$E(X) = A + B/(X + \frac{1}{2})^4. \quad (103)$$

The method is denoted ccCA-P when Equation (102) is used with the aug-cc-pVXZ basis sets with $X=2, 3, 4$, and it is denoted ccCA-S4 when Equation (103) is used with the basis sets aug-cc-pVTZ and aug-cc-pVQZ.

After the frozen-core MP2 basis-set limit $E_{\text{MP2}/\infty}$ has been determined, a number of corrections are added to yield the final ccCA-CBS energy,

$$\begin{aligned} E(\text{ccCA-CBS}) = & E_{\text{MP2}/\infty} + \Delta E_{\text{CC}} + \Delta E_{\text{DKH}} \\ & + \Delta E_{\text{CV}} + \Delta E_{\text{ZPVE}}. \end{aligned} \quad (104)$$

The ΔE_{CC} term is obtained from CCSD(T)/cc-pVTZ calculations,

$$\begin{aligned} \Delta E_{\text{CC}} = & E(\text{CCSD(T)}/\text{cc-pVTZ}) - E(\text{MP2}/\text{cc-pVTZ}). \end{aligned} \quad (105)$$

ΔE_{DKH} is a correction for scalar relativistic effects obtained from MP2 calculations using the DKH Hamiltonian,

$$\begin{aligned} \Delta E_{\text{DKH}} = & E(\text{MP2}/\text{cc-pVTZ-DK}) - E(\text{MP2}/\text{cc-pVTZ}), \end{aligned} \quad (106)$$

and ΔE_{CV} is a correction for core–valence correlation effects,

$$\Delta E_{CV} = E(\text{MP2}(\text{full})/\text{aug-cc-pCVTZ}) - E(\text{MP2}/\text{aug-cc-pVTZ}). \quad (107)$$

ΔE_{ZPVE} is calculated from the scaled B3LYP/cc-pVTZ harmonic frequencies and, for atomic species (and some linear molecules), a spin–orbit correction is taken into account.

The ccCA-P and ccCA-S4 methods appear to perform similarly to the G4 model chemistry – that is, with a mean absolute deviation of about 3.4 kJ mol^{-1} , but they are free of empirical parameters (except for the scaling of the B3LYP frequencies). The correlation-consistent composite approach has been applied to alkali and alkaline-earth metal oxides and hydroxides [288], s-block thermochemistry [288,289], transition-metal compounds [290], reaction barriers [291] and molecules with up to third-row elements [292]. Since the most time-consuming calculations are the MP2/aug-cc-pVQZ, MP2(full)/aug-cc-pCVTZ and CCSD(T)/cc-pVTZ calculations, the ccCA scheme could be applied to molecules as large as tetranitromethane, $\text{C}(\text{NO}_2)_4$, and adamantane, $\text{C}_{10}\text{H}_{16}$.

6.3. CBS methods

The complete-basis-set (CBS) methods employ non-linear pair NO extrapolations to the complete basis-set limit. These methods come in various flavours such as CBS-4 (and CBS-4M), CBS-q, CBS-Q (and CBS-QB3) and CBS-APNO. See, for example, [266] and [267].

The CBS models CBS-4, CBS-q, CBS-Q, and CBS-APNO are unique in their use of the N^{-1} asymptotic convergence of second-order Møller–Plesset pair energies, calculated from pair NO expansions to extrapolate to the complete basis-set limit. This N^{-1} asymptotic convergence is related to the X^{-3} convergence of the principal expansion since N increases as X^3 – see Section 3. Another unique aspect of the CBS methods is that the extra MP2 correlation energy obtained by extrapolating to the basis-set limit is scaled with an ‘interference factor’ to account for the fact that the MP2 energies converge more slowly to the basis-set limit than the correlation energies at higher levels such as the QCI level. For a discussion of the interference factor, see also [25].

The CBS-4, CBS-q, and CBS-Q models were found to have mean absolute deviations of 8.4, 7.1, and 4.2 kJ mol^{-1} , respectively, for the G2 test set. Today, the CBS-QB3 model developed in [266] is very popular. This is a modification of the CBS-Q model, using

B3LYP geometries and frequencies in place of Hartree–Fock/MP2 data and the CCSD(T) method in place of the QCISD(T) method. The CBS-QB3 model also contains empirical energy terms and parameters as well as a spin–orbit correction. The CBS-APNO method, which is more sophisticated and expensive than the other CBS methods, has a mean absolute deviation of about 2 kJ mol^{-1} for the 64 first-row examples in the G2 test set [266]. The CBS-4M method is less accurate than CBS-QB3 method but is less expensive and can be used for calculations on larger systems.

6.4. Focal-point analysis

The philosophy of the focal-point analysis of Allen and co-workers is to systematically explore the convergence of *ab initio* predictions to the one- and n -particle limits. Electron-correlation treatments as extensive as fifth-order Møller–Plesset perturbation theory (MP5), the full CCSDT method, and Brueckner-doubles theory including perturbational corrections for triple and quadruple excitations, BD(TQ), were used for example in [269,295]. Also CI sequences such as CISD, CISDT, CISDTQ, CISDTQ5 up to the FCI limit have been employed [295]. Furthermore, very large correlation-consistent basis sets are used in conjunction with basis-set extrapolation, and ZPVEs, DBOCs and corrections for relativistic effects and core–valence correlation are taken into account.

A recent example of a focal-point analysis is given by the work of Wheeler *et al.* [188] on the thermochemistry of C_3H_3 isomers. Geometries were optimised at the frozen-core restricted open-shell Hartree–Fock (ROHF)-CCSD(T)/TZ(2d1f,2p1d) level. Hartree–Fock, MP2, CCSD and CCSD(T) calculations were performed in basis sets ranging from cc-pVDZ to cc-pV6Z, and the individual increments of the n -particle hierarchy were extrapolated to the basis-set limit using the $a + b\exp(-cX)$ and $a + bX^{-3}$ extrapolation formulae for Hartree–Fock and correlation energies, respectively, if at least the cc-pVQZ basis was affordable. CCSDT and CCSDT(Q) increments were computed in the cc-pVTZ and cc-pVDZ basis sets, respectively, and anharmonic ZPVEs (including the G_0 term), DBOCs (at the CISD level) and corrections for relativistic effects and core–valence correlation were added.

This focal-point analysis is similar to other high-accuracy composite methods such as W4 and HEAT (*vide infra*), but the emphasis is to push basis-sets and methods for a given problem to their limits rather than to define a model chemistry with prescribed methods and basis sets.

6.5. Multicoefficient correlation methods

The multicoefficient correlation methods of Truhlar and co-workers [270–272, 296–299] are composite methods in the sense that the total electronic energy is calculated as a sum of various contributions. The authors have noted that improved results can be obtained by introducing empirical scaling factors (coefficients) for the additive terms in a given composite method. For example, in the multicoefficient G3 method (MCG3), the various correlation (Hartree–Fock → MP2 → MP4(SDQ) → MP4 → QCISD(T)) and basis-set increments are scaled with optimised coefficients. Composite approaches clearly offer many opportunities for introducing empirical scaling parameters and a large variety of multicoefficient correlation methods (MCCMs) have been developed. Recently, the performance of fifteen MCCMs was assessed by Dahlke *et al.* [299].

The MCCMs have not only been designed for thermochemistry but also for thermochemical kinetics, empirically mixing wave-function correlation methods and DFT [297, 298]. A similar approach was taken by Schwabe and Grimme in combining DFT with MP2 theory [300]. Moreover, it was soon recognised by Truhlar and co-workers that not only single-point energies can be calculated in a composite approach but that also equilibrium geometries can be optimised in such a scheme [301]. A similar multilevel geometry optimisation at the coupled-cluster level (including quadruple excitations) has been published by Heckert *et al.* [302].

6.6. HEAT protocol

Recently, a theoretical model chemistry known as ‘high-accuracy extrapolated *ab-initio* thermochemistry’ (HEAT) has been developed by Tajti *et al.* [190] and Bomble *et al.* [189]. According to the most recent HEAT protocol, total energies are computed as a sum of eight terms,

$$E_{\text{HEAT}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT}} + \Delta E_{\text{HLC}} + \Delta E_{\text{rel}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{SO}} + \Delta E_{\text{ZPVE}}. \quad (108)$$

The E_{HF}^{∞} and $\Delta E_{\text{CCSD(T)}}^{\infty}$ terms are Hartree–Fock and CCSD(T) energies that are extrapolated to the limit of a complete basis set by exploiting the approximate exponential ($\exp(-bX)$) and inverse-cubic (X^{-3}) convergence of the aug-cc-pCVXZ correlation-consistent basis sets in Hartree–Fock and coupled-cluster theories, respectively. Usually, the basis sets with $X=3, 4$ and 5 are used, leading to the HEAT345 model chemistry,

but also basis sets up to $X=6$ have been used (HEAT456 method). Core orbitals are included in the CCSD(T) calculations but not in the CCSDT calculations that define the E_{CCSDT} increment (extrapolated from cc-pVTZ and cc-pVQZ energies). The ‘higher-level correction’ E_{HLC} is obtained from CCSDT(Q), CCSDTQ or CCSDTQ5 calculations in a cc-pVDZ basis. The corresponding protocols are denoted HEAT345-(Q), HEAT345-Q and HEAT345-QP (with P for pentuple excitations). Scalar relativistic corrections include the mass–velocity and one- and two-electron Darwin terms (ΔE_{rel}) and the DBOC is computed at the Hartree–Fock level. ΔE_{SO} is calculated using a spin–orbit CI program and relativistic effective core potentials. Anharmonic corrections and the G_0 term are included in the zero-point vibrational energy ΔE_{ZPVE} . All geometries are optimised at the ae-CCSD(T)/cc-pVQZ level, using restricted and unrestricted Hartree–Fock reference wave functions for closed-shell systems and radicals, respectively.

The HEAT protocol is a very expensive method (in terms of computer resources) but also very accurate. The HEAT345-QP method has been reported to have root-mean-square and maximum errors of about 0.3 and 0.5 kJ mol⁻¹ [189]. Today, the HEAT345-(Q) protocol can be applied to systems such as vinyl chloride (C₂H₃Cl), in this particular case optimising the geometry at the CCSD(T) level in the cc-pCVQZ basis to account for the presence of a second-row atom. For this molecule, the protocol is able to predict a standard heat of formation of $\Delta_f H^0$ (298.15 K) = 20.9 ± 2 kJ mol⁻¹, while the experimental data scatter between 21.0 and 38.1 kJ mol⁻¹ [303].

6.7. Weizmann-4 theory

In the latest version (W4.4) of the highly accurate Weizmann-4 theory [201], the total molecular energy is computed as

$$E_{\text{W4.4}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT}} + \Delta E_{\hat{T}_4} + \Delta E_{\hat{T}_5} + \Delta E_{\text{CV}} + \Delta E_{\text{DKH}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{SO}} + \Delta E_{\text{ZPVE}}. \quad (109)$$

Omitting the details concerning open-shell systems and unrestricted versus restricted reference wave functions, the individual terms are computed in the following manner. The Hartree–Fock basis-set limit is obtained from an extrapolation based on the aug-cc-pV5Z and aug-cc-pV6Z basis sets and the formula [304–307]

$$E_{\text{HF}}(X) = A + B(X+1)\exp(-9X^{1/2}). \quad (110)$$

The CCSD correlation energy $\Delta E_{\text{CCSD}}^{\infty}$ is also extrapolated from the aug-cc-pV5Z and aug-cc-pV6Z basis-set results, using Equation (30). The term $\Delta E_{\text{CCSD}}^{\infty}(T)$ is either (variant W4.4a) extrapolated from the same aug-cc-pV5Z and aug-cc-pV6Z basis set with $\alpha = 3$ or by means of Schwenke's extrapolation (variant W4.4b, see also [308]). The full-triples term ΔE_{CCSDT} is extrapolated from the cc-pVDZ and cc-pVTZ basis sets using

$$\Delta E_{\text{CCSDT}}(X) = A + B/X^{2.5}, \quad (111)$$

based on the observations made in [201]. The contribution due to quadruple excitations is computed as

$$\Delta E_{\hat{T}_4} = 1.1(\Delta E_{\text{CCSDTQ}} + \Delta E_{\text{CCSDT(Q)}}), \quad (112)$$

where ΔE_{CCSDTQ} is the difference between CCSDTQ and CCSDT(Q) in the cc-pVDZ basis and $\Delta E_{\text{CCSDT(Q)}}$ the difference between CCSDT(Q) and CCSDT in the cc-pVTZ basis. $\Delta E_{\hat{T}_3}$ is obtained in the cc-pVDZ basis and the core-valence term ΔE_{CV} contains two contributions: X^{-3} -extrapolated CCSD(T)/aug-cc-pwCV(TQ)Z plus a CCSDT(Q)/cc-pwCVTZ correction. ΔE_{DKH} is obtained at the CCSD(T) level in a quadruple-zeta basis using the DKH Hamiltonian and ΔE_{DBOC} is computed at the CISD/cc-pVDZ level.

For a set of 18 molecules including C_2H_2 , SO, H_2S and Cl_2 , the root-mean-square deviation from the ATcT reference values was as small as 0.2 kJ mol^{-1} , implying a 95% confidence interval of about 0.5 kJ mol^{-1} . The target accuracy of ' $3\sigma \leq 1 \text{ kJ mol}^{-1}$ ' had thus been reached. Remaining errors, in particular for the second-row systems, could be due to the fact that the reference geometries were computed at the fc-CCSD(T)/cc-pV(Q+d)Z level – and not at the ae-CCSD(T) level – or due to the neglect of high-order relativistic corrections such as the second-order spin-orbit coupling.

7. Concluding remarks

Through the development of coupled-cluster methods, correlation-consistent basis sets, extrapolation techniques, explicitly correlated methods and additivity schemes, computational chemistry is today useful for predicting and testing experimental data for many molecular properties. Near basis-set limit CCSD(T) energies are usually sufficient to give agreement with experiment to within a few kJ mol^{-1} , and in this respect, the recent explicitly correlated CCSD(T)(F12) method is likely to find a wide range of application since it removes the necessity for large orbital basis sets. Analytic derivatives for this method have the same structure as for the conventional CCSD(T)

method and, once these have been implemented, highly accurate properties will also become available. The Dunning basis sets are far from optimal for use with explicitly correlated methods and new correlation-consistent basis sets optimised specifically for F12 methods are being developed, further enhancing the advantages of this approach [309].

For agreement within the experimental error bars, sub- kJ mol^{-1} accuracy is often required. This level of precision is extremely difficult to reach since the basis-set limit of the CCSD(T) model is insufficient and many expensive small terms must be accounted for. The most expensive of the terms that enter the composite methods for high accuracy are the treatment of CCSDT(Q) correlation, and higher-level corrections, particularly because the convergence of these terms with the orbital basis size is also very slow. The extension of the R12 methodology to correlate explicitly the triples and higher excitations would alleviate this problem, just as it does for the doubles excitations, and such methods are already being developed. Until now, explicitly correlated methods have not been used in general composite methods, although they have been used in additivity approximations for specific investigations. Since the CCSD(T) (F12) methods are now a cheap alternative to CCSD(T) calculations in large basis sets, it is likely that these and related approaches will be incorporated in composite model chemistries in the future, extending of the range of molecules that may be treated accurately.

Acknowledgements

This work was supported by the Norwegian Research Council through the CeO Centre for Theoretical and Computational Chemistry (Grant No. 179568/V30) and by the Deutsche Forschungsgemeinschaft (KL 721/2-2 and TE 644/1-1). The authors would like to thank Pål Dahle for comments.

References

- [1] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
- [2] R.J. Bartlett, in *Modern Electronic-Structure Theory*, edited by D.R. Yarkony (World Scientific, Singapore, 1995), p. 1047.
- [3] T.H. Dunning Jr, K.A. Peterson, and D.E. Woon, in *Encyclopedia of Computational Chemistry*, edited by P.v.R. Schleyer, *et al.* (Wiley, New York, 1998), Vol. 1, p. 88.
- [4] G.D. Purvis III and R.J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).

- [5] J. Noga and R.J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987).
- [6] G.E. Scuseria and H.F. Schaefer III, *Chem. Phys. Lett.* **152**, 382 (1988).
- [7] J.A. Pople, M. Head-Gordon, and K. Raghavachari, *J. Chem. Phys.* **87**, 5968 (1987).
- [8] R.J. Bartlett and D.M. Silver, *Phys. Rev.* **A10**, 1927 (1974).
- [9] J.S. Binkley and J.A. Pople, *Int. J. Quantum Chem.* **9**, 229 (1975).
- [10] K. Raghavachari, G.W. Trucks, J.A. Pople, *et al.*, *Chem. Phys. Lett.* **157**, 479 (1989).
- [11] J.F. Stanton and J. Gauss, *J. Chem. Phys.* **103**, 1064 (1995).
- [12] J.F. Stanton and J. Gauss, *Theoret. Chim. Acta* **93**, 303 (1996).
- [13] J.F. Stanton and J. Gauss, *Theoret. Chim. Acta* **95**, 97 (1997).
- [14] T. Shiozaki, K. Hirao, and S. Hirata, *J. Chem. Phys.* **126**, 244106 (2007).
- [15] O. Christiansen, H. Koch, and P. Jørgensen, *Chem. Phys. Lett.* **243**, 409 (1995).
- [16] H. Koch, O. Christiansen, P. Jørgensen, *et al.*, *J. Chem. Phys.* **106**, 1808 (1997).
- [17] T. Helgaker, T.A. Ruden, P. Jørgensen, *et al.*, *J. Phys. Organic Chem.* **17**, 913 (2004).
- [18] C. Schwartz, *Phys. Rev.* **126**, 1015 (1962).
- [19] E.A. Hylleraas, *Z. Phys.* **65**, 209 (1930).
- [20] B. Klahn and J.D. Morgan III, *J. Chem. Phys.* **81**, 410 (1984).
- [21] R.N. Hill, *J. Chem. Phys.* **83**, 1173 (1985).
- [22] W. Kutzelnigg and J.D. Morgan III, *J. Chem. Phys.* **96**, 4484 (1992).
- [23] P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955).
- [24] D.P. Carroll, H.J. Silverstone, and R.M. Metzger, *J. Chem. Phys.* **71**, 4142 (1979).
- [25] W. Klopper, K.L. Bak, P. Jørgensen, *et al.*, *J. Phys. B* **32**, R103 (1999).
- [26] W. Kutzelnigg, *Phys. Chem. Chem. Phys.* **10**, 3460 (2008).
- [27] M.R. Nyden and G.A. Petterson, *J. Chem. Phys.* **75**, 1843 (1981).
- [28] G.A. Petterson and M.R. Nyden, *J. Chem. Phys.* **75**, 3423 (1981).
- [29] J. Almlöf and P.R. Taylor, *J. Chem. Phys.* **86**, 4070 (1987).
- [30] J. Almlöf and P.R. Taylor, *Adv. Quantum Chem.* **22**, 301 (1991).
- [31] T.H. Dunning Jr, *J. Chem. Phys.* **90**, 1007 (1989).
- [32] D.E. Woon and T.H. Dunning Jr, *J. Chem. Phys.* **98**, 1358 (1993).
- [33] D.E. Woon and T.H. Dunning Jr, *J. Chem. Phys.* **103**, 4572 (1995).
- [34] R.A. Kendall, T.H. Dunning Jr, and R.J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [35] T. Helgaker, W. Klopper, H. Koch, *et al.*, *J. Chem. Phys.* **106**, 9639 (1997).
- [36] A. Halkier, T. Helgaker, P. Jørgensen, *et al.*, *Chem. Phys. Lett.* **286**, 243 (1998).
- [37] W. Klopper, *Mol. Phys.* **99**, 481 (2001).
- [38] D. Bakowies, *J. Chem. Phys.* **127**, 084105 (2007).
- [39] D. Bakowies, *J. Chem. Phys.* **127**, 164109 (2007).
- [40] H.-J. Werner and W. Meyer, *Mol. Phys.* **31**, 855 (1976).
- [41] J. Rychlewski, *Explicitly Correlated Wave Functions in Chemistry and Physics Theory and Applications* (Kluwer Academic Publishers, Dordrecht, 2003).
- [42] H.M. James and A.S. Coolidge, *J. Chem. Phys.* **1**, 825 (1933).
- [43] R.D. Present, *J. Chem. Phys.* **3**, 122 (1935).
- [44] A.S. Coolidge, H.M. James, and R.D. Present, *J. Chem. Phys.* **4**, 193 (1936).
- [45] A.S. Coolidge and H.M. James, *J. Chem. Phys.* **6**, 730 (1938).
- [46] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **43**, 2429 (1965).
- [47] D.C. Clary, *Mol. Phys.* **34**, 793 (1977).
- [48] D.C. Clary and N.C. Handy, *Chem. Phys. Lett.* **51**, 483 (1977).
- [49] S.F. Boys, *Proc. R. Soc. Ser. A* **258**, 402 (1960).
- [50] K. Singer, *Proc. R. Soc. Ser. A* **258**, 412 (1960).
- [51] K. Singer and J.V.L. Longstaff, *Proc. R. Soc. Ser. A* **258**, 421 (1960).
- [52] S.F. Boys, *Proc. R. Soc. Ser. A* **200**, 542 (1950).
- [53] W.A. Lester Jr and M. Krauss, *J. Chem. Phys.* **41**, 1407 (1964).
- [54] W.A. Lester Jr and M. Krauss, *J. Chem. Phys.* **44**, 207 (1966).
- [55] N.C. Handy, *Mol. Phys.* **26**, 169 (1973).
- [56] K.M. Karunakaran and R.E. Christoffersen, *J. Chem. Phys.* **62**, 1992 (1975).
- [57] J. Rychlewski, W. Cencek, and J. Komasa, *Chem. Phys. Lett.* **229**, 657 (1994).
- [58] J.O. Hirschfelder, *J. Chem. Phys.* **39**, 3145 (1963).
- [59] K. Jankowski, *Acta Phys. Pol.* **421**, XXXII (1967).
- [60] K. Jankowski, *Acta Phys. Pol. A* **669**, 37 (1970).
- [61] S.F. Boys and N.C. Handy, *Proc. R. Soc. Ser. A* **309**, 209 (1969).
- [62] S.F. Boys and N.C. Handy, *Proc. R. Soc. Ser. A* **310**, 43 (1969).
- [63] S.F. Boys and N.C. Handy, *Proc. R. Soc. Ser. A* **310**, 63 (1969).
- [64] S.F. Boys and N.C. Handy, *Proc. R. Soc. Ser. A* **311**, 309 (1969).
- [65] N.C. Handy, *Mol. Phys.* **21**, 817 (1971).
- [66] N.C. Handy, *Mol. Phys.* **23**, 1 (1972).
- [67] O. Sinanoglu, *Adv. Chem. Phys.* **VI**, 315 (1964).
- [68] H.F. King, *J. Chem. Phys.* **46**, 705 (1967).
- [69] K.-C. Pan and H.F. King, *J. Chem. Phys.* **53**, 4397 (1970).
- [70] K.-C. Pan and H.F. King, *J. Chem. Phys.* **56**, 4667 (1972).
- [71] L. Adamowicz and A.J. Sadlej, *J. Chem. Phys.* **67**, 4298 (1977).
- [72] L. Adamowicz and A.J. Sadlej, *J. Chem. Phys.* **69**, 3992 (1978).
- [73] K. Szalewicz, B. Jeziorski, H.J. Monkhorst, *et al.*, *Chem. Phys. Lett.* **91**, 169 (1982).

- [74] K. Szalewicz, B. Jeziorski, H.J. Monkhorst, *et al.*, J. Chem. Phys. **78**, 1420 (1983).
- [75] K. Szalewicz, B. Jeziorski, H.J. Monkhorst, *et al.*, J. Chem. Phys. **79**, 5543 (1983).
- [76] B. Jeziorski, H.J. Monkhorst, K. Szalewicz, *et al.*, J. Chem. Phys. **81**, 368 (1984).
- [77] K. Szalewicz, J.G. Zabolitzky, B. Jeziorski, *et al.*, J. Chem. Phys. **81**, 2723 (1984).
- [78] K.B. Wenzel, J.G. Zabolitzky, K. Szalewicz, *et al.*, J. Chem. Phys. **85**, 3964 (1986).
- [79] S.A. Alexander, H.J. Monkhorst, and K. Szalewicz, J. Chem. Phys. **85**, 5821 (1986).
- [80] S.A. Alexander, H.J. Monkhorst, and K. Szalewicz, J. Chem. Phys. **87**, 3976 (1987).
- [81] S.A. Alexander, H.J. Monkhorst, and K. Szalewicz, J. Chem. Phys. **89**, 355 (1988).
- [82] R. Bukowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **108**, 7946 (1998).
- [83] R. Bukowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **110**, 4165 (1999).
- [84] R. Bukowski, B. Jeziorski, S. Rybak, *et al.*, J. Chem. Phys. **102**, 888 (1995).
- [85] W. Cencek, M. Jeziorska, R. Bukowski, *et al.*, J. Phys. Chem. A **108**, 3211 (2004).
- [86] K. Patkowski, W. Cencek, M. Jeziorska, *et al.*, J. Phys. Chem. A **111**, 7611 (2007).
- [87] W. Kutzelnigg, Theoret. Chem. Accounts: Theory Computat. Model. (Theoret. Chim. Acta) **68**, 445 (1985).
- [88] W. Klopper and W. Kutzelnigg, Chem. Phys. Lett. **134**, 17 (1987).
- [89] W. Klopper and J. Almlöf, J. Chem. Phys. **99**, 5167 (1993).
- [90] W. Klopper, H.P. Lüthi, T. Brupbacher, *et al.*, J. Chem. Phys. **101**, 9747 (1994).
- [91] W. Klopper and H.P. Lüthi, Chem. Phys. Lett. **262**, 546 (1996).
- [92] W. Klopper, R. Röhse, and W. Kutzelnigg, Chem. Phys. Lett. **178**, 455 (1991).
- [93] J. Noga, W. Kutzelnigg, and W. Klopper, Chem. Phys. Lett. **199**, 497 (1992).
- [94] R.J. Gdanitz, Chem. Phys. Lett. **210**, 253 (1993).
- [95] W. Klopper and J. Noga, J. Chem. Phys. **103**, 6127 (1995).
- [96] J. Noga and P. Valiron, Chem. Phys. Lett. **324**, 166 (2000).
- [97] H. Müller, W. Kutzelnigg, J. Noga, *et al.*, J. Chem. Phys. **106**, 1863 (1997).
- [98] H. Müller and W. Kutzelnigg, Phys. Chem. Chem. Phys. **2**, 2061 (2000).
- [99] D. Tunega, J. Noga, and W. Klopper, Chem. Phys. Lett. **269**, 435 (1997).
- [100] R. Franke, H. Müller, and J. Noga, J. Chem. Phys. **114**, 7746 (2001).
- [101] J. Noga, P. Valiron, and W. Klopper, J. Chem. Phys. **115**, 2022 (2001).
- [102] M. Heckert, M. Kállay, D.P. Tew, *et al.*, J. Chem. Phys. **125**, 044108 (2006).
- [103] D.P. Tew, W. Klopper, M. Heckert, *et al.*, J. Phys. Chem. A **111**, 11242 (2007).
- [104] J. Aguilera-Iparraguirre, A.D. Boese, W. Klopper, *et al.*, Chem. Phys. **346**, 56 (2008).
- [105] W. Cencek and J. Rychlewski, J. Chem. Phys. **98**, 1252 (1993).
- [106] W. Cencek, J. Komasa, and J. Rychlewski, Chem. Phys. Lett. **246**, 417 (1995).
- [107] W. Cencek and J. Rychlewski, J. Chem. Phys. **102**, 2533 (1995).
- [108] J. Komasa and J. Rychlewski, Mol. Phys. **91**, 909 (1997).
- [109] W. Cencek and J. Rychlewski, Chem. Phys. Lett. **320**, 549 (2000).
- [110] J. Komasa, J. Rychlewski, and K. Jankowski, Phys. Rev. A **65**, 042507 (2002).
- [111] J. Komasa and J. Rychlewski, Chem. Phys. Lett. **342**, 185 (2001).
- [112] W. Cencek and W. Kutzelnigg, J. Chem. Phys. **105**, 5878 (1996).
- [113] W. Cencek, J. Rychlewski, R. Jaquet, *et al.*, J. Chem. Phys. **108**, 2831 (1998).
- [114] K. Pachucki and J. Komasa, J. Chem. Phys. **125**, 204304 (2006).
- [115] B.J. Persson and P.R. Taylor, J. Chem. Phys. **105**, 5915 (1996).
- [116] P. Dahle, dr.philos thesis, University of Oslo, May 2004.
- [117] P. Dahle, T. Helgaker, D. Jonsson, *et al.*, Phys. Chem. Chem. Phys. **9**, 3112 (2007).
- [118] R. Polly, H.-J. Werner, P. Dahle, *et al.*, J. Chem. Phys. **124**, 234107 (2006).
- [119] P. Dahle, T. Helgaker, D. Jonsson, *et al.*, Phys. Chem. Chem. Phys. **10**, 3377 (2008).
- [120] M. Nooijen and R.J. Bartlett, J. Chem. Phys. **109**, 8232 (1998).
- [121] S. Ten-no, Chem. Phys. Lett. **330**, 169 (2000).
- [122] S. Ten-no, Chem. Phys. Lett. **330**, 175 (2000).
- [123] Y. Imamura and G.E. Scuseria, J. Chem. Phys. **118**, 2464 (2003).
- [124] N. Umezawa and T. Chikyow, Int. J. Quantum Chem. **106**, 1477 (2006).
- [125] W. Klopper and C.C.M. Samson, J. Chem. Phys. **116**, 6397 (2002).
- [126] F.R. Manby, J. Chem. Phys. **119**, 4607 (2003).
- [127] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
- [128] A.J. May, E. Valeev, R. Polly, *et al.*, Phys. Chem. Chem. Phys. **7**, 2710 (2005).
- [129] D.P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
- [130] E.F. Valeev, J. Chem. Phys. **125**, 244106 (2006).
- [131] H.-J. Werner, T.B. Adler, and F.R. Manby, J. Chem. Phys. **126**, 164102 (2007).
- [132] H. Flieggl, C. Hättig, and W. Klopper, J. Chem. Phys. **124**, 044112 (2006).
- [133] C. Neiss, C. Hättig, and W. Klopper, J. Chem. Phys. **125**, 064111 (2006).

- [134] H. Fliegl, W. Klopper, and C. Hättig, *J. Chem. Phys.* **122**, 084107 (2005).
- [135] C. Neiss and C. Hättig, *J. Chem. Phys.* **126**, 154101 (2007).
- [136] H. Fliegl, C. Hättig, and W. Klopper, *Int. J. Quantum Chem.* **106**, 2306 (2006).
- [137] D.P. Tew, W. Klopper, C. Neiss, *et al.*, *Phys. Chem. Chem. Phys.* **9**, 1921 (2007).
- [138] D.P. Tew, W. Klopper, and C. Hättig, *Chem. Phys. Lett.* **452**, 326 (2008).
- [139] E.F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 106 (2008).
- [140] T.B. Adler, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **127**, 221106 (2007).
- [141] C.L. Pekeris, *Phys. Rev.* **112**, 1649 (1958).
- [142] T. Kinoshita, *Phys. Rev.* **105**, 1490 (1957).
- [143] A.J. Thakkar and T. Koga, *Phys. Rev. A* **50**, 854 (1994).
- [144] K. Frankowski and C.L. Pekeris, *Phys. Rev.* **146**, 46 (1966).
- [145] C. Schwartz, *Int. J. Mod. Phys. E* **15**, 887 (2006).
- [146] H. Nakashima and H. Nakatsuji, *J. Chem. Phys.* **127**, 224104 (2007).
- [147] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, *et al.*, *Commun. Math. Phys.* **255**, 183 (2005).
- [148] E.A. Burke, *Phys. Rev.* **130**, 1871 (1963).
- [149] L. Wolniewicz, *J. Chem. Phys.* **103**, 1792 (1995).
- [150] Z.-C. Yan, M. Tambasco, and G.W.F. Drake, *Phys. Rev. A* **57**, 1652 (1998).
- [151] J.S. Sims and S. Hagstrom, *Phys. Rev. A* **4**, 908 (1971).
- [152] E. Clementi, G. Corongiu, D. Estrin, *et al.*, *Int. J. Quantum Chem.* **45**, 511 (1993).
- [153] J. Komasa, *Chem. Phys. Lett.* **363**, 307 (2002).
- [154] J. Komasa, *J. Chem. Phys.* **115**, 158 (2001).
- [155] O. Sinanoglu, *J. Chem. Phys.* **36**, 3198 (1962).
- [156] B. Jeziorski, R. Bukowski, and K. Szalewicz, *Int. J. Quantum Chem.* **61**, 769 (1997).
- [157] D.P. Tew, W. Klopper, and F.R. Manby, *J. Chem. Phys.* **127**, 174105 (2007).
- [158] E.F. Valeev, *Chem. Phys. Lett.* **395**, 190 (2004).
- [159] P. Wind, W. Klopper, and T. Helgaker, *Theoret. Chim. Acc.* **107**, 173 (2002).
- [160] S. Kedžuch, M. Milko, and J. Noga, *Int. J. Quantum Chem.* **105**, 929 (2005).
- [161] D.P. Tew and W. Klopper, *J. Chem. Phys.* **125**, 094302 (2006).
- [162] S. Ten-no, *J. Chem. Phys.* **121**, 117 (2004).
- [163] D. Yamaki, H. Koch, and S. Ten-no, *J. Chem. Phys.* **127**, 144104 (2007).
- [164] W. Klopper, *J. Chem. Phys.* **102**, 6168 (1995).
- [165] J. Noga, S. Kedžuch, and J. Šimunek, *J. Chem. Phys.* **127**, 034106 (2007).
- [166] C. Villani and W. Klopper, *J. Phys. B: At. Mol. Opt. Phys.* **38**, 2555 (2005).
- [167] H.-J. Werner and F.R. Manby, *J. Chem. Phys.* **124**, 054114 (2006).
- [168] F.R. Manby, H.-J. Werner, T.B. Adler, *et al.*, *J. Chem. Phys.* **124**, 094103 (2006).
- [169] R.A. Bachorz, W. Klopper, and M. Gutowski, *J. Chem. Phys.* **126**, 085101 (2007).
- [170] S. Ten-no, *Chem. Phys. Lett.* **447**, 175 (2007).
- [171] G. Knizia and H.-J. Werner, *J. Chem. Phys.* **128**, 154103 (2008).
- [172] E. Kordel, C. Villani, and W. Klopper, *J. Chem. Phys.* **122**, 214306 (2005).
- [173] E. Kordel, C. Villani, and W. Klopper, *Mol. Phys.* **105**, 2565 (2007).
- [174] J. Noga and W. Kutzelnigg, *J. Chem. Phys.* **101**, 7738 (1994).
- [175] B. Ruscic, R.E. Pinzon, M.L. Morton, *et al.*, *J. Phys. Chem. A* **108**, 9979 (2004).
- [176] B. Ruscic, R.E. Pinzon, M.L. Morton, *et al.*, *J. Phys. Chem. A* **110**, 6592 (2006).
- [177] M.W. Chase Jr, editor, *NIST-JANAF Thermochemical Tables*, 4th ed. (ACS, Washington DC/AIP, New York, 1998).
- [178] T.A. Ruden, T. Helgaker, P. Jørgensen, *et al.*, *J. Chem. Phys.* **121**, 5874 (2004).
- [179] K.P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
- [180] H. Müller, R. Franke, S. Vogtner, *et al.*, *Theoret. Chim. Acc.* **100**, 85 (1998).
- [181] G. Tasi, R. Izsák, G. Matisz, *et al.*, *ChemPhysChem* **7**, 1664 (2006).
- [182] T.A. Ruden, P.R. Taylor, and T. Helgaker, *J. Chem. Phys.* **119**, 1951 (2003).
- [183] M.S. Schuurman, W.D. Allen, and H.F. Schaefer III, *J. Comput. Chem.* **26**, 1106 (2005).
- [184] M.S. Schuurman, W.D. Allen, P. von Ragué Schleyer, *et al.*, *J. Chem. Phys.* **122**, 104302 (2005).
- [185] K. Aarset, A.G. Császár, E.L. Sibert III, *et al.*, *J. Chem. Phys.* **112**, 4053 (2000).
- [186] A.D. Boese, W. Klopper, and J.M.L. Martin, *Mol. Phys.* **103**, 863 (2005).
- [187] A.D. Boese, W. Klopper, and J.M.L. Martin, *Int. J. Quantum Chem.* **104**, 830 (2005).
- [188] S.E. Wheeler, K.A. Robertson, W.D. Allen, *et al.*, *J. Phys. Chem. A* **111**, 3819 (2007).
- [189] Y.J. Bomble, J. Vázquez, M. Kállay, *et al.*, *J. Chem. Phys.* **125**, 064108 (2006).
- [190] A. Tajti, P.G. Szalay, A.G. Császár, *et al.*, *J. Chem. Phys.* **121**, 11599 (2004).
- [191] A. Karton, B. Ruscic, and J.M.L. Martin, *J. Mol. Struct.: THEOCHEM* **811**, 345 (2007).
- [192] C. Puzzarini and P.R. Taylor, *J. Chem. Phys.* **122**, 054315 (2005).
- [193] T.A. Ruden, T. Helgaker, P. Jørgensen, *et al.*, *Chem. Phys. Lett.* **371**, 62 (2003).
- [194] J.F. Stanton, *Chem. Phys. Lett.* **281**, 130 (1997).
- [195] Y.J. Bomble, J.F. Stanton, M. Kállay, *et al.*, *J. Chem. Phys.* **123**, 054101 (2005).
- [196] A.D. Boese, M. Oren, O. Atasoylu, *et al.*, *J. Chem. Phys.* **120**, 4129 (2004).

- [197] A. Karton, E. Rabinovich, J.M.L. Martin, *et al.*, J. Chem. Phys. **125**, 144108 (2006).
- [198] M. Kállay and P.R. Surján, J. Chem. Phys. **115**, 2945 (2001).
- [199] J. Olsen, unpublished.
- [200] S. Hirata, J. Phys. Chem. A **107**, 9887 (2003).
- [201] A. Karton, P.R. Taylor, and J.M.L. Martin, J. Chem. Phys. **127**, 064104 (2007).
- [202] J.M.L. Martin and O. de Oliveira, J. Chem. Phys. **111**, 1843 (1999).
- [203] R.D. Cowan and D.C. Griffin, J. Opt. Soc. Am. **66**, 1010 (1976).
- [204] A. Rutkowski, J. Phys. B: At. Mol. Opt. Phys. **19**, 149 (1986).
- [205] A. Rutkowski, J. Phys. B: At. Mol. Opt. Phys. **19**, 3431 (1986).
- [206] A. Rutkowski, J. Phys. B: At. Mol. Opt. Phys. **19**, 3443 (1986).
- [207] W. Kutzelnigg, Z. Phys. D **11**, 15 (1989).
- [208] W. Kutzelnigg, E. Ottschofski, and R. Franke, J. Chem. Phys. **102**, 1752 (1995).
- [209] E. Ottschofski and W. Kutzelnigg, J. Chem. Phys. **106**, 6634 (1997).
- [210] W. Klopper, J. Comput. Chem. **18**, 20 (1997).
- [211] W. Klopper, S. Coriani, T. Helgaker, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **37**, 3753 (2004).
- [212] G. Tarczay, A.G. Császár, W. Klopper, *et al.*, Mol. Phys. **99**, 1769 (2001).
- [213] A.C. Hennum, W. Klopper, and T. Helgaker, J. Chem. Phys. **115**, 7356 (2001).
- [214] C. Michauk and J. Gauss, J. Chem. Phys. **127**, 044106 (2007).
- [215] T. Helgaker, A.C. Hennum, and W. Klopper, J. Chem. Phys. **125**, 024102 (2006).
- [216] S. Coriani, T. Helgaker, P. Jørgensen, *et al.*, J. Chem. Phys. **121**, 6591 (2004).
- [217] M. Douglas and N.M. Kroll, Ann. Phys. (NY) **82**, 89 (1974).
- [218] B.A. Hess, Phys. Rev. A **32**, 756 (1985).
- [219] B.A. Hess, Phys. Rev. A **33**, 3742 (1986).
- [220] G. Jansen and B.A. Hess, Phys. Rev. A **39**, 6016 (1989).
- [221] C. Chang, M. Pelissier, and P. Durand, Phys. Scr. **34**, 394 (1986).
- [222] E. van Lenthe, E.J. Baerends, and J.G. Snijders, J. Chem. Phys. **99**, 4597 (1993).
- [223] E. van Lenthe, E.J. Baerends, and J.G. Snijders, J. Chem. Phys. **101**, 9783 (1994).
- [224] W. Klopper, J.H. van Lenthe, and A.C. Hennum, J. Chem. Phys. **113**, 9957 (2000).
- [225] K.G. Dyall, J. Chem. Phys. **106**, 9618 (1997).
- [226] K.G. Dyall, J. Chem. Phys. **109**, 4301 (1998).
- [227] K.G. Dyall and T. Enevoldsen, J. Chem. Phys. **111**, 10000 (1999).
- [228] K.G. Dyall, J. Chem. Phys. **115**, 9136 (2001).
- [229] W. Kutzelnigg and W. Liu, J. Chem. Phys. **123**, 241102 (2005).
- [230] W. Kutzelnigg and W. Liu, Mol. Phys. **104**, 2225 (2006).
- [231] W. Liu and W. Kutzelnigg, J. Chem. Phys. **126**, 114107 (2007).
- [232] W. Kutzelnigg and W. Liu, Chem. Phys. **349**, 133 (2008).
- [233] C. van Wüllen and C. Michauk, J. Chem. Phys. **123**, 204113 (2005).
- [234] D. Feller, K.A. Peterson, W.A. de Jong, *et al.*, J. Chem. Phys. **118**, 3510 (2003).
- [235] O.L. Polyansky, A.G. Császár, S.V. Shirin, *et al.*, Science **299**, 539 (2003).
- [236] P. Pyykkö, K.G. Dyall, A.G. Császár, *et al.*, Phys. Rev. A **63**, 024502 (2001).
- [237] K.G. Dyall, C.W. Bauschlicher Jr, D.W. Schwenke, *et al.*, Chem. Phys. Lett. **348**, 497 (2001).
- [238] P. Pyykkö, M. Tokman, and L.N. Labzowsky, Phys. Rev. A **57**, R689 (1998).
- [239] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1956).
- [240] W. Kołos and L. Wolniewicz, J. Chem. Phys. **41**, 3663 (1964).
- [241] H. Sellers and P. Pulay, Chem. Phys. Lett. **103**, 463 (1984).
- [242] H. Sellers, Chem. Phys. Lett. **109**, 339 (1984).
- [243] N.C. Handy, Y. Yamaguchi, and H.F. Schaefer III, J. Chem. Phys. **84**, 4481 (1986).
- [244] N.C. Handy and A.M. Lee, Chem. Phys. Lett. **252**, 425 (1996).
- [245] W. Kutzelnigg, Mol. Phys. **90**, 909 (1997).
- [246] J.O. Jensen and D.R. Yarkony, J. Chem. Phys. **89**, 975 (1988).
- [247] W. Cencek and W. Kutzelnigg, Chem. Phys. Lett. **266**, 383 (1997).
- [248] D. Schwenke, J. Phys. Chem. A **105**, 2352 (2001).
- [249] E.F. Valeev and C.D. Sherill, J. Chem. Phys. **118**, 3921 (2003).
- [250] S.L. Mielke, D.W. Schwenke, and K.A. Peterson, J. Chem. Phys. **122**, 224313 (2005).
- [251] M. Przybytek and B. Jeziorski, J. Chem. Phys. **123**, 134315 (2005).
- [252] J. Gauss, A. Tajti, M. Kállay, *et al.*, J. Chem. Phys. **125**, 144111 (2006).
- [253] H.-J. Werner, M. Kállay, and J. Gauss, J. Chem. Phys. **128**, 034305 (2008).
- [254] J.A. Pople, M. Head-Gordon, D.J. Fox, *et al.*, J. Chem. Phys. **90**, 5622 (1989).
- [255] L.A. Curtiss, C. Jones, G.W. Trucks, *et al.*, J. Chem. Phys. **93**, 2537 (1990).
- [256] L.A. Curtiss, K. Raghavachari, G.W. Trucks, *et al.*, J. Chem. Phys. **94**, 7221 (1991).
- [257] L.A. Curtiss, K. Raghavachari, P.C. Redfern, *et al.*, J. Chem. Phys. **109**, 7764 (1998).
- [258] L.A. Curtiss, P.C. Redfern, and K. Raghavachari, J. Chem. Phys. **126**, 084108 (2007).
- [259] N.J. DeYonker, T.R. Cundari, and A.K. Wilson, J. Chem. Phys. **124**, 114104 (2006).
- [260] N.J. DeYonker, T. Grimes, S. Yockel, *et al.*, J. Chem. Phys. **125**, 104111 (2006).

- [261] G.A. Petersson and M.A. Al-Laham, *J. Chem. Phys.* **94**, 6081 (1991).
- [262] G.A. Petersson, T.G. Tensfeldt, and J.A. Montgomery Jr, *J. Chem. Phys.* **94**, 6091 (1991).
- [263] J.A. Montgomery Jr, J.W. Ochterski, and G.A. Petersson, *J. Chem. Phys.* **101**, 5900 (1994).
- [264] J.W. Ochterski, G.A. Petersson, and J.A. Montgomery Jr, *J. Chem. Phys.* **104**, 2598 (1996).
- [265] G.A. Petersson, D.K. Malick, W.G. Wilson, *et al.*, *J. Chem. Phys.* **109**, 10570 (1998).
- [266] J.A. Montgomery Jr, M.J. Frisch, J.W. Ochterski, *et al.*, *J. Chem. Phys.* **110**, 2822 (1999).
- [267] J.A. Montgomery Jr, M.J. Frisch, J.W. Ochterski, *et al.*, *J. Chem. Phys.* **112**, 6532 (2000).
- [268] A.L.L. East and W.D. Allen, *J. Chem. Phys.* **99**, 4638 (1993).
- [269] A.G. Császár, W.D. Allen, and H.F. Schaefer III, *J. Chem. Phys.* **108**, 9751 (1998).
- [270] P.L. Fast, M.L. Sánchez, and D.G. Truhlar, *Chem. Phys. Lett.* **306**, 407 (1999).
- [271] P.L. Fast, J. Corchado, M.L. Sánchez, *et al.*, *J. Phys. Chem. A* **103**, 3139 (1999).
- [272] P.L. Fast, J.C. Corchado, M.L. Sánchez, *et al.*, *J. Phys. Chem. A* **103**, 5129 (1999).
- [273] P. Ho and C.F. Melius, *J. Phys. Chem.* **99**, 2166 (1995).
- [274] J.A. Miller and S.J. Klippenstein, *J. Phys. Chem. A* **107**, 2680 (2003).
- [275] W. Klopper and H.P. Lüthi, *Mol. Phys.* **96**, 559 (1999).
- [276] W. Klopper, J.G.C.M. van Duijneveldt-van de Rijdt, and F.B. van Duijneveldt, *Chem. Phys. Phys. Chem.* **2**, 2227 (2000).
- [277] M.S. Gordon and D.G. Truhlar, *J. Am. Chem. Soc.* **108**, 5412 (1986).
- [278] P.E.M. Siegbahn, M.R.A. Blomberg, and M. Svensson, *Chem. Phys. Lett.* **223**, 35 (1994).
- [279] D. Feller and D.A. Dixon, *J. Phys. Chem. A* **103**, 6413 (1999).
- [280] D. Feller and D.A. Dixon, *J. Phys. Chem. A* **107**, 9641 (2003).
- [281] L. Pollack, T.L. Windus, W.A. de Jong, *et al.*, *J. Phys. Chem. A* **109**, 6934 (2005).
- [282] M.T. Nguyen, M.H. Matus, W.A. Lester, *et al.*, *J. Phys. Chem. A* **112**, 2082 (2008).
- [283] D. Feller and K.A. Peterson, *J. Chem. Phys.* **108**, 154 (1998).
- [284] D. Feller and K.A. Peterson, *J. Chem. Phys.* **110**, 8384 (1999).
- [285] L. Bytautas and K. Ruedenberg, *J. Chem. Phys.* **121**, 10905 (2004).
- [286] L. Bytautas and K. Ruedenberg, *J. Chem. Phys.* **124**, 174304 (2006).
- [287] L.A. Curtiss, P.C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **123**, 124107 (2005).
- [288] D.S. Ho, N.J. DeYonker, A.K. Wilson, *et al.*, *J. Phys. Chem. A* **110**, 9767 (2006).
- [289] N.J. DeYonker, D.S. Ho, A.K. Wilson, *et al.*, *J. Phys. Chem. A* **111**, 10776 (2007).
- [290] N.J. DeYonker, K.A. Peterson, G. Steyl, *et al.*, *J. Phys. Chem. A* **111**, 11269 (2007).
- [291] T.V. Grimes, A.K. Wilson, N.J. DeYonker, *et al.*, *J. Chem. Phys.* **127**, 154117 (2007).
- [292] N.J. DeYonker, B. Mintz, T.R. Cundari, *et al.*, *J. Chem. Theory Comput.* **4**, 328 (2008).
- [293] K.A. Peterson, D.E. Woon, and T.H. Dunning Jr, *J. Chem. Phys.* **100**, 7410 (1994).
- [294] J.M.L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- [295] G. Tarczay, A.G. Császár, W. Klopper, *et al.*, *J. Chem. Phys.* **110**, 11971 (1999).
- [296] P.L. Fast, M.L. Sánchez, J.C. Corchado, *et al.*, *J. Chem. Phys.* **110**, 11679 (1999).
- [297] Y. Zhao, B.J. Lynch, and D.G. Truhlar, *J. Phys. Chem. A* **108**, 4786 (2004).
- [298] Y. Zhao, B.J. Lynch, and D.G. Truhlar, *Phys. Chem. Chem. Phys.* **7**, 43 (2005).
- [299] E.E. Dahlke, M.A. Orthmeyer, and D.G. Truhlar, *J. Phys. Chem. B* **112**, 2372 (2008).
- [300] T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.* **8**, 4398 (2006).
- [301] J.M. Rodgers, P.L. Fast, and D.G. Truhlar, *J. Chem. Phys.* **112**, 3141 (2000).
- [302] M. Heckert, M. Kállay, and J. Gauss, *Mol. Phys.* **103**, 2109 (2005).
- [303] M.E. Harding, J. Gauss, K. Pflüger, *et al.*, *J. Phys. Chem. A* **111**, 13623 (2007).
- [304] A. Karton and J.M.L. Martin, *Theor. Chem. Acc.* **115**, 330 (2006).
- [305] F. Jensen, *Theor. Chem. Acc.* **113**, 267 (2005).
- [306] W. Kutzelnigg, *Int. J. Quantum Chem.* **51**, 447 (1994).
- [307] W. Klopper and W. Kutzelnigg, *J. Mol. Struct.: THEOCHEM* **135**, 339 (1986).
- [308] D.W. Schwenke, *J. Chem. Phys.* **122**, 014107 (2005).
- [309] K.A. Peterson, T.B. Adler, and H.-J. Werner, *J. Chem. Phys.* **128**, 084102 (2008).