INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

QUANTUM CHEMISTRY IN THE 21st CENTURY

(Special Topic Article)

Prepared for publication by
CHRISTOPHER J. BARDEN AND HENRY F. SCHAEFER III

Center for Computational Quantum Chemistry, University of Georgia, Athens, Georgia 30602 USA

Quantum chemistry in the 21st century
(Special Topic Article)

Abstract: Quantum chemistry is the field in which solutions to the Schrödinger equation are used to predict the properties of molecules and solve chemical problems. This paper considers possible future research directions in light of the discipline's past successes. After decades of incremental development—accompanied by a healthy dose of skepticism from the experimental community—the ready availability of fast computers has ushered in a "golden age" of quantum chemistry. In this new era of acceptance, theoretical predictions often precede experiment in small molecule chemistry, and quantum chemical methods play an ever greater role in biochemical and other larger systems. Quantum chemists increasingly divide their efforts along three fronts: high-level (spectroscopic) accuracy for small molecules, characterized by such techniques as Brueckner methods, r_{12} formalisms, and multireference calculations; parameterization- or extrapolation-based intermediate-level schemes (such as Gaussian-N theory) for medium molecules; and lower-level (chemical) accuracy for large molecules, characterized by density functional theory and linear scaling techniques. These tools, and quantum chemistry as a whole, are examined here from a historical perspective and with a view toward their future applications.

INTRODUCTION

For the second quarter of the 20th century, quantum chemistry was in an almost prenatal stage, gestating in the womb of quantum mechanics. In 1927, only a year after the appearance of Schrödinger's much-lauded wave equation, Heitler and London published the first theoretical treatment of the hydrogen molecule [1]. The Heitler–London valence bond model, which is still used today, was followed shortly by groundwork for the Hartree–Fock approximation utilizing molecular orbitals [2,3]. Further developments had their start in this time period as well, including the Configuration Interaction [4] method, r_{12} methods [5,6], and the Perturbation Theory treatment described by Møller and Plesset [7]. There was even a paper describing a multireference calculation (a precursor to modern Multi-Configuration Self-Consistent-Field theory) on two states of the oxygen atom [8]. Nevertheless, there were many scientists who believed that a diatomic system was about as large as anyone could theoretically contemplate due to the tremendous calculational undertaking. In 1929, Paul Dirac, one of the pioneers of the quantum theory, noted his opinion on prospects for chemical computation thusly:

> The fundamental laws necessary for the mathematical treatment of large parts of physics and the whole of chemistry are thus fully known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved [9].

In the quantum chemical community, this quote has become a necessary preface to any discussion about the field; indeed, it appears near the top of the 1998 Nobel prize citation for John Pople and Walter Kohn as their contributions to quantum chemistry are narrated. It holds this position in the oral history of our field because it is a dramatic example of the overall pessimism and scepticism that has plagued quantum chemistry since the beginning, despite its many successes [10–13] and present-day respectability. An important lesson can be learned here about forecasting in a field thought to be impossible even at the time of its inception. In the better part of a century since Dirac's ill-fated pronouncement, the thing which has made quantum chemistry possible has been the improbably large improvements in computer performance. This paper is devoted to an examination of research directions which

may bear fruit, but the true advances in quantum chemistry in the next century could very well arrive from some equally unlikely external development.

THE SEGMENTED WORLD OF CHEMICAL COMPUTATION

From its primordial origins in theoretical chemistry, quantum chemistry has matured into a robust field with many different features and research directions. What was once only "theoretical chemistry" now comprises a number of different research areas: quantum chemistry, computational chemistry, molecular modeling, and even interdisciplinary specializations such as structural biology and bioinformatics. Today, the term "quantum chemistry" has a much more circumscribed use than suggested thus far in this article; indeed, many techniques for use with large molecules do not rely upon quantum mechanics at all! Other related fields are more concerned with the classification and statistical analysis of large amounts of data than with calculations of molecular properties. Such areas, while certainly important, will not be detailed here as their roots lie principally in other regimes (such as computer science or Newtonian physics). Historically, a chief tenet of quantum chemistry has been that predictions about the structure and properties of molecules ought to be based on the quantum theory and not on parameterizations, heuristics, or empirical correlations, however accurate they may appear to be. That sentiment has changed quite a bit as the problems have become more complicated, as will be discussed below. Regardless of the approximations being made now and in the future, however, there will always be a difference of perception as to what constitutes a "large molecule".

The *status quo* is an artificial barrier between various sizes of molecules and the level of detail in their structure and property predictions. This is a necessary outgrowth of the finite computational power available. While we may expect computer technology to continually improve, in the meantime we must strike a balance between accuracy of results and speed of computation. The methods currently in vogue and the nascent techniques of the future are discussed in these practical categories below.

SMALL MOLECULES

We begin by examining small molecule chemistry, for it is here that the most powerful methods of quantum chemistry can be utilized. Some background is necessary at this point. Nearly all *ab initio* quantum chemical methods attempt to solve the eigenvalue problem defined by the time-independent Schrödinger equation:

$$H|\Psi\rangle = E|\Psi\rangle \tag{1}$$

Ψ is taken to be the exact wavefunction, with two caveats: relativistic effects account for a small part of the energy, as do assumptions made by keeping nuclei fixed in space with respect to the electrons (the Born-Oppenheimer approximation) [14]. As in any eigenvalue problem, the use of an operator upon an eigenfunction yields a numerical multiple of the eigenfunction, the multiple being the eigenvalue. Here the operator is typically the molecular electronic Hamiltonian, consisting of all of the kinetic and potential energy terms that act upon a system of moving electrons and fixed nuclei:

$$\left[-\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{r_{ij}} \right] |\Psi\rangle = E|\Psi\rangle \tag{2}$$

In this equation N is the number of electrons while M is the number of nuclei. The first term represents the electronic kinetic energy, while the second and third terms are the nuclear-electronic and electronic-electronic potential energy terms, respectively. (The nuclear kinetic and nuclear-nuclear potential energy terms are added to the total energy after computing the electronic part.) It is the inter-

electronic term (called r_{12}) that makes analytical solution of the Schrödinger equation impossible for many-electron systems.

So if it cannot be solved analytically, how can anything be accomplished with this relation? The first important approximation in general use was introduced near the beginning of the era—the Hartree–Fock (HF) approximation [2,3]. In the HF approximation, the electronic Hamiltonian is replaced with an approximate Hamiltonian identical in the first two terms, but with an effective potential υ^{HF} instead of the exact r_{12} term:

$$H_{HF} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \upsilon^{HF} \tag{3}$$

The introduction of this effective potential means that each electron is subjected to a field that models the effect of the other electrons in the system. This is a reasonable first approximation to the way that electrons interact with each other, and many simple bonding theories are perfectly consistent with this view, although it would be, of course, preferable to retain the interelectronic r_{12} interactions (more on this later). At first the relation was solved in a strictly numerical fashion as a differential equation. The method was made vastly more applicable with the introduction of the Roothaan formalism [15]. In 1951, Roothaan showed that there exists a procedure by which a matrix representation of the Hartree–Fock problem may be solved iteratively, so long as a further approximation is made. Instead of calculating the orbitals as points on a finite grid, in the analytic approximation they are expanded algebraically as a finite set of *basis functions*. These basis functions introduce a degree of chemical intuition into quantum chemistry by way of the almost-universally-known orbital concept. Unfortunately, the orbital picture is only formally true in a quantum mechanical sense for one-electron systems (such as the hydrogen atom or H_2^+). Thus there are two very distinct conjectures being made in a conventional HF calculation.

From a practical standpoint, these simplifications allow an initial guess at the electron configuration of the chemical system to produce a guess at the effective potential, which can then be used as an improvement to the guess of the orbitals, etc. Once the orbitals and the potential no longer change, the iterative calculation is complete; it is said to be *self-consistent*, which is why the Hartree–Fock approximation as practiced today is generally synonymous with self-consistent field (SCF) in the literature. There is also a built-in advantage to this procedure in that the *variational theorem* will be true for this method:

$$\int \Psi_{HF}^* H_{HF} \Psi_{HF} d\tau \geq E \tag{4}$$

such that a lowering of the energy is necessarily an improvement of the electronic description. A variational method is generally advantageous due to this straightforward figure of merit.

Basis sets in quantum chemistry are generally built according to the notion that they should reflect both the chemical ideas of orbitals as well as the physical reality of long-range interactions (such as hydrogen bonding) that may be more diffuse than most conventional bonding theories would suggest. Early research showed that Slater functions describe an orbital very well, as they maintain the cusp of the electronic potential near the nucleus; but they are also computationally rather expensive. An alternative that is in general use today is to utilize sets of easier-to-calculate linear combinations of Gaussian functions rather than Slater functions (the former are then called "contracted Gaussian basis functions") [16]. Other chemical hypotheses are in effect in these calculations as well, most notably in the area of electronic spin. At first, HF/SCF calculations operated under an ironclad definition of orbital spin, restricting the energies of α and β spin orbitals to be identical as implied by most simple models of a chemical system. Later it was found that sometimes the electronic energy would improve if the α and β parts of orbitals were allowed to vary (but this approach muddles the intuitive picture of clearly defined electron configurations). This method is called the Unrestricted Hartree–Fock (UHF) method and has been in general use for about 40 years [17].

The HF/SCF method, while only a starting point, is an enormously effective one. It provides a mathematically guaranteed improvement condition via the variational principle, but it remains compu-

tationally quite efficient. Computational methods are often rated by how well they *scale*, that is, how much slower a calculation will become if its size is increased. For HF/SCF, the standard scaling is $O(N^2–N^3)$ where N is the number of electrons, which means that a system with twice as many electrons will typically take somewhere between four and eight times as long to compute. Today, it is possible to do quite large *ab initio* calculations using this method (even proteins may be tractable for isolated investigations by 2010). Nevertheless, the conceptual underpinnings of HF/SCF are both its strength and its downfall, for this chemical picture comes at the price of theoretical accuracy in many cases. Because of the central field imposed by the Hartree–Fock potential, each electron is not explicitly aware of the others' presence. It is usually said that the HF/SCF procedure does not *correlate* electrons because there is a finite probability that two electrons will occupy the same point in space. Higher levels of theory are necessary to resolve this issue by recovering the correlation energy:

$$E_{corr} = E - E_{HF} \tag{5}$$

This energy will always be negative because the Hartree–Fock energy is an upper bound to the exact energy (as guaranteed by the variational theorem). Most of the standard high-level techniques in quantum chemistry utilize the Hartree–Fock approximation as a starting point and then attempt to correlate electrons by more rigorous (hence more expensive) computational methods. The more correlation energy that is recovered, the better the result (theoretically, at least; this certainly does not mean the answer is closer to an experimental value!). In principle, also, the more complete the basis set, the better the result and the longer the computation. Combine an essentially complete basis set (the "Hartree–Fock limit") with a completely correlated electron picture (usually called full configuration interaction after the technique commonly used to obtain it) and the exact nonrelativistic, clamped-nuclei energy is obtained. From there, relativistic corrections, Born–Oppenheimer diagonal corrections, and such may be applied to find the most accurate results. Unfortunately, this ultimate solution will not be possible for more than a diatomic molecule without increases in computational power just as dramatic as those in the past half-century. Even then, the tremendous $O(N!)$ scaling of a fully correlated calculation seems to be an insurmountable difficulty for large systems.

At this point we must therefore consider what will be the best use of resources in finding the majority of the correlation energy. The two classes of solution methods applicable to this problem are *variational* and *perturbational*. Mathematically speaking, variational methods treat the exact solution as a linear combination of discrete solution sets, while perturbative methods separate the problem into an exactly soluble part and a difficult part with no general analytic solution. Since variational methods, by their very nature, provide an upper bound to the exact energy of the system, they are often easier to explain as well as evaluate. On the other hand, while not variational, perturbative methods mathematically satisfy another important principle. If we were to calculate the energy of, let us say, two water molecules separated by an infinite distance, we would formally expect no interaction between them. In any (incompletely-correlated) conventional variational method, this is unfortunately not the case. We say, then, that the principle of *size-consistency* is satisfied in a perturbative method, even though its correlation may not be complete. The most popular variational post-HF method is called configuration interaction (CI). There are two important perturbative post-HF methods: many-body perturbation theory (MBPT) and coupled cluster (CC).

Configuration interaction is a conceptually simple idea that became feasible with the construction of early computers in the 1950's [18]. It was becoming known that the Hartree–Fock approximation was qualitatively sound for most ground state properties of molecules. It remained, therefore, to include the interaction energy that was due to excited states in a deterministic fashion. The standard CI method relies on partitioning the exact wavefunction Ψ into a selected collection of smaller functions Φ_i and their corresponding coefficients c_i:

$$\Psi = \sum_i c_i \left| \Phi_i \right\rangle \tag{6}$$

Using the method of undetermined multipliers, the contributions of these other determinants (or, in some implementations, special spin-free configuration state functions) can then be determined utilizing the constraint of the already-known Hartree–Fock wavefunction. If this set is complete, it is a full CI (FCI) calculation and is both variational and size-consistent (but also extremely expensive). More realistically, this set of determinants usually includes all single and double excitations out of the reference determinant (denoted CISD). CISD has had a long history of providing a first step toward correlating the electrons; many of the studies that were "high-level" in 1980 were of this variety. Quite a lot of chemistry can be modeled satisfactorily using CISD, but nowadays it is typically augmented with triple and/or quadruple excitations in some fashion. The most popular variant is entitled quadratic configuration interaction (QCISD) [19]; it is widely used in the Gaussian-N model chemistries (see below). As a side note, it is worth mentioning that truncated CI is not size-consistent, but that an approximate size-consistency formula (called the Davidson correction) is defined for CISD [20].

Many-body perturbation theory (MBPT) was the first of the perturbative theories to evolve. Using MBPT, we wish to solve the eigenvalue problem

$$H|\Psi\rangle = H_0|\Phi_0\rangle + V|\Phi_1...\Phi_n\rangle = E|\Psi\rangle \tag{7}$$

where we know the eigenfunctions and eigenvalues of H_0

$$H_0|\Phi_0\rangle = E_0|\Phi_0\rangle \tag{8}$$

and will systematically improve the eigenfunctions and eigenvalues of H_0 such that they approach those of H. This can be accomplished by producing a set of separable equations that represent various orders of a MacLaurin series; by solving each equation the exact energy can be determined by summing the values:

$$E = E_0 + E_1 + ... + E_n \tag{9}$$

In practice, the most logical choice for H_0 is H_{HF}, and the series is usually truncated at the second order. This variant of MBPT is denoted MP2 after the Møller–Plesset formalism on which it is based [7]. The MP2 method is a well-tested quantum chemical technique and is currently most valuable for medium-accuracy calculations where the size of the molecule is a factor, since standard MP2 scales only as $O(N^5)$. MP2 (and the generally more accurate MP4) are also key components of several Gaussian-N schemes (see below). The more computationally advantageous density functional theory (DFT) may largely supplant MP2, however, due to the methods' similar overall empirical performance.

Today, by far the most popular high-level *ab initio* quantum chemical method is coupled cluster (CC) theory. Since its introduction into quantum chemistry in the late 1960s by Čížek and Paldus [21–23], coupled cluster theory has emerged as perhaps the most reliable method for the prediction of molecular properties. At first the method seemed overly abstruse to most quantum chemists, as it had its formal basis in areas of mathematics not typically studied by chemists. But Hurley presented a re-derivation of the coupled cluster doubles (CCD) equations [24] in terms which were more familiar, and soon thereafter Monkhorst [25] formulated a general coupled cluster theory for calculating molecular properties. By 1978, computer implementations of the theory for realistic systems were developed by researchers led by Pople [26] and Bartlett [27]. Later, a coupled cluster singles and doubles (CCSD) method became available [28]. Today, CCSD(T), which uses perturbation theory to compute the connected triples, is probably the most popular high-level *ab initio* method [29]. The theory is complicated and will not be completely explained here, but it begins by suggesting that the exact wavefunction may be described as

$$\Psi = e^T|\Phi_0\rangle \tag{10}$$

in which Φ_0 is a suitable reference function, usually Hartree–Fock. *T* is an excitation operator:

$$T = T_1 + T_2 + ... + T_n \tag{11}$$

This exponential form makes coupled cluster theory unique among all techniques in quantum chemistry. Unlike a CI calculation, in which the terms "single" and "double" apply to excitations out of the reference function, in a CC calculation those terms refer to all interactions associated with that quantity of particle (i.e., one- and two-body interactions). Armed with this powerful technique, something like 97% of the correlation energy can be extracted from just the first three terms in *T*. All this power is computationally expensive, though; CCSD(T) scales as $O(N^7)$, meaning that doubling the size of the chemical system results in *128 times* the computer time required in a straightforward implementation.

The accuracy that can be expected in a large-basis CCSD(T) computation on a small molecule is very good indeed; for a state-of-the-art example by Martin and Taylor, see Fig. 1 [30]. Their study of N_2H_2 utilized basis sets employing *h* orbitals, careful evaluation of core-valence effects, and anharmonic contributions to the zero-point energy. As a result, rotational constants were obtained to within better than 0.1% of experiment. Studies like theirs can be expected to generate relative energies to within 0.2 kcal/mol of experiment, which in many cases is probably more authoritative than experiment itself! Such studies are very common on molecules in the 2–6 atom domain. It is regrettably true that most organic chemistry is undertaken with larger molecules than these, but such research is still quite valuable in determining the chemistry of important radical intermediates and atmospheric species. Of particular usefulness to astronomers is that many interstellar molecules found in distant molecular clouds and stars are in this size range.

We have now largely completed our discussion of conventional *ab initio* quantum chemical methods. It is worth pointing out that whatever the limitations of these methods, they remain the foundation of the field and the typical starting point for future development. The limitations of traditional quantum chemistry drive researchers to develop new techniques which no longer have these limitations and can thus be used to model chemistry that was previously not theoretically well described. A key area of ongoing methods development, for example, is in the chemistry of electronically excited states. Many reactions are photochemically activated rather than thermally activated, and so understanding the nature of excited states is essential for such situations. Furthermore, there are a number of research directions in spectroscopy that attempt to record the spectra of highly excited and/or short-lived species. Conventional *ab initio* theories are at a loss to explain much of these phenomena, for they are formally only applicable to ground states of molecules (or, more precisely, the lowest-energy state of a given multiplicity and symmetry class). For that matter, predicting the structure of a short-lived species using the time-*independent* Schrödinger equation may be counterintuitive. For these kinds of systems, high-
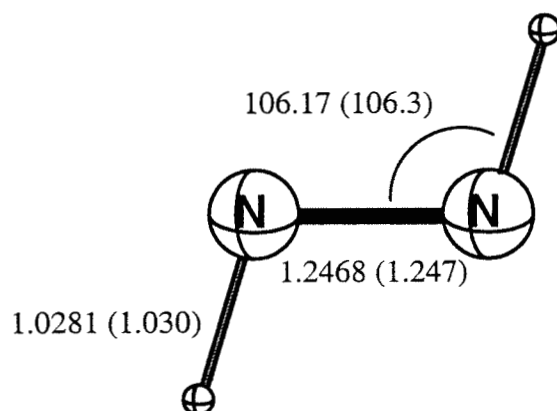
Fig. 1 *trans*-HNNH geometry in Å and degrees (experimental values in parentheses) determined at the CCSD(T)/cc-pVQZ level plus core-valence interaction. See J. M. L. Martin, P. R. Taylor. *Mol. Phys.* **96**, 681 (1999). For the purposes of the present paper, N_2H_2 is considered a "small" molecule.

lying excited state quantum chemical methods are under development. The power of coupled cluster theory makes it especially useful to modify into acceptable excited state methods. One popular variant of the coupled cluster approximation called EOM-CCSD uses an equation of motion formalism to build an effective Hamiltonian for the excited-state wavefunction [31–33]. Other *ab initio* methods have excited-state algorithms as well; most of these are linear response theories [34] that are related to solving time-dependent problems and are valuable for determining excitation energies (such as configuration interaction singles, time-dependent Hartree–Fock, and Green's function methods). Such calculations are extremely versatile in the types of phenomena they can predict; they can even be used to compute the properties of extended systems (such as polymers or crystals) using only a "unit cell" [35]. Ultimately, the development of true multireference methods may be the key to modeling molecular systems regardless of the electronic state, as will be discussed below.

Another problem that manifests itself mostly in excited-state systems is a rather technical but important theoretical weakness. The Born–Oppenheimer approximation constrains the nuclei to remain fixed, and most computer implementations of conventional *ab initio* methods use the symmetry of the molecule to reduce computational cost whenever possible. This optimization is also an aid in the interpretation of results. However, the electrons are not fixed, and their density is not required to obey the same symmetry rules; the result being that in some systems the electronic wavefunction may "break symmetry" and certain properties computed from such a system will be unphysical. This issue can probably be resolved, but it will require the attention of researchers in the field. In isolated high-spin cases, quasi-restricted HF references appear to solve symmetry-breaking problems. A more general solution in coupled cluster theory utilizes so-called "Brueckner orbitals" which can be transformed in such a way as to make the T_1 excitation operator vanish [36,37]. The Brueckner method has already been shown to work extremely well on a number of difficult symmetry-breaking problems. It is unfortunately not a complete solution to the problem, though, as other symmetry-breaking phenomena may still come into play (for example, in the case of Jahn–Teller distortions as recently shown by Crawford and Stanton) [38].

Many of the most interesting research directions in quantum chemistry have involved the solution of multireference problems. Chemically speaking, a multireference system is a molecule or reaction in which the one-electron approximation is not only quantitatively but also qualitatively wrong. Practically all reactions that involve bond-breaking are of this type, as are some low-lying excited electronic states of many molecules. In such situations, the typical HF wavefunction is not a good description of the system, and something more complex is necessary to reproduce the chemistry. This idea has a long lineage in the field, starting with Hartree [8] and continuing to the present day. The simplest idea is the multiconfiguration self-consistent-field (MCSCF) technique, which performs an HF/SCF calculation upon a set of configurations chosen for their chemical importance. More general techniques mainly differ as to how the various determinants are included in the wavefunction; perhaps the most popular method is the complete active space self-consistent Field (CASSCF) method developed by Ruedenberg and co-workers and Roos and co-workers [39,40]. Using CASSCF, the references are selected by choosing an "active space" of several chemically important orbitals and performing a FCI in the span of the active space. For example, if there were eight orbitals, five of them filled, and the highest occupied molecular orbital (HOMO; containing two electrons) and lowest unoccupied molecular orbital (LUMO) were chosen to be active, then there would be four references chosen. Configurations resulting from excitations out of the filled-core orbitals and/or into the inactive unoccupied virtual orbitals should in principle represent high-lying excited states of lesser importance. If it appears necessary to include excitations into the higher orbitals, a perturbational calculation is performed on each reference to improve the correlation description. Such a calculation is referred to as CASSCF plus second-order perturbation theory, or CASPT2. If a completely variational method is desired, then a multireference version of the CI method (MRCI) is the method of choice. It requires, of course, a great deal more computational power to properly elucidate the structure and properties of multireference systems and reaction mechanisms.

What is the future of multireference methods? There are a number of intriguing possibilities. MRCI is valuable, to be sure, but size-consistency corrections for multireference CISD are not trivial to determine. CASPT2 is an excellent general-purpose method for handling many systems, but it is still not general enough to treat certain problems with quasidegenerate high-lying excited states such as some Rydberg states which are of interest to spectroscopists. It would be especially nice to have other readily applicable multireference methods to go along with MRCI and CASPT2. Multireference coupled cluster would be the most attractive, as one could then apply CCSD and CCSD(T) with more confidence toward systems with low-lying excited states. In the MRCC realm, there are formalisms which are state-selective (Hilbert space) [41] and thus more computationally efficient, and there are also formalisms which are state-universal (Fock space; e.g., they can provide energies and properties for a manifold of excited states *simultaneously* in each iteration)[42]. At this point, many of the implementations are based upon Møller–Plesset partitionings which are limited to two-configuration CC from a practical computational standpoint (although they are nevertheless quite useful for describing open-shell singlet excited configurations). Recently, however, some very promising improvements were made in this area; Hubač and co-workers have successfully demonstrated a Brillouin–Wigner partitioning formalism that can be applied to larger molecular systems [43,44]. These calculations are currently at an experimental stage, but such work may eventually provide an alternative to the established multireference methods. It is not yet clear if the CC advantage will be offset by the increased complexity of the MRCC method—only time will tell. A similar situation is probably in store for quasidegenerate perturbation theory (QDPT), which in some forms is truly MRMP2 and in others state-selective [45–48]. QDPT, like MRCC, is still in a developmental stage. There have been some impressive results on small molecules using the state-specific methods, but the progress on state-universal methods is slow. Probably QDPT will never be as popular as MRCI or even MRCC, since MP2 may slowly disappear as density functional theory gains ground.

As a final topic of discussion in small molecule *ab initio* methods, let us return to the fundamental approximations of Hartree–Fock theory. All of the work described above depends on the notion of an effective potential to replace the interelectronic term in the electronic Hamiltonian. Such a potential is smoothly defined from a mathematical perspective, and it is because of this smoothness that it is fundamentally an unsatisfactory description of the electronic Schrödinger equation. It has been long acknowledged that the true Hamiltonian form displays singularities in the Coulomb potential at the points where electrons would exist at the same position. These are called "coalescence points", and they imply that there will be irregularities in the first and higher derivatives of the wavefunction with respect to the interelectronic distance r_{12}. These electron–electron and electron–nuclear cusp conditions were shown by Kato in 1957 [49]. A smooth potential will not intrinsically model this cusp and therefore will be subject to slow convergence behavior. With a finite amount of time and resources on our hands, we are always interested in finding methods with faster convergence properties.

Fortunately, a solution to this problem has been known since the early days of quantum mechanics. Hylleraas attacked the problem of the ground-state helium atom with expansions that were explicitly dependent on r_{12}. His three-term trial wave function including a linear r_{12} term variationally optimized by hand [5] yielded an energy within ≈ 0.6 kcal/mol of the exact nonrelativistic value—it would require ≈ 30 terms in a CI calculation (and a computer, obviously) to match that feat! Calculations in 1994 incorporating more complicated expansions have achieved results approaching the exact nonrelativistic energy of He atom to within 10^{-12} kcal/mol [50]. Earlier, Kolos and Wolniewicz reported accurate r_{12} energies for the hydrogen molecule, including nuclear motion beyond the Born–Oppenheimer approximation and relativistic and radiative effects [51]. The reason for this fast convergence is that the r_{12} terms provide the necessary electron–electron cusp behavior.

Can we extend this technique to the province of small molecule quantum chemistry? It appears the answer is "yes". Extending Hylleraas' approach to systems containing three or more electrons results in some very difficult mathematics. More realistic explicitly correlated methods are of two types. The first type uses Boys' Gaussian-type geminal (GTG) functions instead of normal basis functions,

with the explicit dependence on the interelectronic coordinate built into the exponent of the integral [52]:

$$\phi(i,j) = e^{(-\alpha r_{iA}^2 - \beta r_{jB}^2 - \gamma r_{ij}^2)} \tag{12}$$

While the resulting integrals are certainly less bothersome than in a Hylleraas-type CI expansion, a number of technical problems remain to be solved before the method will be applicable to chemical systems. A Gaussian-type geminal is nothing but a product of two normal Gaussian basis functions with an additional factor of the form $e^{-\gamma r_{ij}^2}$; hence expansions in terms of GTGs do not rigorously satisfy the cusp condition. Another variation of interest is due to Kutzelnigg, in which a degree of dependence on r_{12} is introduced into the theoretical method (i.e., MP2 or CC) itself, using the idea that such terms will improve the traditional high-level methods to the point that the electron–electron cusp problem will be properly treated [53]. These "linear r_{12}" methods utilize a modified version of molecular integral codes to avoid some of the problems associated with explicit r_{12} terms. Practical computations have already been performed using the MP2–r_{12} method on molecules as large as ferrocene. Both of these techniques show significant promise and could very well be in standard usage in the next quarter-century.

MEDIUM MOLECULES

Slightly larger molecules, which include most of the interesting species in organic chemistry (ring systems and dendrimers, for example), cannot easily be treated with the highest-level methods. Furthermore, if a system includes metal atoms (like many active sites of proteins or organometallic compounds), the systems can become intractable at higher levels of theory. Thus the compromise begins. First approximations appear in the sense of purely theoretical examinations of basis set convergence. For small systems, very large basis sets may be used, but for larger systems (the cutoff being as few as 12 atoms presently) such computations are impractical. If some extrapolation were possible, the smaller basis set would suffice as a good estimate of the final properties could be determined by applying the extrapolation. This is the premise behind the infinite-basis-set extrapolation techniques. Most of these methods use the "correlation-consistent" [54] basis sets developed by Dunning and co-workers or the atomic natural orbital (ANO) [55] basis sets introduced by Almlöf and Taylor. Extrapolation techniques can be more computationally demanding than the other intermediate-level methods, but from a quantum chemical perspective, they are also more aesthetically pleasing.

It is also at this point that empiricism becomes more acceptable as well, for assumptions may be made as to the properties of a molecule based on molecules similar to it. Various theories, following the example of the Gaussian-N (G2, G3, etc.) [56] and the complete basis set (CBS) [57] methods, apply certain low-level methods in a specific order to the molecule and on fragments of it in certain categories (i.e., homodesmic reactions and such), in an effort to emulate a higher-level method by inductive reasoning. These methods have been quite successful for thermochemical predictions, an area in which molecular mechanics methods are often less reliable. One example of such a theory, the conventional G2 theory, utilizes the following quantum chemical techniques, basis sets, and empirical parameters to determine the model chemistry energy:

$$\Delta E_+ = E[MP4 / 6-311+G(d,p)] - E[MP4 / 6-311G(d,p)] \tag{13}$$

$$\Delta E_{2df} = E[MP4 / 6-311G(2df,p)] - E[MP4 / 6-311G(d,p)] \tag{14}$$

$$\Delta E_{QCI} = E[QCISD(T) / 6-311G(d,p)] - E[MP4 / 6-311G(d,p)] \tag{15}$$

$$\Delta E_{+3df,2p} = E[MP2 / 6-311+G(3df,2p)] - E[MP2 / 6-311G(2df,p)]$$
$$- E[MP2 / 6-311+G(d,p)] + E[MP2 / 6-311G(d,p)] \tag{16}$$

$$\Delta_{HLC} = -4.81 N_\beta - 0.19 N_\alpha \tag{17}$$

$$E_{G2} = E[MP4/6-311G(d,p)] + \Delta E_+ + \Delta E_{2df} + \Delta E_{QCI} + \Delta E_{+3df,2p} + \Delta E_{HLC} \tag{18}$$

Other versions of G2 and later theories [58] use DFT [59] or even CCSD(T) [60] calculations in the model. Effective core potentials have also been used instead of conventional basis sets to reduce computation times [61]. The rationale behind each model chemistry is determined by a fitting to a library of compounds, called the test set [62]. The G2 test set as originally formulated is comprised of 125 molecules representing a variety of general and difficult cases in organic chemistry; later test sets contain more compounds, but a difficult balance must be struck between wide applicability and overall accuracy. This is due to the uncertainties associated with the experimental results used in each parameterization, as well as a general phenomenon of all parameterization schemes. It is, of course, possible to add more and more methods with increasing numbers of terms so as to better predict a greater variety of compounds. Unfortunately, such a scheme may become "overtrained" and be generally applicable only to those compounds in the set—which is hardly the outcome desired! These methods are nevertheless a subject of great interest and importance, for we cannot expect to use our most powerful methods on larger systems anytime soon, if at all. By contrast, CBS and Gaussian-N are already computationally feasible for medium molecules and have the potential to become the *de facto* standards, as they are implemented in the popular Gaussian suite of programs. At present, G2 theory can provide thermochemical energies within 1.4 kcal/mol and CBS-Q to about 0.8 kcal/mol of experiment, which is significantly less accurate than the CCSD(T) work detailed in Fig. 1 [30]. Nevertheless, there is little doubt that since they are capable of providing relative energies and properties approaching chemical accuracy, they will account for many accurate computations during the next 10–20 years.

The truly spectacular development in this new quantum chemical era is density functional theory (DFT). Density functional studies of organic and inorganic systems abound in today's literature. For example, as early as 1998 it was possible to report all-electron complete geometry optimizations for $Fe_3(CO)_{12}$ with two DFT methods and a flexible basis set [63]. Extension to systems as large as $Co_6(CO)_{16}$ is already feasible, again with good basis sets, explicit consideration of all electrons, and full geometry optimization. Especially common are studies of organic molecules and reactions that have a number of conformations and/or fragments; DFT can compute the structure and properties of these species with reasonable accuracy but cannot always be relied upon for relative energetics of such conformers, as was recently shown in the case of [10]annulene (see Fig. 2) [64]. DFT represents somewhat of a departure from conventional *ab initio* quantum chemical methods in that it does not formally deal with wavefunctions, but rather with the electron density. In some revolutionary work done the early 1960s, Kohn and co-workers discovered that all ground-state properties of a molecular system could be completely described using only the electron density and the (fixed) nuclear positions. The first Hohenberg–Kohn theorem [65] proved the existence of a unique functional relationship between the electron density and the electronic Hamiltonian. Thus, if the exact relationship were known we could compute all of the molecular properties using only the electron density (which depends on only three coordinates) and avoid the building of a wavefunction (which depends on $3n$ coordinates). Much of the development work in modern-day quantum chemistry is directed toward discovering the true nature of this functional relationship.

From a practical perspective, the most important advance was made in 1965 by Kohn and Sham. Utilizing the well-developed iterative techniques of the HF/SCF method, a computationally feasible approach to DFT was found which avoided the explicit construction of the electronic kinetic energy portion of the energy functional. This formalism is called Kohn–Sham density functional theory (KS-DFT) [66] and it can be fairly easily implemented with computation times comparable to conventional SCF procedures. The chief difference between the SCF and KS-DFT approaches lies in the exchange-correlation (XC) density functional, which accounts for both electron correlation and nonclassical exchange effects on the energy. If the exact form for the XC functional were known, the KS-DFT approach would give the exact energy. KS-DFT therefore requires the construction of accurate XC functionals. Its com-
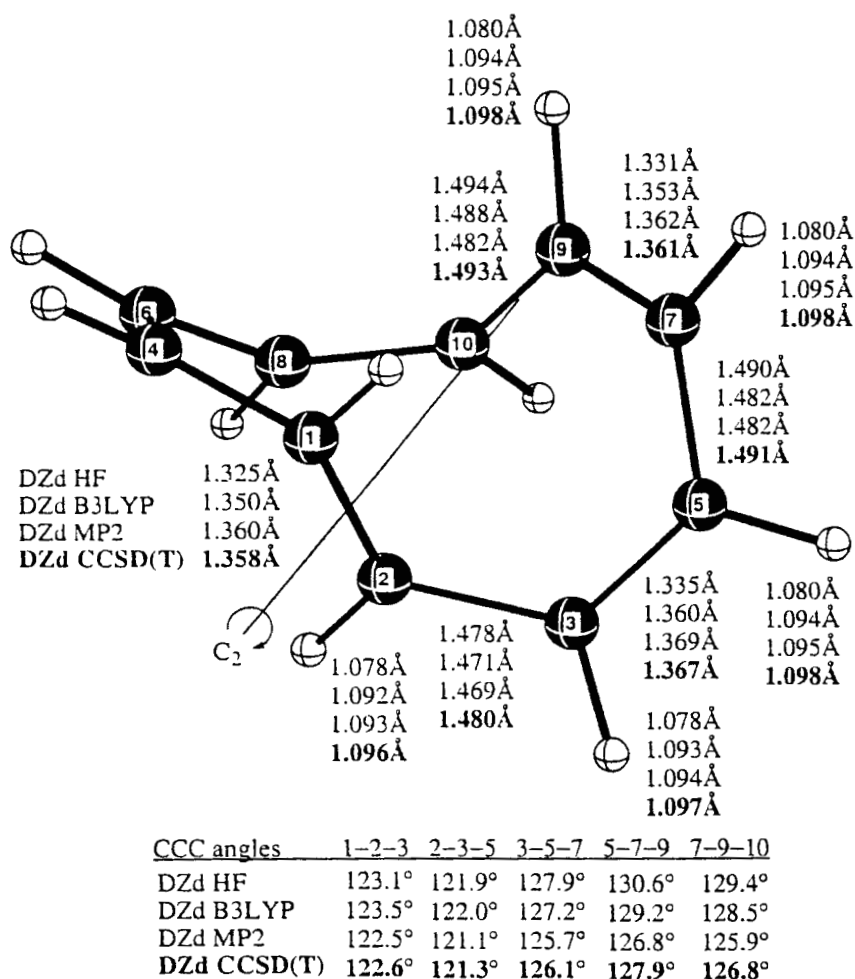
Fig. 2 Fully optimized structures of the "twist" conformation of [10]annulene, $C_{10}H_{10}$, computed using a Dzd basis set and several different quantum chemical methods. See R. A. King, T. D. Crawford, J. F. Stanton, H. F. Schaefer. *J. Am. Chem. Soc.* **121**, 10788 (1999). For the purposes of the present paper, [10]annulene is considered a "medium" molecule.

putational efficiency rivals that of the HF/SCF procedure while incorporating (at least approximately) electron correlation and exchange.

Early DFT-related methods such as Thomas–Fermi theory, Xα, and the local density approximation (LDA) were in large part based upon fitting the functionals to the theoretical electron gas. Not surprisingly, these were useful in solid-state physics (where periodic boundary conditions dictate a significant part of the electronic structure) but not nearly as impressive for chemical systems. The first truly successful DFT functionals, the generalized gradient approximation (GGA) functionals, were of the KS-DFT variety and relied on XC terms designed using density gradient corrections. Some of the most popular functionals are of this type, including the Becke-88 (B88) exchange functional [67] and the Lee–Yang–Parr (LYP) [68] and Perdew-86 (P86) [69] correlation functionals. These functionals were still found to be lacking for nonlocal molecular phenomena, and so hybrid functionals were built that contained some proportion of exact (i.e., HF) exchange. Of this type, the Becke three-parameter exchange functional (B3) [70] is still the most common. If exchange and correlation functionals are

combined, modern DFT functionals (B3LYP, BP86, etc.) are the result. Currently, DFT enjoys an enormous amount of acceptance for predicting molecular structures and properties due to its relative accuracy and increased speed over other quantum chemical methods. Systems involving bond-breaking or weak bonds are still problematic, however. The next phase of DFT functionals may be the meta-GGA or τ-dependent functionals, which remove the necessity of exact exchange by substituting the kinetic energy gradient [71–74]. Certain improvements can be realized from a more theoretical route by performing the calculations numerically (basis-set-free) as in Becke's NUMOL program [75], but this appears to be very expensive when high precision is required. Though DFT is formally a theory of the ground state, time-dependent DFT (TD-DFT) methods [76,77] have been implemented using established linear-response formalisms from HF/SCF and CC theories, providing encouraging results for the energies and time-dependent properties of singly excited states. In addition, certain implementation-related optimizations are possible. An example is grid-free DFT, which allows calculation without a numerical grid but instead using a nonlinear equation solver to fit the integrable functional form [78]. Further improvements to KS-DFT are clearly an area of frenetic theoretical interest.

Unfortunately, while in principle DFT theory is not empirical, we do not have the exact functional and should not expect to. We must therefore continue to develop functionals that perform increasingly well on the compounds of choice, often using parameters found in a manner similar to the above discussion of theoretical model chemistries [79]. Recent work in this direction has produced functionals fitted to as many as 429 chemical systems [80]. There is some promising work in forming a strategy for improvement of functionals based on coupled cluster theory and other high-level methods [81]. Regardless, to date there have been no successful attempts to systematically improve a functional on a purely theoretical basis in an effort to find the exact functional. Until such time as a method may be found, we must be content with the rather *ad hoc* functionals being developed today. However, we cannot dispute the enormous significance of this method, for it dramatically reduces the computational cost of a calculation while still satisfying many of the sound principles of *ab initio* quantum chemical methods. DFT has every indication of being the dominant medium-to-large molecular method for some time to come.

LARGE MOLECULES

There is little question that the great advances in medicine have been one of the most impressive developments of the 20th century. The turn of the twentieth century brought humankind a rudimentary understanding of organic chemistry and how medicines might be developed, and during the rest of the century (peacetime or otherwise) a great deal of effort was expended in advancing chemical and biological knowledge so that disease might be eradicated and the overall quality of life elevated. With that in mind, the turn of this new century appears to be a time when understanding has become great enough that never-before-seen molecules may be designed with a view toward producing new drugs. Unfortunately, knowledge is not yet to the point that it is understood what makes some molecules biologically active and others biologically inert. Much of the work in large-molecule chemical computation involves molecular mechanics, genetic algorithms, or semiempirical schemes that are very fast compared to high level *ab initio* quantum chemical methods but typically not as accurate; such techniques, ranging from statistics-based quantitative structure–activity relationships (QSAR) to classical-physics-based molecular dynamics (MD) [82], belong to computational chemistry. These methods rely on simple models and empiricism mostly by necessity, for even a Hartree–Fock calculation is difficult on a biomolecular system such as a protein. How can quantum chemistry play a role in this? We, too, must make some adjustments.

Some of the methods available today attempt to explicitly handle certain interactions while replacing other interactions with computationally simpler approximations. Pseudospectral methods achieve computational advantages by treating the electron–electron repulsion term numerically, while other terms are treated in a conventional manner [83]. Hybrid methods that are quantum mechanical in

nature can perform, say, a CCSD(T) calculation on the active site of a molecule and a Hartree–Fock (or molecular mechanics) calculation on the rest, for much better performance in the local area of particular interest [84]. Other techniques focus more on the intrinsic differences between small and large molecular systems, deliberately neglecting the fine points that would be important in small molecules and instead examining the effects that more significantly affect a large molecule. For example, the concept of an "equilibrium geometry" is meaningful for a triatomic molecule existing in a vacuum at near 0 K but is less valuable for a DNA strand existing in solution at body temperature. Solvation methods are in development for precisely this reason. Some of these methods postulate a mathematical relationship between solvent and solute in a manner similar to QSAR [85]. Other models are more explicit as to their theoretical description, as in local-density-functional MD [86,87] or in the electrostatic polarizable continuum model (PCM) [88]. Another possibility lies in applying a simple "reaction field" correction to an SCF procedure, much like KS-DFT adds its exchange–correlation terms; such Self-Consistent Reaction Field (SCRF) methods [89] have been extended to coupled cluster theory [90] and incorporated into the popular SMx semiempirical solvation technique [91]. Further development into large-molecule-specific methods will most likely be guided by specialists in other fields and will be crucial to continued progress in those areas.

The above discussion notwithstanding, the foundation of *ab initio* quantum chemistry remains sound even if the questions it answers may be more specific than actually necessary for most large molecules. Researchers continue to increase the practical size for such computations through efforts unrelated to external improvements in computers. Indeed, quantum chemistry has a history of improving algorithms on a fundamental level in order to eke out more computational power; the analytic derivative techniques are an excellent historical example [92]. At first, most geometry optimizations of the HF/SCF variety were done by hand, point by point, until an energy minimum was found. It was eventually realized that optimizations could be performed much faster if derivatives of the energy with respect to nuclear coordinates could be calculated analytically. In fact, many molecular properties could be more easily computed if derivative codes were readily available. Analytic derivatives for HF/SCF were pioneered by Pulay in 1969, and similar methods for MP2 and CISD were fairly quick in following [93–95]. Today, just about every method in general use either has analytic derivatives implemented or is in the process of being implemented. Similarly, direct methods followed this same type of evolution. As soon as it was realized that it would be possible to compute molecular integrals on-the-fly rather than store them, the size of chemical systems available for study increased tremendously [96]. CI theory also benefited greatly from a conceptual reformulation called the graphical unitary group approach (GUGA) [97,98].

Linear scaling techniques appear to be the vanguard of this type of algorithmic research. The high scaling associated with coupled cluster (i.e., $O(N^7)$ for CCSD(T)) and other theories can be alleviated by sacrificing a little accuracy along the way or more efficiently implementing the methods, allowing them to be used on much larger systems than otherwise would be practical (or even possible). Such work hinges on the revelation that while two-electron repulsion integral evaluation scales as $O(N^4)$, the number of important two-electron integrals scales as only $O(N^2)$. Therefore it would appear that there is no physical reason why a program could not be written that would compute only those integrals which are truly necessary to perform the calculation. Most of these algorithms, currently under examination primarily for integrals in HF/SCF or DFT methods, utilize a polynomial (or "multipole") expansion of the r_{12}^{-1} integral component. By expanding and partitioning the integral overlap density into interacting components, these techniques (such as the Fast Multipole Method or the Recursive Bisection Method) [99–101] can achieve linear or near-linear scaling with very little loss of rigor. The future of large-scale quantum chemical calculations may be found in the roots of this preliminary development, a most impressive example of which was recently reported by Scuseria (see Fig. 3) [102,103].

Ultimately, highly parallel machines may very well be the most cost-effective way to solve many of the computational bottlenecks [104]. By combining large numbers of powerful computers (a term which now includes most personal computers as well), as long as an efficient parallel algorithm can be

found, we should be able to extend some of the higher level *ab initio* quantum chemical methods into the biomolecular and metal cluster regime. With the aforementioned theoretical developments combined with the ever-increasing power of computers, some of the methods listed in the previous section should become tenable even for large molecules in the next decade or so.

THE GREAT UNKNOWN: COMPUTING TECHNOLOGIES

This paper began by pointing out quantum chemistry's greatest ally in its development: constant improvement in computing technology. The increase in computing power has been exponential in
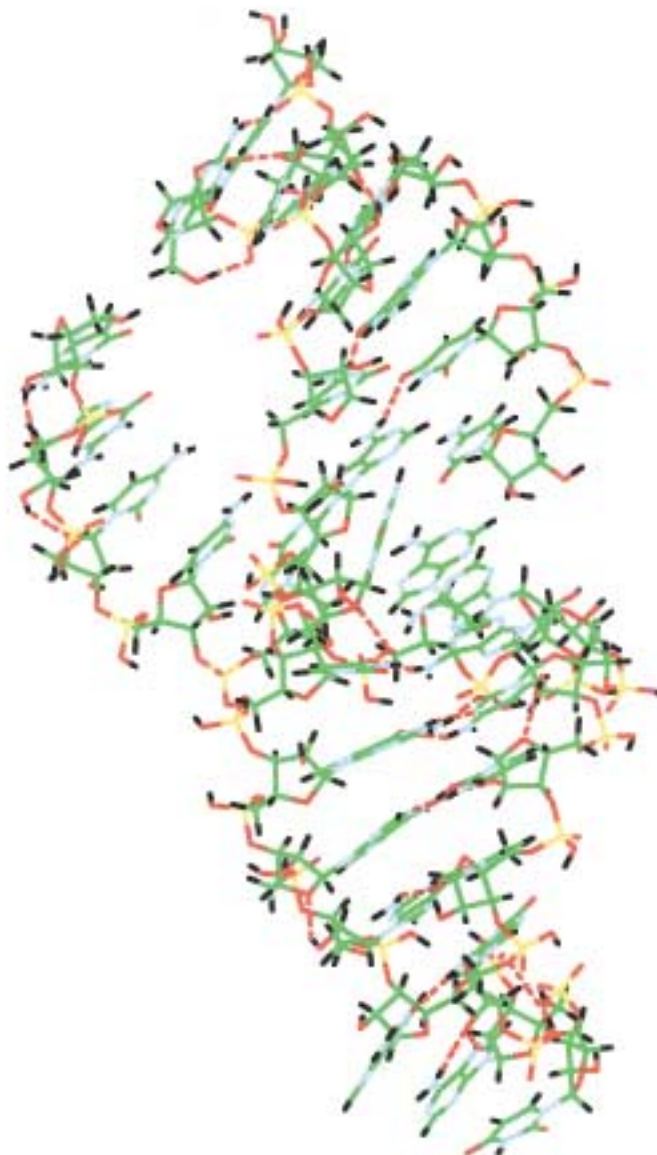
Fig. 3 A 1026 atom fragment of RNA, calculated using linear scaling DFT at the LSDA/3-21G level. See G. E. Scuseria. *J. Phys. Chem.* A **103**, 4782 (1999). For the purposes of the present paper, this RNA fragment is considered a "large" molecule.

nature for quite some time, while costs continue to plummet as older technology becomes common-place. Indeed, there is a prediction coined "Moore's Law" which states that the speed of computers will double approximately every eighteen months with no additional cost. It is not, of course, truly a law, but only a surprisingly accurate picture of the current technology sector. How much longer will this growth continue? It would appear that the answer is "until chip manufacturers reach the 157-nm barrier." Below that wavelength, there are currently no known lasers. Even before reaching 157 nm, smaller photolithography for processor designs will necessitate moving production into the vacuum ultraviolet, and costs will increase considerably for such conditions as compounds in the air will interfere with the transmission of radiation. Many experts believe that Moore's Law will cease to be true in the next 10 years [105].

Is this the end, then, for improvements in computing speed? It would certainly be shortsighted to suggest so! It is true that the current situation seems to suggest a future plateau, but it is most certainly not a permanent one. New lasers will probably be found to continue the miniaturization down to the point of theoretical breakdowns due to the increasingly quantum–mechanical effects such as tunneling. At that point some other method will have to be used to improve speed further. Some of the possibilities are in directions varied, far-reaching, or bizarre. Quantum computers, which could potentially use the superposition of states to perform many calculations simultaneously, have been demonstrated on a prototype scale [106]. It is unclear how algorithms would be written for such a "nondeterministic" computer, but the idea is promising. Other possibilities include biocomputing (using bacteria or DNA to perform computer-like functions) [107], nanocomputing (small chemical-based Boolean logic circuits) [108], and optical computing (using photons to perform switching functions and construct logic devices with nearly speed-of-light efficiency) [109]. These possibilities (or some equally exciting technologies as yet unknown) will undoubtedly become useful in the century to come.

CONCLUSION

Theoretical predictions about the structure and properties of molecules have come a long way since 1929. Even as early as 1959, the advances made in quantum chemistry were significant enough to warrant a review. In their famous "Broken Bottlenecks" article, Mulliken and Roothaan made their own pronouncement about the efficacy of quantum chemistry:

> Looking toward the future, it seems certain that colossal rewards lie ahead from large quantum–mechanical calculations of the structure of matter [110].

In 1984, one of us commented in a review of the quantum chemical literature:

> We are confident that by the year 2000, essentially all fields of chemistry will acknowledge the accuracy of Mulliken and Roothaan's prophecy [111].

It is now time to note the accuracy of both prophecies, and to make a new one: We, as quantum chemists, now predict that by the year 2100, those "colossal rewards" will have been largely realized, and their consequences will be so striking that essentially all fields of *science* will acknowledge the accuracy of Mulliken and Roothaan's prophecy.

ACKNOWLEDGMENTS

REFERENCES

1. W. Heitler and F. London. *Z. Physik* **44**, 455 (1927).
2. D. R. Hartree. *Proc. Cambridge Philos. Soc.* **24**, 328 (1928).
3. V. A. Fock. *Z. Phys.* **15**, 126 (1930).
4. E. A. Hylleraas. *Z. Physik* **48**, 469 (1928).
5. E. A. Hylleraas. *Z. Physik* **54**, 347 (1929).
6. H. M. James and A. S. Coolidge. *J. Chem. Phys.* **1**, 825 (1933).
7. C. Møller and M. S. Plesset. *Phys. Rev.* **46**, 618 (1934).
8. D. R. Hartree, W. Hartree, B. Swirles. *Phil. Trans. Roy. Soc. (London)* **A238**, 229 (1939).
9. P. A. M. Dirac. *Proc. Roy. Soc. (London)* **123**, 714 (1929).
10. H. F. Schaefer. *The Electronic Structure of Atoms and Molecules: A Survey of Rigorous Quantum Mechanical Results*. Addison-Wesley, Reading, Massachusetts (1972).
11. H. F. Schaefer. *Science* **231**, 1100 (1986).
12. H. F. Schaefer. *Chimia* **43**, 1 (1989).
13. H. F. Schaefer. *J. Mol. Struct.* **398**, 199 (1997).
14. M. Born and J. R. Oppenheimer. *Ann. Phys.* **79**, 361 (1927).
15. C. C. J. Roothaan. *Rev. Mod. Phys.* **23**, 69 (1951).
16. S. F. Boys. *Proc. Roy. Soc. (London)* **A200**, 542 (1950).
17. J. A. Pople, R. K. Nesbet. *J. Chem. Phys.* **22**, 571 (1954).
18. A. Meckler. *J. Chem. Phys.* **21**, 1750 (1953).
19. J. A. Pople, M. Head-Gordon, K. Raghavachari. *J. Chem. Phys.* **10**, 5968 (1987).
20. S. R. Langhoff and E. R. Davidson. *Int. J. Quantum Chem.* **8**, 61 (1974).
21. J. Čížek. *J. Chem. Phys.* **45**, 4526 (1966).
22. J. Čížek. *Adv. Chem. Phys.* **14**, 35 (1969).
23. J. Čížek and J. Paldus. *Int. J. Quantum Chem.* **5**, 359 (1971).
24. A. C. Hurley. *Electron Correlation in Small Molecules*. Academic Press, London (1976).
25. H. J. Monkhorst. *Int. J. Quantum Chem.* **11**, 421 (1977).
26. J. A. Pople, R. Krishnan, H. B. Schlegel, J. S. Binkley. *Int. J. Quantum Chem. Symp.* **14**, 545 (1978).
27. R. J. Bartlett and G. D. Purvis. *Int. J. Quantum Chem.* **14**, 561 (1978).
28. G. D. Purvis and R. J. Bartlett. *J. Chem. Phys.* **76**, 1910 (1982).
29. K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon. *Chem. Phys. Lett.* **157**, 479 (1989).
30. J. M. L. Martin and P. R. Taylor. *Mol. Phys.* **96**, 681 (1999).
31. J. F. Stanton and R. J. Bartlett. *J. Chem. Phys.* **98**, 7029 (1993).
32. R. J. Bartlett and J. F. Stanton. *Rev. Comput. Chem.* **5**, 65 (1994).
33. T. D. Crawford and H. F. Schaefer. *Rev. Comput. Chem.* **14**, 33 (1999).
34. H. Koch and P. Jørgensen. *J. Chem. Phys.* **93**, 3333 (1990).
35. S. Hirata, M. Head-Gordon, R. J. Bartlett. *J. Chem. Phys.* **111**, 10774 (1999).
36. N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, G. W. Trucks. *Chem. Phys. Lett.* **164**, 185 (1989).
37. J. F. Stanton, J. Gauss, R. J. Bartlett. *J. Chem. Phys.* **97**, 5554 (1992).
38. T. D. Crawford and J. F. Stanton. *J. Chem. Phys.* **112**, 7873 (2000).
39. B. O. Roos, P. R. Taylor, P. E. M. Siegbahn. *Chem. Phys.* **48**, 157 (1980).
40. K. Ruedenberg, M. Schmidt, M. M. Gilbert, S. T. Elbert. *Chem. Phys.* **71**, 41 (1982).
41. B. Jeziorski and H. J. Monkhorst. *Phys. Rev. A* **24**, 1668 (1981).
42. D. Mukherjee, R. K. Moitra, A. Mukhopadhyay. *Mol. Phys.* **30**, 1861 (1975).
43. I. Hubač and P. Neogrády. *Phys. Rev. A*, **50**, 4558 (1994).
44. J. Pittner, P. Nachtigall, P. Carsky, J. Masik, I. Hubač. *J. Chem. Phys.* **110**, 10275 (1999).
45. Y. G. Khait and M. R. Hoffmann. *J. Chem. Phys.* **108**, 8317 (1998).

46. H. Nakano, K. Hirao, M. S. Gordon. *J. Chem. Phys.* **108**, 5660 (1998).
47. J. Malrieu, J. Heully, A. Zaitsevskii. *Theor. Chim. Acta* **90**, 167 (1995).
48. K. F. Freed. In *Many-Body Methods in Quantum Chemistry*, U. Kaldor (Ed.), Springer, Berlin (1989).
49. T. Kato. *Pure Appl. Math.* **10**, 151 (1957).
50. G. W. F. Drake and Z. Yan. *Chem. Phys. Lett.* **229**, 486 (1994).
51. W. Kolos and L. Wolniewicz. *J. Chem. Phys.* **49**, 404 (1968).
52. S. F. Boys. *Proc. Roy. Soc. (London)* **A258**, 402 (1960).
53. W. Kutzelnigg. *Theor. Chim. Acta* **68**, 445 (1985).
54. T. H. Dunning. *J. Chem. Phys.* **90**, 1007 (1989).
55. J. Almlöf and P. R. Taylor. *J. Chem. Phys.* **86**, 4070 (1987).
56. J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, L. A. Curtiss. *J. Chem. Phys.* **90**, 5622 (1989).
57. J. W. Ochterski, G. A. Petersson, J. A. Montgomery. *J. Chem. Phys.* **104**, 2598 (1996).
58. L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, J. A. Pople. *J. Chem. Phys.* **109**, 7764 (1998).
59. C. W. Bauschlicher and H. Partridge. *J. Chem. Phys.* **103**, 1788 (1995).
60. L. A. Curtiss, K. Raghavachari, P. C. Redfern, A. G. Baboul, J. A. Pople. *Chem. Phys. Lett.* **314**, 101 (1999).
61. M. N. Glukhovtsev, A. Pross, M. P. McGrath, L. Radom. *J. Chem. Phys.* **103**, 1878 (1995).
62. L. A. Curtiss, K. Raghavachari, G. W. Trucks, J. A. Pople. *J. Chem. Phys.* **94**, 7221 (1991).
63. J. H. Jang, J. G. Lee, H. Lee, Y. Xie, H. F. Schaefer. *J. Phys. Chem. A* **102**, 5298 (1998).
64. R. A. King, T. D. Crawford, J. F. Stanton, H. F. Schaefer. *J. Am. Chem. Soc.* **121**, 10788 (1999).
65. P. Hohenberg and W. Kohn. *Phys. Rev.* **136**, B864 (1964).
66. W. Kohn and L. J. Sham. *Phys. Rev.* **140**, A1133 (1965).
67. A. D. Becke. *Phys. Rev. A* **38**, 3098 (1988).
68. C. Lee, W. Yang, R. G. Parr. *Phys. Rev. B* **37**, 785 (1988).
69. J. P. Perdew. *Phys. Rev. B* **33**, 8822 (1986).
70. A. D. Becke. *J. Chem. Phys.* **98**, 5648 (1993).
71. T. Van Voorhis, G. E. Scuseria. *J. Chem. Phys.* **109**, 400 (1998).
72. A. D. Becke. *J. Chem. Phys.*, **109**, 2092 (1998).
73. J. P. Perdew, S. Kurth, A. Zupan, P. Blaha. *Int. J. Quantum Chem.* **75**, 889 (1999).
74. M. Ernzerhof, G. E. Scuseria. *J. Chem. Phys.* **111**, 911 (1999).
75. A. D. Becke. *J. Chem. Phys.* **88**, 2547 (1988).
76. C. Jamorski, M. E. Casida, D. R. Salahub. *J. Chem. Phys.* **104**, 5134 (1996).
77. R. Bauernschmitt and R. Ahlrichs. *Chem. Phys. Lett.*, **256**, 454 (1996).
78. K. S. Werpetinski and M. Cook. *J. Chem. Phys.* **106**, 7124 (1997).
79. A. D. Becke. *J. Chem. Phys.* **107**, 8554 (1997).
80. G. K. L. Chan and N. C. Handy. *J. Chem. Phys.* **112**, 5639 (2000).
81. V. E. Ingamells and N. C. Handy. *Chem. Phys. Lett.* **248**, 373 (1996).
82. *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, K. Binder (Ed.), Oxford University Press, New York (1995).
83. R. A. Friesner. *Chem. Phys. Lett.* **116**, 39 (1985).
84. J. Aqvist and A. Warshel. *Chem. Rev.* **93**, 2523 (1993).
85. *Quantitative Treatment of Solute/Solvent Interactions*, P. Politzer and J. S. Murray (Eds.), Elsevier, New York (1994).
86. R. Car and M. Parinnello. *Phys. Rev. Lett.* **55**, 247 (1985).
87. M. Tuckerman, K. Laasonen, M. Sprik, M. Parrinello. *J. Chem. Phys.* **103**, 150 (1995).
88. J. Tomasi and M. Persico. *Chem. Rev.* **94**, 2027 (1994).
89. M. W. Wong, M. J. Frisch, K. B. Wiberg. *J. Am. Chem. Soc.* **113**, 4776 (1991).

90. O. Christiansen and K. V. Mikkelsen. *J. Chem. Phys.* **110**, 1365 (1999).
91. C. J. Cramer and D. G. Truhlar. *J. Comp.-Aided Mol. Design* **6**, 629 (1992).
92. Y. Yamaguchi, Y. Osamura, J. D. Goddard, H. F. Schaefer. *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory.* Oxford University Press, New York (1994).
93. P. Pulay. *Mol. Phys.* **17**, 197 (1969).
94. J. A. Pople, R. Krishnan, H. B. Schlegel, J. S. Binkley. *Int. J. Quantum Chem. Symp.* **13**, 225 (1979).
95. B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, H. F. Schaefer. *J. Chem. Phys.* **72**, 4652 (1980).
96. B. Roos. *Chem. Phys. Lett.* **15**, 153 (1972).
97. I. Shavitt. *Int. J. Quantum Chem. Symp.* **11**, 131 (1977).
98. B. R. Brooks and H. F. Schaefer. *J. Chem. Phys.* **70**, 5092 (1979).
99. C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon. *Chem. Phys. Lett.* **230**, 8 (1994).
100. C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon. *Chem. Phys. Lett.* **253**, 268 (1996).
101. M. C. Strain, G. E. Scuseria, M. J. Frisch. *Science* **271**, 51 (1996).
102. G. E. Scuseria. *J. Phys. Chem. A* **103**, 4782 (1999).
103. G. E. Scuseria and P. Y. Ayala. *J. Chem. Phys.* **111**, 8330 (1999).
104. R. A. Kendall, R. J. Harrison, R. J. Littlefield, M. F. Guest. *Rev. Comput. Chem.* **6**, 209 (1995).
105. C. C. Mann. *Technol. Rev.* **103**, 42 (2000).
106. M. M. Waldrup. *Technol. Rev.* **103**, 60 (2000).
107. M. Gerstein. *Nature* **399**, 101 (1999).
108. D. H. Lowndes. *Adv. Mater. Process* **157**, 48 (2000).
109. L. J. Irakliotis and P. A. Mitkas. *Computer* **31**, 36 (1998).
110. R. S. Mulliken and C. C. J. Roothaan. *Proc. U. S. Natl. Acad. Sci.* **45**, 394 (1959).
111. H. F. Schaefer. *Quantum Chemistry: The Development of Ab Initio Methods in Molecular Electronic Structure Theory*. Clarendon Press, Oxford (1984).