

Q-Chem: Quantum Chemistry Software for Large Systems

Peter M.W. Gill

Q-Chem, Inc.
Four Triangle Drive
Export, PA 15632, USA

and

Department of Chemistry
University of Cambridge
Cambridge, CB2 1EW, England

Introduction

The 1998 Nobel Prize for Chemistry was awarded jointly to Professors Pople and Kohn for the development of computational methods in quantum chemistry and, in particular, for the development of density-functional theory (DFT). The award acknowledges the enormous strides that these men and their disciples have taken during the last three decades and celebrates the revolution in chemical philosophy and practice that their ideas have wrought.

Quantum chemical calculations are being utilised by a continually increasing range of computational chemists and, facilitated by impressive improvements in computer hardware, they can now be used to solve much more complicated problems than one would have thought possible a few years ago. Furthermore, it seems likely that these trends will continue in the coming years: computers will become faster and cheaper and the algorithms and software of quantum chemistry will be radically transformed as researchers strive to apply first-principles quantum mechanics to problems of biological scope and complexity.

In this article, we will review both the classical and DFT approaches to quantum chemistry and we will discuss the theoretical and implementational problems that have restricted the application of these approaches to small systems. Finally, we will describe some of the progress that the Q-Chem collaborators have recently made toward the solution of these problems.

Wavefunction-based approaches

In 1926, Schrödinger introduced wave mechanics, proposing that any physical system in a stationary state possesses a wavefunction Ψ and, for many years, the goal of classical quantum chemistry was to obtain useful approximate solutions to the time-independent electronic Schrödinger Equation

$$\hat{H}\Psi = E\Psi \quad (1)$$

The resulting models (Hartree-Fock, configuration interaction, coupled-cluster and Møller-Plesset theories) were the standard methods of the '70s and '80s and were widely used, especially by academic researchers. It is useful to express the total electronic energy E as

$$E = E_T + E_J + E_X + E_C \quad (2)$$

where these terms are the kinetic, Coulomb, exchange and correlation energies, respectively. In wavefunction-based approaches, each of these four energies is calculated as a functional of the wavefunction. The Fock exchange functional, for example, is given by

$$E_X^{\text{Fock}} = \frac{1}{2} \sum_i \sum_j \iint \Psi_i(\mathbf{r}_1)\Psi_j(\mathbf{r}_1) \frac{1}{r_{12}} \Psi_i(\mathbf{r}_2)\Psi_j(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (3)$$

Density-based approaches

In the mid-60s, Kohn and his collaborators proved that it is theoretically possible to calculate E as a functional of the electron density $\rho(\mathbf{r})$ and, thereby, to avoid consideration of the wavefunction entirely. This was an important discovery because $\rho(\mathbf{r})$ is a much simpler object than Ψ . In Kohn-Sham theory, the kinetic energy E_T is computed from the orbitals (as it is in Hartree-Fock theory) but the remaining terms in (2) are calculated as functionals of $\rho(\mathbf{r})$. The Dirac exchange functional, for example, is given by

$$E_X^{\text{Dirac}} = \frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \int \rho(\mathbf{r})^{4/3} d\mathbf{r} \quad (4)$$

Many exchange/correlation density functionals have been developed.

Cost and Accuracy

As of 1998, the most sophisticated wavefunction-based methods are more accurate than the most sophisticated density-based methods. Nonetheless, it is believed that roughly 90% of modern quantum chemistry computations use one of the DFT schemes. What is it that makes DFT so attractive?

The answer lies in the cost/accuracy ratios of the various methodologies. The computational cost of the very sophisticated wavefunction-based methods (such as the popular CCSD(T) model) increases extremely quickly as the system grows. In fact, the cost of CCSD(T) scales with the *seventh* power of the size of the system and, not surprisingly, CCSD(T) can therefore be applied only to relatively small molecules. This has been the key weakness of wavefunction-based approaches and even spectacular improvements in computer speed will not significantly improve the situation.

In contrast, standard DFT calculations typically yield errors that are 2–3 times larger than those from CCSD(T) and yet their cost grows only with the *third* power of system size. To a pragmatic scientist, this combination of reasonable accuracy at a modest computational cost is irresistible.

The Q-CHEM Software

Although DFT leapt into prominence in the early '90s and was quickly adopted by the majority of computational chemists, a number of significant problems remained unsolved and, in 1993, Q-Chem Inc. was founded to address these. The company is currently owned by Peter Gill (University of Cambridge), Martin Head-Gordon (University of California, Berkeley) and Benny Johnson. The Board of Directors is growing but presently includes Dr Gill, Fritz Schaefer (University of Georgia) and John Stott (Eastside Holdings Ltd).

It was recognised from the onset that there were many reasons why it was best to plan the Q-CHEM package from scratch, rather than basing its structure and design on earlier programs. Modern software engineering practice argued that we should adopt an object-oriented programming language but the retention of FORTRAN was favoured on the grounds of compiler efficiency. As a result, the code is a mixture of high-level C++ and low-level FORTRAN routines. The code was designed to function efficiently on vector and superscalar architectures and, although the initial versions do not run on multiple processors, we were careful to employ algorithms and data structures that are well suited to the parallelisation efforts that we are now pursuing.

The first major problem that the Q-Chem collaboration sought to solve was the so-called Coulomb problem that arises when standard algorithms are used to compute the Coulomb energy

$$E_J = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \quad (5)$$

Since the density $\rho(\mathbf{r})$ is usually represented as a sum of gaussian functions, straightforward methods for the evaluation of (5) involve a double sum over all pairs of gaussians and therefore scale quadratically with system size.

Q-Chem Inc has received funding from the US Department of Energy and the National Science Foundation to conduct research into the design of radically new algorithms that compute E_J with a computational cost that scales only *linearly* with system size. These research efforts have been successful and, in 1994, the Q-Chem collaborators published the first algorithm, the CFMM, that computes the Coulomb energy of a molecular electron density in linear work. Such algorithms are known as $O(N)$ methods. More recently, the Cambridge team has developed another $O(N)$ Coulomb method, the KWIK algorithm, that can compute exchange, as well as Coulomb, energies.

Q-Chem Inc is also involved in the development of local correlation methods. This research, which is supported by the US Air Force, is attempting to develop modified versions of the sophisticated wavefunction-based approaches (such as Møller-Plesset perturbation theory) whose computational costs grow with system size more slowly than the original versions. The Berkeley group is devising well-defined localisation schemes that will recover a large fraction of the correlation energy to be in a small fraction of the time that a normal calculation would require.

Another of the problems associated with conventional DFT algorithms is their reliance on numerical grids to estimate the difficult integrals that arise when one uses any of the standard density functionals (such as B-LYP, PW91, etc). The Cambridge group is about to commence a major investigation into the development of “gridless” density functionals, that is, ones that can be integrated without the need for a molecular grid. This work will be funded by the National Science Foundation.

Density functional theory, in its original form, was a ground-state paradigm. However, there is great interest, of course, in methods that enable molecular excited states to be studied. The Berkeley group has been conducting ground-breaking research in this area and is also currently working on the development of appropriate extensions of DFT for excited states.

The development of new, more accurate density functionals is one of the most important areas of research in quantum chemistry today and Q-Chem is playing a role in this unfolding story. Two years ago, we introduced the G96 functional which, although much simpler than the popular Becke 88 functional, yields similar model chemistries. More recently, in collaboration with Prof Pople, we have developed the EDF1 functional which is superior to the widely used B3LYP functional when used with a basis set of medium size.

Inorganic chemists, who need to understand and predict the chemistry of metal-containing systems, require quantum chemical software that offers a highly efficient treatment of effective core potentials (ECPs) or pseudopotentials. The Cambridge group is currently developing a new algorithm for the calculation of ECP integrals, using a suitably modified version of the PRISM algorithm for two-electron integrals that they introduced some years ago.

The first parallel version of Q-CHEM is now in use and has shown impressive speed-up efficiencies on SGI Origin computers. Continued development of this aspect of the code is currently being supported by the National Institutes of Health and we expect that future versions of Q-CHEM will be fully able to exploit the enormous potential of modern parallel machines.

Q-Chem Inc. has recently released the Q-CHEM 1.2 package and it is expected that Q-CHEM 1.3 will be released in the summer of 1999. Important theoretical breakthroughs are being made every few months and the Q-Chem collaboration is building vigorously on the solid foundations that Kohn and Pople have laid.

Literature references, timing benchmarks and other information can be found on the Q-Chem web pages at <http://www.q-chem.com>