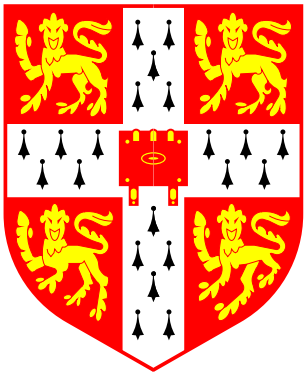
Novel Methods for Large Molecules in Quantum Chemistry

Ross David Adamson



A thesis submitted to the University of Cambridge
for the degree of Doctor of Philosophy

Trinity College

October 1, 1998

Declaration

This thesis describes research carried out in the Department of Chemistry at Massey University, New Zealand, and the University of Cambridge, between April 1995 and October 1998. The contents are the original work of the author except where indicated otherwise. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration unless stated otherwise.

Acknowledgements

During my Ph.D. I have been lucky enough to work in two universities, with very different departments, meeting a great deal of people who have helped me in my research. Throughout this time I have been extremely fortunate to have a supervisor of the calibre of Dr Peter Gill. For his constant encouragement, scientific guidance and friendship I am very grateful. I would also like to thank the patient supervision afforded to me by Professor John Pople, whose periodic visits proved very productive.

Help and comradeship from members of the "sweatbox" at Massey were invaluable, especially Dr Jeremy Dombroski's assistance with system administration and Dr Terry Adams' (not forgetting Gwen and Emily's) unending support, including short-notice accommodation in Pittsburgh.

I also owe a debt of thanks to the members of the Q-Chem group at Cambridge: Aaron, Katrina, Oliver, Andrew, Nikhil and Simon have all in some way contributed to this thesis. *Prosit!*

To all the gang at Trinity, particularly my three flatmates, who have all helped make my time at Cambridge an enjoyable one, I am most grateful. A special thank you to Sharon, who has not only helped immensely with proofreading, but has had to put up with the stress and bad moods associated with completing a Ph.D.

Finally, I would like to acknowledge the financial support from Massey University, the ORS awards scheme, the Cambridge Commonwealth Trust and Trinity College.

Abstract

With the rapid increase in computing power Quantum Chemists are looking towards larger and larger molecules. This thesis presents new ways to reduce the expensive scaling of computational cost with system size, thus allowing the advances in computer science to be utilized. The first chapter is an introduction to Hartree-Fock theory and the traditional methods of calculating electron correlation. This is followed by an introduction to Density Functional Theory, concentrating on Kohn-Sham density functional theory.

Chapter 3 presents a new way of assessing the accuracy of a density functional by partitioning the density and examining the energy of the component pieces. Chapter 4 describes a new density functional (EDF1), designed especially for small basis sets, thus making it ideal for large systems. The functional is formed from several other common functionals, grouped together in a way to minimize the error of the chemical energetics of a selection of molecules.

Chapter 5 gives an introduction to modern two-electron integral theory and then describes a new method for efficiently calculating integrals arising from charges that are well separated. The new algorithm does not scale with the contraction of the basis set.

The efficient algorithm of Chapter 5 is $O(N^2)$ overall, and therefore still too slow. To truly examine large molecules $O(N)$ methods are required. Chapter 6 provides an introduction to these linear methods and also presents a new method, the CASE approximation, which neglects long-range interactions. How to implement this new method (in $O(N)$ work) is described in Chapter 7. The method is extended to density functional theory in Chapter 8 by attenuating the Dirac functional. Chapter 9 presents a second way to reduce the magnitude (and speed) of the approximation, and also a correction for the main failure of the original approximation. Chapter 10 examines the accuracy of the approximation on a variety of chemical properties. The final chapter describes a way to improve the accuracy of CASE by correcting for the neglected terms in only $O(N)$ work. This correction however is not without its own problems and work continues in this area.

List of Publications

The following papers have been published as a direct consequence of the work undertaken for this thesis:

1. "Density functional partitions" John A. Pople, Ross D. Adamson and Peter M. W. Gill. *Journal of Physical Chemistry*, (1996) **100**, 6348.

2. "Chemistry without Coulomb tails" Ross D. Adamson, Jeremy P. Dombroski and Peter M. W. Gill. *Chemical Physics Letters*, (1996) **254**, 329.

3. "Coulomb-attenuated exchange energy density functionals" Peter M. W. Gill, Ross D. Adamson and John A. Pople. *Molecular Physics*, (1996) **88**, 1005.

4. "A family of attenuated Coulomb operators" Peter M. W. Gill and Ross D. Adamson. *Chemical Physics Letters*, (1996) **261**, 105.

5. "Effects of Coulomb attenuation on chemical properties" Ross D. Adamson and Peter M. W. Gill. *Journal of Molecular Structure (Theochem)*, (1997) **398** 45.

6. "A tensor approach to two-electron matrix elements" Terry R. Adams, Ross D. Adamson and Peter M. W. Gill. *Journal of Chemical Physics*, (1997) **107**, 124.

7. "Empirical density functionals" Ross D. Adamson, Peter M. W. Gill and John A. Pople. *Chemical Physics Letters*, (1998) **284** 6.

8. "Efficient calculation of short-range Coulomb energies" Ross D. Adamson, Jeremy P. Dombroski and Peter M. W. Gill. *Journal of Computational Chemistry*, submitted.

9. "Empirical density functionals dependence on basis set" John A. Pople, Ross D. Adamson and Peter M. W. Gill. To be submitted.

10. "Coulombic energies by analytic continuation" Aaron M. Lee, Ross D. Adamson and Peter M. W. Gill. In preparation.

11. "Accuracy of the CASE approximation" Ross D. Adamson and Peter M. W. Gill. In preparation.

Contents

Chapter 1

Introduction

> "I think it is true to say that nobody understands quantum mechanics" - Richard Feynman

1.1 The Schrödinger Equation

It can be argued that the field of quantum chemistry began in 1925, with de Broglie's postulate that electrons possess both wave-like and particle-like characteristics [1]. Later in 1925, Davisson and Germer [2] confirmed de Broglie's conjecture experimentally by producing a diffraction pattern of electrons consistent with the de Broglie relation. Schrödinger then expanded on de Broglie's work, forming the non-relativistic Schrödinger wave equation [3], a mathematical model powerful enough to describe all non-relativistic chemical phenomena. This idea was not fully understood, however, until Heisenberg introduced the Uncertainty Principle (that it is impossible to specify both the linear momentum and position of a particle to arbitrary precision) [4].

The time-dependent Schrödinger equation is

$$\hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian operator, representing all energy contributions of the system. The wavefunction, Ψ, is a function of the nuclear and electron positions, electron spins and time. It is easiest viewed using the Born interpretation [5]:

> The probability that a system, described by Ψ, will be found in a given state is proportional to $\Psi^*\Psi$.

This interpretation generates a constraint that the wavefunction must be square integrable, that is

$$\int |\Psi|^2 \, d\tau < \infty. \tag{1.2}$$

If we consider a system of M nuclei (each of charge Z_A) and N electrons in the absence of any external field, the Hamiltonian, in atomic units, is given by

$$\hat{H} = \hat{T} + \hat{V}, \tag{1.3}$$

where

$$\hat{T} = -\frac{1}{2}\sum_{A}^{M} \nabla_A^2 - \frac{1}{2}\sum_{i}^{N} \nabla_i^2 \,, \tag{1.4}$$

representing the kinetic energy of the nuclei and electrons respectively, and

$$\hat{V} = -\sum_{A}^{M}\sum_{i}^{N} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_{A}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} + \sum_{i}^{N}\sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{1.5}$$

describing the potentials due to nuclear-electron attraction, nuclear-nuclear repulsion and electron-electron repulsion.

The Hamiltonian above is independent of time, which allows a separation of variables. The wavefunction can then be of the form

$$\Psi(t) = \Psi e^{\frac{-iEt}{\hbar}} \tag{1.6}$$

where Ψ does not depend on time, and the Schrödinger equation reduces to

$$\hat{H}\Psi = E\Psi, \tag{1.7}$$

also independent of time. This is an important result: if the potential is independent of time and the system is in a state of energy E, all that is required to construct the time-dependent wavefunction from the time-independent wavefunction is multiplication by $e^{-iEt/\hbar}$, which is simply a modulation of its phase.

1.2 The Born-Oppenheimer Approximation

Perhaps the great disappointment of quantum chemistry is that, whilst the Schrödinger equation *is* powerful enough to describe almost all chemistry, it is too complex to solve for all but the simplest of systems. Even the simplest molecule, H_2^+, consists of three particles,

thus producing a Schrödinger equation that is impossible to solve analytically. To overcome this difficulty a variety of approximations are made, the most common of which is the Born-Oppenheimer approximation [6].

The masses of the nuclei are much greater than the electrons, hence the electrons can respond almost instantaneously to any change in the nuclear positions. Thus, to a good approximation, we can think of the electrons as moving in a field of fixed nuclei. Within this approximation, the nuclear kinetic energy term can be neglected and the nuclear-nuclear repulsion term can be considered a constant. These two terms can therefore be removed[1] to form the electronic Hamiltonian, $\hat{H}_{elec}$.

What remains is termed the electronic Schrödinger equation,

$$\hat{H}_{elec}\,\Psi_{elec}(\mathbf{r}_i;\mathbf{R}_A) = E_{elec}(\mathbf{R}_A)\,\Psi_{elec}(\mathbf{r}_i;\mathbf{R}_A). \tag{1.8}$$

The notation $\Psi_{elec}(\mathbf{r}_i;\mathbf{R}_A)$ implies that the electronic wavefunction depends on the nuclear positions only parametrically — a different wavefunction is defined for each nuclear configuration.

$E_{elec}(\mathbf{R}_A)$ is only the electronic energy; to regain the total energy (for fixed nuclei) we must add the nuclear-nuclear repulsion constant.

$$E_{tot} = E_{elec} + \sum_{A}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{1.9}$$

Repeating the calculation with a different nuclear arrangement allows the potential energy surface to be mapped out and the equilibrium geometry to be found. The work of this thesis is entirely within the Born-Oppenheimer approximation, so for clarity, the 'tot' and 'elec' subscripts will be dropped and only electronic Hamiltonians and wavefunctions will be considered.

1.3 The Hartree Approximation

Although the Born-Oppenheimer approximation considerably reduces the complexity of the Schrödinger equation, the resulting electronic Schrödinger equation is still extremely complex, due to the electron-electron interactions. It is possible to use wavefunctions which explicitly include inter-electronic distance [7–9], but this approach is computationally infeasible for all but the smallest systems.

[1] A constant added to an operator does not affect the wavefunction, it only adds to the operator eigenvalue.

A more satisfying solution is to introduce the molecular orbital approximation, the simplest of which is the independent-particle, or Hartree, approximation [10–12] wherein the total wavefunction is approximated by a product of orthonormal molecular orbitals (MOs). This idea closely follows the chemists' view of electrons occupying orbitals. The Hartree approximation assumes that each electron moves independently within its own orbital and sees only the *average* field generated by all the other electrons. The Hartree wavefunction (for an N electron system) is

$$\Psi = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\ldots\chi_{N-1}(\mathbf{x}_{N-1})\chi_N(\mathbf{x}_N), \tag{1.10}$$

where each χ_i is a spin orbital containing one electron. The χ_i are orthonormal, consisting of a spatial orbital, $\psi_i(\mathbf{r})$, and one of two spin functions, $\alpha(s)$ and $\beta(s)$, representing spin up and spin down states. $\mathbf{x}$ is the space-spin coordinate, containing both the position, $\mathbf{r}$, and spin, s, of a particle.

1.4 The Variational Method

By rearranging equation (1.7) it is possible to obtain an expression for the energy, E_{app}, associated with an approximate wavefunction, Ψ_{app}.

$$E_{app} = \frac{\int \Psi^*_{app}\hat{H}\Psi_{app}\,d\tau}{\int \Psi^*_{app}\Psi_{app}\,d\tau} = \frac{\langle\Psi_{app}|\hat{H}|\Psi_{app}\rangle}{\langle\Psi_{app}|\Psi_{app}\rangle} \tag{1.11}$$

This approximate wavefunction can also be written as a linear combination of the exact eigenstates of $\hat{H}$:

$$\Psi_{app} = \sum_i c_i\Psi_i. \tag{1.12}$$

Now, consider the integral

$$\begin{aligned}\int \Psi^*_{app}(\hat{H}-E_0)\Psi_{app}\,d\tau &= \sum_i\sum_{i'} c_i^*c_{i'}\int \Psi_i^*(\hat{H}-E_0)\Psi_{i'}\,d\tau \\ &= \sum_i\sum_{i'} c_i^*c_{i'}(E_{i'}-E_0)\int \Psi_i^*\Psi_{i'}\,d\tau \\ &= \sum_i c_i^*c_i(E_i-E_0) \geq 0\end{aligned} \tag{1.13}$$

as $|c_i|^2 \geq 0$ and $E_i \geq E_0$. Therefore,

$$\int \Psi^*_{app}(\hat{H}-E_0)\Psi_{app}\,d\tau \geq 0, \tag{1.14}$$

from which follows the Variational Principle, that $E_{app} \geq E_0$. This also gives insight into how to find the best approximate wavefunction, as

$$E_{app} = E_0 \iff \Psi_{app} = \Psi_0. \tag{1.15}$$

Hence, minimization of E_{app} with respect to all allowed Ψ_{app} will give the exact ground state energy and wavefunction. Unfortunately, this is not practical and what is usually done is to expand the molecular orbitals as *linear* combinations of basis functions,

$$\psi_i = \sum_{\mu} c_{\mu i} \phi_{\mu}, \tag{1.16}$$

reducing the problem to finding the optimum set of coefficients, $c_{\mu i}$, which minimize E_{app}, something which can be achieved via matrix diagonalisation (see section 1.5). Obviously, the set of functions ϕ_μ cannot be complete and so an approximation has been made. The number and type of functions chosen has a large effect on the overall accuracy (and speed) of a calculation. Some of the more common functions used will be given in the section on Basis Sets (section 1.7).

1.5 Hartree-Fock Theory

In 1930 Fock [13] pointed out that the Hartree wavefunction was invalid as it did not satisfy the Pauli Exclusion Principle — that the wavefunction must be antisymmetric with respect to electron interchange [14]. Fock also showed that a Hartree product could be made antisymmetric by appropriately adding and subtracting all possible permutations of the Hartree product, thereby forming the Hartree-Fock (HF) wavefunction. Later, Slater showed that the resulting wavefunction is simply the determinant of a matrix, called a Slater determinant [15, 16]

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}. \tag{1.17}$$

The prefactor normalizes the HF wavefunction (remembering that the χ_i are orthonormal).

1.5.1 Matrix Elements and Notation

Before presenting the HF equations it is useful to define some matrix elements common to quantum chemistry. The overlap matrix, $\mathbf{S}$, represents the overlap of basis functions

$$S_{\mu\nu} = \int \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\, d\mathbf{r}, \tag{1.18}$$

A familiar concept to chemists is the electron density, $\rho(\mathbf{r})$, which can be obtained by integrating the square of the wavefunction over all N electrons but one, and all spin variables.

$$\rho(\mathbf{r}) = N \int \Psi^2\, ds_1\, d\mathbf{x}_2\, d\mathbf{x}_3\, \ldots\, d\mathbf{x}_N \tag{1.19}$$

which is simply the square of the sum over all the occupied orbitals:

$$\begin{aligned} \rho(\mathbf{r}) =& \sum_i^{occ} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}) \\ =& \sum_\mu \sum_\nu \left(\sum_i^{occ} C_{\mu i} C_{\nu i} \right) \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}). \end{aligned} \tag{1.20}$$

The bracketed term is frequently required in solving the HF equations, and hence is usually precomputed and stored as the density matrix, $\mathbf{P}$

$$P_{\mu\nu} = \sum_i^{occ} C_{\mu i} C_{\nu i}. \tag{1.21}$$

The one-electron operators are also used to form their own matrices, $\mathbf{T}$ and $\mathbf{V}$.

$$T_{\mu\nu} = \int \phi_\mu(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 \right) \phi_\nu(\mathbf{r})\, d\mathbf{r} \tag{1.22}$$

$$V_{\mu\nu} = \int \phi_\mu(\mathbf{r}) \left(-\sum_A \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} \right) \phi_\nu(\mathbf{r})\, d\mathbf{r}. \tag{1.23}$$

These two matrices are added together to form the core Hamiltonian matrix, $\mathbf{H}$, where $H_{\mu\nu}$ represents the energy of an isolated electron in the distribution $\phi_\mu\phi_\nu$.

The repulsion between an electron in the MO distribution $\psi_i\psi_j$ and the MO distribution $\psi_k\psi_l$ has a shorthand notation:

$$(ij|kl) = \iint \frac{\psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\psi_k^*(\mathbf{r}_2)\psi_l(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_1 d\mathbf{r}_2, \tag{1.24}$$

where

$$r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|. \tag{1.25}$$

This shorthand also exists for repulsions between basis function distributions:

$$(\mu\nu|\lambda\sigma) = \iint \frac{\phi^*_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)\phi^*_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_1 d\mathbf{r}_2. \tag{1.26}$$

The relationship between the two is, of course

$$(ij|kl) = \sum_{\mu\nu\lambda\sigma} C_{\mu i} C_{\nu j} C_{\lambda k} C_{\sigma l} (\mu\nu|\lambda\sigma). \tag{1.27}$$

1.5.2 Initial Guess

HF Theory assumes that an electron moves in a potential which is the average of the potentials due to all the other electrons and nuclei, so a trial wavefunction is required for the potential *before* the energy can be calculated. The optimal wavefunction is then found by iteratively solving the Schrödinger equation. This initial guess must be close enough to the optimum wavefunction for the equations to converge to the correct electronic state.

There are several ways to obtain a guess for the wavefunction, the simplest of which is to guess the MO coefficients by diagonalising the core Hamiltonian matrix [17]. In 1952 Wolfsberg and Helmholtz presented a more sophisticated procedure which also uses the overlap matrix [18]

$$H_{\mu\nu} = c_x S_{\mu\nu}(H_{\mu\mu} + H_{\nu\nu})/2 \tag{1.28}$$

where c_x is a constant. Obviously, with the guess playing such a crucial role in the efficiency of HF theory, it is an area of much research; one of the latest (and more successful) methods uses a superposition of atomic densities (SAD) [19].

1.5.3 Restricted Closed Shell Hartree-Fock

In the restricted formalism each spatial orbital contains two electrons, one spin up, the other spin down. That is,

$$\begin{aligned} \chi_{2i-1}(\mathbf{x}) &= \psi_i(\mathbf{r})\alpha(s) \\ \chi_{2i}(\mathbf{x}) &= \psi_i(\mathbf{r})\beta(s). \end{aligned} \tag{1.29}$$

Using these orbitals the HF energy is, in terms of the spatial components of molecular orbitals (remembering that the wavefunction is normalized and that the orbitals are orthonormal),

$$\begin{aligned} E =& \langle\Psi|\hat{H}|\Psi\rangle \\ =& \sum_i^{N/2} 2(\psi_i|\hat{h}|\psi_i) + \sum_{ij}^{N/2}[2(\psi_i\psi_i|\psi_j\psi_j) - (\psi_i\psi_j|\psi_i\psi_j)] \end{aligned} \tag{1.30}$$

where $\hat{h}$ represents the one-electron operators and $2\sum(\psi_i\psi_i|\psi_j\psi_j)$ is the Coulombic repulsion between all electrons. Note that this term includes a spurious self-repulsion. The final term of equation (1.30), $-\sum(\psi_i\psi_j|\psi_i\psi_j)$, has arisen from making the wavefunction antisymmetric. It is termed the exchange energy and has no classical analogue. Most importantly, the exchange term contains elements which exactly cancel the spurious self-interaction of the Coulomb energy. Another important effect of the exchange term is that, while an electron feels only the average field of all other electrons, it does feel an instantaneous effect of all electrons of the same spin. That is, the probability of finding two electrons at the same point at the same time is non-zero, but the probability of finding two electrons at the same point (at the same time) of the same spin *is* zero.

To find the orbitals which minimize the energy we make the energy stationary with respect to variations of the MO coefficients, $C_{\mu i}$. If there are m basis functions and n occupied orbitals, ψ_i, then solving the Schrödinger equation will produce $(m-n)$ unoccupied (or virtual) orbitals, ψ_a, which obey $(\psi_a|\psi_i) = 0$ (with the standard notation of using i, j to denote occupied orbitals; a, b for virtual; and p, q to denote any MOs).

At the minimum the energy is stationary with respect to the variation

$$\psi_i \rightarrow \psi_i + \lambda\psi_a \qquad (i = 1, \ldots, n; a = n+1, \ldots, m). \tag{1.31}$$

This variation preserves orbital orthonormality through first order in λ. Substituting equation (1.31) into equation (1.30), picking out the coefficient of λ and setting it to zero yields the stationary condition;

$$(\psi_a|\hat{h}|\psi_i) + \sum_j [2(ai|jj) - (aj|ij)] = 0 \tag{1.32}$$

It is easier to use these equations when expressed in operator form. We define the Fock operator such that

$$\begin{aligned} F_{ai} =& (\psi_a|\hat{F}|\psi_i) \\ =& (\psi_a|\hat{h}|\psi_i) + \sum_j [2(ai|jj) - (aj|ij)] = 0 \end{aligned} \tag{1.33}$$

or, in operator form

$$\hat{F}(\mathbf{r}_1) = \hat{h}(\mathbf{r}_1) + 2\hat{J}(\mathbf{r}_1) - \hat{K}(\mathbf{r}_1), \tag{1.34}$$

where the Coulomb operator, $\hat{J}$, is

$$\hat{J}(\mathbf{r}_1) = \sum_j \int \frac{\psi_j^2(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_2 \tag{1.35}$$

and the exchange operator, $\hat{K}$, is

$$\hat{K}(\mathbf{r}_1) = \sum_j \int \frac{\psi_j(\mathbf{r}_2)\psi_j(\mathbf{r}_1)}{r_{12}} \hat{P}_{12}\, d\mathbf{r}_2, \tag{1.36}$$

where $\hat{P}_{12}$ is the permutation operator, that is

$$\hat{P}_{12}\psi_p(\mathbf{r}_1) = \psi_p(\mathbf{r}_2). \tag{1.37}$$

The Fock operator is therefore an effective one-electron Hamiltonian.

Orbitals that satisfy the condition $F_{ai} = 0$ are obtained by solving the Roothaan-Hall equations [20, 21]

$$\sum_\nu (\phi_\mu|\hat{F} - \varepsilon_i|\phi_\nu) C_{\nu i} = 0. \tag{1.38}$$

The resulting orbitals will not only satisfy $F_{ai} = 0$, but also

$$F_{ij} = \varepsilon_i \delta_{ij}. \tag{1.39}$$

This, however, does not matter as the SCF energy is invariant to a mixing of the occupied orbitals. When the Fock matrix is completely diagonal the orbitals are termed canonical.

1.5.4 Restricted Open Shell Hartree-Fock

The Roothaan-Hall equations are unsuitable for an open-shell system, and require some modification. One approach is to allow some orbitals to contain only an electron of α spin [22, 23]. Under such a scheme the energy expression becomes

$$\begin{aligned} E = &\sum_i 2(\psi_i|\hat{h}|\psi_i) + \sum_{ij}[2(\psi_i\psi_i|\psi_j\psi_j) - (\psi_i\psi_j|\psi_i\psi_j)] \\ &+ \sum_s (\psi_s|\hat{h}|\psi_s) + \frac{1}{2}\sum_{st}[(\psi_s\psi_s|\psi_t\psi_t) - (\psi_s\psi_t|\psi_s\psi_t)] \\ &+ \sum_{is}[2(\psi_s\psi_s|\psi_i\psi_i) - (\psi_i\psi_s|\psi_i\psi_s)] \end{aligned} \tag{1.40}$$

where i, j denote doubly occupied orbitals and s, t denote singly occupied orbitals.

Proceeding as before, we consider the variations

$$\begin{aligned} \psi_i &\rightarrow \psi_i + \lambda\psi_a \\ \psi_s &\rightarrow \psi_s + \lambda\psi_a \\ \psi_i &\rightarrow \psi_i + \lambda\psi_s \\ \psi_s &\rightarrow \psi_s - \lambda\psi_i. \end{aligned} \tag{1.41}$$

Substituting these into equation (1.40) and minimizing the energy gives the self consistent conditions

$$\begin{aligned} F_{ai} &= 0 \\ F_{sa} - \frac{1}{2}K^O_{sa} &= 0 \\ F_{si} + \frac{1}{2}K^O_{si} &= 0 \end{aligned} \tag{1.42}$$

where

$$\mathbf{F} = \mathbf{H} + 2\mathbf{J}^C - \mathbf{K}^C + \mathbf{J}^O - \frac{1}{2}\mathbf{K}^O \tag{1.43}$$

with the superscripts C and O denoting summation over closed shell and open shell orbitals respectively. From this we can define Fock matrices for the α and β electrons

$$\mathbf{F}^\alpha = \mathbf{F} - \frac{1}{2}\mathbf{K}^O \tag{1.44}$$

$$\mathbf{F}^\beta = \mathbf{F} + \frac{1}{2}\mathbf{K}^O. \tag{1.45}$$

Orbitals can then be found that satisfy the conditions (1.42) by diagonalisation of the block matrix

$$\begin{pmatrix} (\mathbf{F} - \mathbf{K}^O) & \mathbf{F}^\beta & \frac{1}{2}(\mathbf{F}^\alpha + \mathbf{F}^\beta) \\ \mathbf{F}^\beta & \mathbf{F} & \mathbf{F}^\alpha \\ \frac{1}{2}(\mathbf{F}^\alpha + \mathbf{F}^\beta) & \mathbf{F}^\alpha & (\mathbf{F} - \mathbf{K}^O) \end{pmatrix} \tag{1.46}$$

where the three blocks refer to doubly occupied, singly occupied and virtual orbitals. The diagonal blocks do not affect the stationary conditions, so can be defined in any desired way. Roothaan's original definition has been used above.

The ROHF scheme places an unphysical constraint on the wavefunction. α and β electrons in an open shell molecule may feel different potentials, yet their spatial orbitals are constrained to be the same. This has the effect of raising the variational energy. The ROHF wavefunction must also be of high spin.

1.5.5 Unrestricted Hartree-Fock

A simpler way to expand RHF to open shell systems is to introduce separate spatial orbitals for electrons of α and β spin:

$$\begin{aligned} \chi_{2i-1}(\mathbf{x}) &= \psi^\alpha_i(\mathbf{r})\alpha(s) \\ \chi_{2i}(\mathbf{x}) &= \psi^\beta_i(\mathbf{r})\beta(s). \end{aligned} \tag{1.47}$$

This allows Hartree-Fock theory to give excess β electron density at points in the molecule, something which has been seen in experiment and is only possible in RHF if the wavefunction is expanded beyond a single determinant. By the variational principle the UHF energy will be lower than (or equal to) the RHF energy.

These orbitals lead to two density matrices,

$$\begin{aligned} P^{\alpha}_{\mu\nu} &= \sum_{i}^{occ} C^{\alpha}_{\mu i} C^{\alpha}_{\nu i} \\ P^{\beta}_{\mu\nu} &= \sum_{i}^{occ} C^{\beta}_{\mu i} C^{\beta}_{\nu i} \end{aligned} \tag{1.48}$$

and two Fock operators,

$$\begin{aligned} \hat{F}^{\alpha}(\mathbf{r}_1) &= \hat{h}(\mathbf{r}_1) + \hat{J}^{\alpha}(\mathbf{r}_1) + \hat{J}^{\beta}(\mathbf{r}_1) - \hat{K}^{\alpha}(\mathbf{r}_1) \\ \hat{F}^{\beta}(\mathbf{r}_1) &= \hat{h}(\mathbf{r}_1) + \hat{J}^{\alpha}(\mathbf{r}_1) + \hat{J}^{\beta}(\mathbf{r}_1) - \hat{K}^{\beta}(\mathbf{r}_1) \end{aligned} \tag{1.49}$$

which are combined to form the Pople-Nesbet [24] equations,

$$\begin{aligned} \sum_{\nu} (\phi_\mu | \hat{F}^{\alpha} - \varepsilon^{\alpha}_i | \phi_\nu) C^{\alpha}_{\nu i} = 0 \\ \sum_{\nu} (\phi_\mu | \hat{F}^{\beta} - \varepsilon^{\beta}_i | \phi_\nu) C^{\beta}_{\nu i} = 0, \end{aligned} \tag{1.50}$$

the solution of which gives the molecular orbitals.

1.5.6 Spin Properties of Hartree-Fock Wavefunctions

The spin operators $\hat{S}_z$ and $\hat{S}^2$ both commute with the non-relativistic Hamiltonian, and therefore eigenfunctions of the Hamiltonian can be found which are also eigenfunctions of these spin operators. The permutation operator (equation (1.37)) commutes with $\hat{S}_z$ so single determinants are eigenfunctions of $\hat{S}_z$. Unfortunately this is not the case for the $\hat{S}^2$ operator. It can be shown [17] that

$$\langle \hat{S}^2 \rangle = \left(\frac{N^{\alpha} - N^{\beta}}{2} \right) \left(\frac{N^{\alpha} - N^{\beta}}{2} + 1 \right) + N^{\beta} - \sum_{ij} |S^{\alpha\beta}_{ij}|^2, \tag{1.51}$$

where N^{α} and N^{β} are the number of α and β electrons ($N^{\alpha} \geq N^{\beta}$), and

$$S^{\alpha\beta}_{ij} = \int \psi^{\alpha}_i \psi^{\beta}_j \, d\mathbf{r}. \tag{1.52}$$

For RHF (open and closed shell) the occupied β orbitals lie within the α orbital's space; therefore,

$$N^{\beta} = \sum_{ij} |S^{\alpha\beta}_{ij}|^2. \tag{1.53}$$

Thus the determinants are eigenfunctions of $\hat{S}^2$. However, for unrestricted determinants, the β orbitals are not constrained to lie within the α space; therefore,

$$N^\beta \geq \sum_{ij} |S_{ij}^{\alpha\beta}|^2. \tag{1.54}$$

These determinants will not be eigenfunctions of $\hat{S}^2$ and are termed spin-contaminated — they contain higher spin multiplicity components. This spin-contamination can allow the UHF function to give the correct dissociation behaviour, as the α and β electrons are no longer forced to occupy the same orbital. However, for methods which build on Hartree-Fock, spin-contamination can have a disastrous effect [25–28].

1.5.7 The cost of HF Theory

The bottleneck for HF calculations is the generation of all the two-electron integrals in the atomic basis, $(\mu\nu|\lambda\sigma)$. The number of these grows as $O(N^4)$, where N is the total number of basis functions in the system. However, this scaling can be drastically reduced by viewing $(\mu\nu|\lambda\sigma)$ as the repulsion between two shell-pairs,

$$(\mu\nu| = \phi_\mu^*(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1) \tag{1.55}$$

and

$$|\lambda\sigma) = \phi_\lambda^*(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2). \tag{1.56}$$

If either $(\mu\nu|$ or $|\lambda\sigma)$ is so small that it is negligible, the integral $(\mu\nu|\lambda\sigma)$ will be negligible. If the two shells of a shell-pair are very far apart (relative to their diffuseness) then their overlap will produce a negligible shell-pair. The number of non-negligible shell-pairs grows only *linearly* with the size of the system. So, forming only the significant shell-pairs will generate a number of integrals which grows only quadratically with molecular size. Hence the cost of HF theory is $O(N^2)$.

There is a separate $O(N^3)$ cost involved in diagonalising the Fock matrix, yet this scaling does not become noticeable until the system is extremely large, and is not expected to be a problem in the medium term [29]. Recent techniques have made HF theory scale only linearly with molecular size for certain types of systems; these are dealt with in later chapters.

1.6 Multiple Determinant Wavefunctions

The main deficiency of HF theory is the inadequate treatment of the correlation between motions of electrons. No account is made for the correlation between electrons of opposite

spin. Electrons of the same spin are partially, but not completely, correlated. This leads to a HF energy above the exact non-relativistic value. This difference is defined to be the correlation energy [30]:

$$E_{corr} = E_{exact} - E_{HF}. \tag{1.57}$$

Unfortunately, the energy change in a reactive chemical processes is often of the same magnitude as the correlation energy, and the correlation energy can change markedly for many chemical processes, especially those where the number of electron pairs change. Hence, HF theory performs well for isodesmic reactions and for locating equilibrium structures (bond lengths usually to within 0.01Å and angles with 1°) [31]. Vibrational frequencies are usually within 10%. For relative energies, however, more accurate calculations are often required.

The are a number of techniques that seek to improve on the HF wavefunction. The method of choice depends very much on the characteristics of the problem. Some of the desirable features of a method involving electron correlation are [32]:

1. it should be well defined, giving a continuous potential surface and a unique energy for any nuclear configuration.
2. it should be "size consistent": the energy of a sum of non-interacting fragments should be exactly the sum of separate calculations on the fragments.
3. it should be exact when applied to a two-electron system.
4. it should be efficient, computational cost scaling slowly with system size.
5. it should be accurate enough to be an adequate approximation to the exact result.
6. it should be variational: that is, the energy is an upper bound to the exact result.

Unfortunately, no current method satisfies all of the above criteria!

1.6.1 Configuration Interaction

To correctly describe the instantaneous interaction of electrons, the inter-electron distance must be introduced. The most conceptually simple way of achieving this is via Configuration Interaction (CI) [33, 34]. CI uses a wavefunction which is a linear combination of the HF determinant and determinants from excitations of electrons,

$$\Psi = C_0\Psi_0 + \sum_i \sum_a C_i^a \Psi_i^a + \sum_{i>j} \sum_{a>b} C_{ij}^{ab} \Psi_{ij}^{ab} + \sum_{i>j>k} \sum_{a>b>c} C_{ijk}^{abc} \Psi_{ijk}^{abc} + \dots \tag{1.58}$$

where Ψ_i^a represents the determinant with an electron excited from the occupied orbital, χ_i to the virtual orbital, χ_a.

The CI expansion is variational and, if the expansion is complete (Full CI), gives the exact correlation energy (within the basis set approximation). The number of determinants in Full CI grows exponentially with the system size, making the method impractical for all but the smallest systems. For this reason the CI expansion is usually truncated at some order, for example CISD, where only singly and doubly excited determinants are considered. Brillouin's Theorem states that singly excited determinants do not mix with the HF determinant [35]. Therefore CISD is the cheapest worthwhile form of CI, yet this method scales as $O(N^6)$ where N is the size of the system.

The other main problem with truncated CI is that it is not size consistent. For CISD, an approximate way to correct for these effects is to introduce the Davidson correction [36]

$$E_{corr} = E_{corr}(\text{CISD}) + (1 - c_0^2)E_{corr}(\text{CISD}) \tag{1.59}$$

where c_0 is the coefficient of the Hartree-Fock wavefunction in the normalized CISD wavefunction.

1.6.2 Quadratic Configuration Interaction

A more acceptable way to make truncated CI size consistent was introduced by Pople *et al.* in 1987 [32]. Termed Quadratic Configuration Interaction (QCISD), it is formed by the addition of higher excitation terms, quadratic in the expansion coefficients, which force size-consistency. The addition of extra terms has not destroyed the correctness for two electrons (property 3 above). At the same time a perturbative treatment of the triple excitations was proposed, giving rise to QCISD(T) theory. This addition has proven to be worthwhile [37] and QCISD(TQ) has even been proposed to include quadruple excitations [38]. QCISD scales as $O(N^6)$ while QCISD(T) requires one iteration of $O(N^7)$.

1.6.3 Coupled Cluster Theory

The theoretical framework of Coupled Cluster (CC) theory was developed in the late 1960s [39, 40], but it was not until the late 1970s that the practical implementation began to take place and until 1982 that the corner stone of modern implementation, CCSD [41] (CC including all single and double excitations), was presented.

CC solves the size consistency problem of CI by forming a wavefunction where the excitation operators are exponentiated,

$$\Psi_{CC} = \exp(\hat{T})\Psi_0 \tag{1.60}$$

where

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots \tag{1.61}$$

and $\hat{T}_n$ is a linear combination of all n-type excitations, for example,

$$\begin{aligned} \hat{T}_1\Psi_0 &= \sum_i \sum_a C_i^a \Psi_i^a \\ \hat{T}_2\Psi_0 &= \sum_{i>j} \sum_{a>b} C_{ij}^{ab} \Psi_{ij}^{ab} \end{aligned} \tag{1.62}$$

where C_i^a and C_{ij}^{ab} are the coefficients to be determined. Substituting equation (1.62) into equation (1.60) yields the CCSD wavefunction

$$\begin{aligned} \Psi_{CCSD} = \Psi_0 + &\sum_a \sum_i C_i^a \Psi_i^a + \sum_{a>b} \sum_{i>j} C_{ij}^{ab} \Psi_{ij}^{ab} \\ &+ \frac{1}{2} \sum_{ab} \sum_{ij} C_i^a C_j^b \Psi_{ij}^{ab} + \frac{1}{2} \sum_{a>b} \sum_{c>d} \sum_{i>j} \sum_{k>l} C_{ab}^{ij} C_{cd}^{kl} \Psi_{ijkl}^{abcd} + \ldots \end{aligned} \tag{1.63}$$

This reveals the advantage of CC theory: higher excitations are partially included, but their coefficients are determined by the lower order excitations. The coefficients are determined by projecting Schrödinger's equation on the left with the configurations generated by the $\hat{T}$ operator. This replaces the eigenvalue problem by a non-linear simultaneous system, requiring iterative solution. Luckily, convergence is fast in most cases [42].

As mentioned above, the addition of pure triple excitations is required for some chemical problems. However, CCSDT scales as $O(N^8)$, which is impractical for all but the simplest of systems. A more practical alternative is CCSD(T) [37] where the effect of triples is estimated through perturbation theory with a non-iterative $O(N^7)$ cost.

With a large enough basis set CCSD typically recovers 95% of the correlation energy for a molecule at equilibrium geometry, while CCSD(T) sees a further five- to ten-fold reduction in error [43]. With such accuracy CC has become the method of choice for accurate small-molecule calculations, even though the method is not variational (property 6, above).

A method closely related to CCSD is Brueckner Doubles (BD) [44], which uses the Brueckner orbitals [45] rather than the HF orbitals for a CCSD treatment. The Brueckner orbitals are defined as the set of orbitals for which the single excitation coefficients are zero. Finding

these orbitals makes the theory slightly more computationally intensive (BD and BD(T) still scale as $O(N^6)$ and $O(N^7)$ respectively). However, BD theory promises a slight increase in accuracy above CCSD.

1.6.4 Møller-Plesset Perturbation Theory

Møller-Plesset Perturbation theory [46, 47] treats the exact Hamiltonian, $\hat{H}$, as a *small* perturbation from the HF Hamiltonian, $\hat{H}_0$ — the sum of the one-electron Fock operators defined by equation (1.34). That is,

$$\hat{H} = \hat{H}_0 + \lambda\hat{V}. \tag{1.64}$$

If we expand the exact energy and wavefunction in terms of the perturbation

$$\begin{aligned} \Psi_i &= \Psi_i^{(0)} + \lambda\Psi_i^{(1)} + \lambda^2\Psi_i^{(2)} + \dots \\ E_i &= E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots , \end{aligned} \tag{1.65}$$

where $\Psi_i^{(n)}$ is the n-th state HF wavefunction. Substituting equations (1.65) into the Schrödinger equation and collating the powers of λ gives the equations:

$$\hat{H}_0\Psi_i^{(0)} = E_i^{(0)} \tag{1.66}$$

$$\hat{H}_0\Psi_i^{(1)} + \hat{V}\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(1)} + E_i^{(1)}\Psi_i^{(0)} \tag{1.67}$$

$$\hat{H}_0\Psi_i^{(2)} + \hat{V}\Psi_i^{(1)} = E_i^{(0)}\Psi_i^{(2)} + E_i^{(1)}\Psi_i^{(1)} + E_i^{(2)}\Psi_i^{(0)} \tag{1.68}$$

and so on. Multiplying each of these equations on the left by Ψ_0 and integrating over all space yields expressions for $E^{(n)}$ in terms of $\hat{V}$ and $\Psi^{(n-1)}$:

$$E_i^{(0)} = \langle\Psi_i^{(0)}|\hat{H}_0|\Psi_i^{(0)}\rangle \tag{1.69}$$

$$E_i^{(1)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(0)}\rangle \tag{1.70}$$

$$E_i^{(2)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(1)}\rangle \tag{1.71}$$

$$E_i^{(3)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(2)}\rangle \tag{1.72}$$

and so on. From this it can be seen that the HF energy is the sum of $E_0^{(0)}$ and $E_0^{(1)}$.

By using the expansion

$$\Psi_i^{(1)} = \sum_n c_n^{(1)}\Psi_n^{(0)} \tag{1.73}$$

in equation (1.67) and rearranging, an expression for the coefficients can be found:

$$c_n^{(1)} = -\frac{\langle\Psi_n^{(0)}|\hat{V}|\Psi_0^{(0)}\rangle}{E_n^{(0)} - E_0^{(0)}}. \tag{1.74}$$

Inserting this expansion into the second-order energy expression gives a readily computable formula for the second-order Møller-Plesset (MP2) energy

$$E_0^{(2)} = \frac{1}{4}\sum_{ijab}\frac{((ia|jb) - (ib|ja))^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}. \tag{1.75}$$

The MPn energies are size consistent, but not variational. Size consistency can be seen by considering the MP2 energy for two widely separated systems A and B. The energy expression will only be non-zero if the orbitals $\psi_i, \psi_j, \psi_a, \psi_b$ are all on A, or all on B. Thus there are no cross-correlation terms.

The computational cost scaling of the MPn energy is $O(N^{(n+3)})$. For MP2 this arises from the need to transform the integrals over atomic orbitals into integrals over molecular orbitals. MP2 is a relatively cheap form of correlation, yet the higher orders become comparitively very expensive, especially considering that a CC or QCI calculation may be more accurate.

Perturbation theory relies on the starting wavefunction being close to the exact wavefunction. When this is the case, convergence of the MP series is rapid. However, when bonds are stretched the MP series becomes oscillatory. Also, if a UHF wavefunction with high spin is used, convergence can be extremely slow [28, 48]. Recent results have suggested that with large basis sets divergence can occur even for systems where HF is a good starting point [49]. For all these reasons it is expected that MP theory will become less popular.

1.6.5 Gaussian-2 Theory

If the energy of a molecule at its equilibrium geometry is all that is required, then a semi-empirical method for determining the correlation energy is available. Gaussian-2 (G2) theory [50–53] is a composite procedure, using HF, MP2, MP4 and QCISD(T) (therefore scaling as $O(N^7)$). It has been parametrized using 125 experimentally well-characterized atomization energies, ionization energies, electron affinities and proton affinities.

The procedure approximates a large basis QCISD(T) calculation using a series of additive approximations. G2 theory is usually within 1-2 kcal/mol of experiment. This combined with its 'black box' nature is making the method very popular — even for correcting experimental energies. However, the method has been shown to fail for some quite simple problems [54].

1.7 Basis Sets

Equation (1.16) states that the orbitals are made up of a linear combination of basis functions. Obviously, we have restricted the orbitals' flexibility unless the basis functions

form a complete set. Each function added to the basis increases the computational cost, so it is vital that the number of functions in a basis is kept as small as possible, while at the same time providing the orbital with as much flexibility as required.

The most convenient way to define a basis set for any nuclear configuration is to define a particular set of functions for each nucleus, depending only on the nuclear charge of that nucleus. There are two main types of basis functions in use today. The first, introduced by Slater in 1930, are termed Slater-Type Atomic Orbitals (STOs) [55]. STOs have exponential radial parts

$$\phi_a(\mathbf{r}) = (x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} e^{-\alpha|\mathbf{r}-\mathbf{A}|} \tag{1.76}$$

with a center $\mathbf{A} = (A_x, A_y, A_z)$, angular momentum $\mathbf{a} = (a_x, a_y, a_z)$ and nuclei dependent exponent α. STOs, like exact wavefunctions, have cusps at the nuclei and decay exponentially. Unfortunately, integrals over STOs are expensive to compute.

In 1950 Boys [56] suggested that basis functions constructed of Gaussian-Type Atomic Orbitals (GTOs) would overcome the computational difficulties of STOs. A GTO has the form

$$\phi_a(\mathbf{r}) = (x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} e^{-\alpha|\mathbf{r}-\mathbf{A}|^2}. \tag{1.77}$$

GTOs decay too fast and have incorrect nuclear cusps, so it is not surprising that many more GTOs than STOs are required to achieve the same accuracy [57]. However, the speed with which integrals over GTOs can be calculated more than compensates for this.

If STO properties are desired, they can be approximated by a sum of Gaussians, a philosophy which led to the introduction of the STO-nG basis sets [58]. These are an example of Contracted GTOs [59]

$$\phi_a(\mathbf{r}) = \sum_k^{K_A} D_{ak}(x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} e^{-\alpha_k|\mathbf{r}-\mathbf{A}|^2} \tag{1.78}$$

where K_A is referred to as the degree of contraction and the D_{ak} are the contraction coefficients. The contraction coefficients are not changed during a calculation, reducing the computational overhead.

There are several different basis sets in common use, all offering different trade-offs between accuracy and speed. Some of the more popular are those by Dunning [60] and the 'split-valence' sets by Pople and co-workers [61–63].

1.8 Molecular Properties

While the energy is undoubtedly the fundamental quantity, chemists usually characterize molecules by other properties, for example, the dipole moment or the molecular structure. The ability to accurately calculate these properties is one of the major strengths of modern electronic structure theory. These calculations are made possible by the fact that the properties are responses of the molecule to external parameters such as the nuclear coordinates, applied electric and magnetic fields, etc. These parameters become variables on which a potential energy surface is mapped out. Therefore analytic derivatives of the energy with respect to these variables yields the familiar molecular properties.

The derivatives with respect to nuclear positions give the nuclear forces, which allows rapid minimization of the energy with respect to nuclear coordinates, providing the molecular structure. Second derivatives with respect to nuclear position reveal the force constants, allowing harmonic frequencies to be calculated. These derivatives also allow the classification of stationary points, greatly facilitating the location of transition structures (which will be first order saddle points).

The various derivatives with respect to electric field, magnetic field and nuclear spin allow determination of a range of properties, including: electric polarizability, infrared intensities, magnetic susceptibility, chemical shielding, spin-spin coupling, Raman intensities and hyperpolarizabilities. However they are beyond the scope of this thesis, and these properties will not be discussed here. What is important is that they all result from derivatives of the energy, and thus fast evaluation of the molecular energy is highly desired.

Chapter 2

Density Functional Theory

"[DFT is] As easy as rolling over in bed" - H. F. Schaefer III

2.1 Introduction

In recent years Density Functional Theory (DFT) has become the most popular method in quantum chemistry, accounting for approximately 90% of all calculations today. The reason for this preference is the extreme computational cost required to obtain chemical accuracy with multiple determinant methods. DFT scales with the same order as HF theory — $O(N)$.

This difference in speed is heightened by the fact that multiple determinant calculations require very large basis sets, with high momentum basis functions, whereas DFT can produce accurate results with relatively small basis sets. This is due to the poor behaviour of the HF wavefunction when the inter-electronic distance becomes very small. The cusp-condition [64–67] states that the wavefunction should increase linearly when moving away from $r_{12} = 0$. The post-HF methods of section (1.6) try to account for this by introducing terms of r_{12}^2 and higher. Thus, the convergence of the correlation energy with the momentum in the basis set can be exceedingly slow, of the order of $(l + \frac{1}{2})^{-4}$ [68].

DFT avoids the expense of the more traditional methods, deriving the energy directly from the electron probability density, rather than the molecular wavefunction,thus drastically reducing the dimensionality of the problem.

2.2 The Hohenberg-Kohn Theorems

DFT was given a formal footing by the two theorems introduced by Hohenberg and Kohn in 1965. It is said that after the two theorems were introduced, the spectroscopist E. Bright Wilson [69] stood up and gave a much more conceptual overview of the theory.

If the exact electron density is known, then the cusps in $\rho(\mathbf{r})$ will provide the positions of the nuclei. The slope of $\rho(\mathbf{r})$ at the nucleus A must obey

$$\frac{\partial}{\partial \mathbf{r}_A}\bar{\rho}(\mathbf{r}_A)\bigg|_{\mathbf{r}_A=0} = -2Z_A\bar{\rho}(0) \tag{2.1}$$

(where $\bar{\rho}$ denotes the spherical average of the density) giving the charge at the nucleus, Z_A. Thus the full Schrödinger Hamiltonian is known, as it is completely defined by the nuclear charges and position. Therefore the wavefunction and energy can be found, and hence, the system can be completely described by the electron density.

A Hamiltonian of the form

$$\hat{H} = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N v(\mathbf{r}_i) + \sum_i^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.2}$$

is completely determined by the external potential, $v(\mathbf{r})$. The first Hohenberg and Kohn theorem [70] states that, for non-degenerate ground states, *the external potential* $v(\mathbf{r})$ *is determined, to within an additive constant, by the electron density,* $\rho(\mathbf{r})$. The theorem has since been extended to include degenerate ground states [71].

The proof is based on the minimum energy principle and begins by considering two external potentials, $v_1(\mathbf{r})$ and $v_2(\mathbf{r})$ arising from the same density. There will be two Hamiltonians, $\hat{H}_1$ and $\hat{H}_2$ with the same density, but different wavefunctions, Ψ_1 and Ψ_2. Now, using the variational principle,

$$E_1^0 < \langle\Psi_2|\hat{H}_1|\Psi_2\rangle = \langle\Psi_2|\hat{H}_2|\Psi_2\rangle + \langle\Psi_2|\hat{H}_1 - \hat{H}_2|\Psi_2\rangle \tag{2.3}$$

$$= E_2^0 + \int \rho(\mathbf{r})[v_1(\mathbf{r}) - v_2(\mathbf{r})]\, d\mathbf{r}. \tag{2.4}$$

Similarly

$$E_2^0 = E_1^0 - \int \rho(\mathbf{r})[v_1(\mathbf{r}) - v_2(\mathbf{r})]\, d\mathbf{r}, \tag{2.5}$$

which leads to the contradiction

$$E_1^0 + E_2^0 < E_2^0 + E_1^0. \tag{2.6}$$

Hence, the external potential is determined by the density and we may thus represent the energy as a functional of the density

$$E[\rho] = \int \rho(\mathbf{r}) v(\mathbf{r})\, dr + T[\rho] + V_{ee}[\rho] \tag{2.7}$$

where $T[\rho]$ is the kinetic energy and $V_{ee}[\rho]$ is the electron-electron repulsion, including the Coulombic interaction, $J[\rho]$:

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\, d\mathbf{r}_1 d\mathbf{r}_2. \tag{2.8}$$

The second Hohenberg-Kohn theorem [70] introduces the variational principle into DFT; *for a trial density* $\tilde{\rho}(\mathbf{r})$*, such that* $\tilde{\rho}(\mathbf{r}) \geq 0$ *and* $\int \tilde{\rho}(\mathbf{r})\, d\mathbf{r} = N$*,*

$$E_0 \leq E[\tilde{\rho}] \tag{2.9}$$

where $E[\tilde{\rho}]$ *is the energy functional from equation* (2.7). The proof is as follows: the first Hohenberg-Kohn theorem allows $\tilde{\rho}$ to determine its own potential $\tilde{v}$, Hamiltonian $\hat{H}$ and wavefunction $\tilde{\Psi}$, which can be used as a trial wavefunction for the problem with external potential v. Therefore,

$$\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle = \int \tilde{\rho}(\mathbf{r}) v(\mathbf{r})\, d\mathbf{r} + T[\tilde{\rho}] + V_{ee}[\tilde{\rho}] = E[\tilde{\rho}] \geq E[\rho]. \tag{2.10}$$

Assuming that $E[\rho]$ is differentiable, equation (2.9) requires that the ground state density be stationary, subject to the constraint that the integral of the density gives the number of electrons,

$$\delta E[\rho] - \mu \delta \left[\int \rho(\mathbf{r})\, d\mathbf{r} - N \right] = 0 \tag{2.11}$$

which leads to the Euler-Lagrange equation

$$\mu = v(\mathbf{r}) + \frac{\delta T[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta V_{ee}[\rho]}{\delta \rho(\mathbf{r})} \tag{2.12}$$

where a familiar property to chemists, the chemical potential μ, has been introduced.

Equation (2.12) would be an exact equation for $\rho(\mathbf{r})$ if the exact form of $T[\rho]$ and $V_{ee}[\rho]$ were known. Unfortunately the Hohenberg-Kohn theorems do not provide this, only that they exist. Also $T[\rho]$ and $V_{ee}[\rho]$ are defined independently of $v(\mathbf{r})$, so once we have a form for these functionals they can be applied to any system.

2.3 The Constrained Search Formulation

The second Hohenberg-Kohn theorem has two drawbacks. Firstly, it assumes that there is no degeneracy in the ground state, and secondly the density must be v-representable: it must arise from a wavefunction with a Hamiltonian able to be written in the form of equation (2.2). The specific conditions that make a density v-representable are unknown, but many 'reasonable' densities have been shown to be non-v-representable [72, 73].

A weaker constraint, that the density is N-representable, can be used if the Levy constrained-search is used [72, 74]. A density is N-representable if it can be obtained from some antisymmetric wavefunction. The theory begins by showing how to distinguish the ground state wavefunction Ψ_0 from a wavefunction Ψ_{ρ_0} that simply integrates to the ground state density $\rho_0(\mathbf{r})$. The variational principle gives

$$\langle \Psi_{\rho_0} | \hat{H} | \Psi_{\rho_0} \rangle \geq \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = E_0. \tag{2.13}$$

Remembering that the potential energy due to the external field $v(\mathbf{r})$ is a function of the density leads to

$$\langle \Psi_{\rho_0} | \hat{T} + \hat{V}_{ee} | \Psi_{\rho_0} \rangle + \int v(\mathbf{r}) \rho_0(\mathbf{r}) \, d\mathbf{r} \geq \langle \Psi_0 | \hat{T} + \hat{V}_{ee} | \Psi_0 \rangle + \int v(\mathbf{r}) \rho_0(\mathbf{r}) \, d\mathbf{r} \tag{2.14}$$

$$\langle \Psi_{\rho_0} | \hat{T} + \hat{V}_{ee} | \Psi_{\rho_0} \rangle \geq \langle \Psi_0 | \hat{T} + \hat{V}_{ee} | \Psi_0 \rangle. \tag{2.15}$$

Thus the Ψ_0 is the wavefunction that integrates to ρ_0 and minimizes the expectation value of $\hat{T} + \hat{V}_{ee}$. Defining our universal functional as

$$F[\rho] = \min_{\Psi \to \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle, \tag{2.16}$$

where $F[\rho]$ searches all Ψ that yield the input density ρ, allows the energy to be expressed as

$$E_0 = \min_{\rho} \left[F[\rho] + \int v(\mathbf{r}) \rho(\mathbf{r}) \, d\mathbf{r} \right] \tag{2.17}$$

$$= \min_{\rho} E[\rho] \tag{2.18}$$

where

$$E[\rho] = F[\rho] + \int v(\mathbf{r}) \rho(\mathbf{r}) \, d\mathbf{r} \tag{2.19}$$

which is a search over all N-representable densities. Thus, the v-representable problem has been removed. The restriction requiring no degeneracy has also been lifted. In degenerate systems the wavefunction giving $\rho(\mathbf{r})$ will be selected. All that remains is to find an accurate form for the functional $E[\rho]$.

2.4 Density Matrices

Before introducing the various density functionals, it is useful to examine density matrices and the exchange correlation-hole. The N-th order density matrix is defined as

$$\gamma_N(\mathbf{x}'_1\mathbf{x}'_2\cdots\mathbf{x}'_N,\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N) = \Psi_N(\mathbf{x}'_1\mathbf{x}'_2\cdots\mathbf{x}'_N)\Psi^*_N(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N). \tag{2.20}$$

From this the first- and second-order reduced density matrices can be defined:

$$\gamma_1(\mathbf{x}'_1,\mathbf{x}_1) = N\int\cdots\int \Psi_N(\mathbf{x}'_1\mathbf{x}'_2\cdots\mathbf{x}'_N)\Psi^*_N(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N)\,d\mathbf{x}_2\ldots d\mathbf{x}_N \tag{2.21}$$

$$\gamma_2(\mathbf{x}'_1\mathbf{x}'_2,\mathbf{x}_1\mathbf{x}_2) = \frac{N(N-1)}{2}\int\cdots\int \Psi_N(\mathbf{x}'_1\mathbf{x}'_2\cdots\mathbf{x}'_N)\Psi^*_N(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N)\,d\mathbf{x}_3\ldots d\mathbf{x}_N. \tag{2.22}$$

Note that the first-order density matrix integrates to the number of electrons, and the second-order density matrix integrates to the number of electron pairs. Obviously, γ_1 can be obtained from γ_2 by integration,

$$\gamma_1(\mathbf{x}'_1,\mathbf{x}_1) = \frac{2}{N-1}\int \gamma_2(\mathbf{x}'_1\mathbf{x}_2,\mathbf{x}_1\mathbf{x}_2)\,d\mathbf{x}_2. \tag{2.23}$$

Most operators of interest do not involve the spin coordinates, so it is common to integrate over spin, forming the spinless density matrices [75],

$$\rho_1(\mathbf{r}'_1,\mathbf{r}_1) = \int \gamma_1(\mathbf{r}'_1s_1,\mathbf{r}_1s_1)\,ds_1 \tag{2.24}$$

$$\rho_2(\mathbf{r}'_1\mathbf{r}'_2,\mathbf{r}_1\mathbf{r}_2) = \iint \gamma_2(\mathbf{r}'_1s_1\mathbf{r}'_2s_2,\mathbf{r}_1s_1\mathbf{r}_2s_2)\,ds_1ds_2. \tag{2.25}$$

The diagonal element of $\rho(\mathbf{r}'_1,\mathbf{r}_1)$ is simply the electron density, $\rho(\mathbf{r}_1)$. There is a shorthand for the diagonal element of ρ_2,

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \rho_2(\mathbf{r}_1\mathbf{r}_2,\mathbf{r}_1\mathbf{r}_2). \tag{2.26}$$

Using this new notation the expectation value of the electronic Hamiltonian can be written as

$$E = \int[-\frac{1}{2}\nabla^2_{\mathbf{r}}\rho_1(\mathbf{r}',\mathbf{r})]_{\mathbf{r}'=\mathbf{r}}\,d\mathbf{r} + \int v(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r} + \iint \frac{1}{r_{12}}\rho_2(\mathbf{r}_1,\mathbf{r}_2)\,d\mathbf{r}_1d\mathbf{r}_2. \tag{2.27}$$

For restricted HF the last term simplifies to

$$J[\rho] - K[\rho_1] = \frac{1}{2}\iint \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\,d\mathbf{r}_1d\mathbf{r}_2 - \frac{1}{4}\iint \frac{1}{r_{12}}|\rho_1(\mathbf{r}_1,\mathbf{r}_2)|^2\,d\mathbf{r}_1d\mathbf{r}_2 \tag{2.28}$$

where the first-order reduced density matrix, the Fock-Dirac density matrix, is defined in terms of the HF orbitals

$$\rho_1(\mathbf{r}_1,\mathbf{r}_2) = 2\sum_i^{N/2}\psi(\mathbf{r}_1)\psi^*(\mathbf{r}_2). \tag{2.29}$$

2.5 The Exchange Correlation Hole

The last term of equation (2.27) can be separated into the classical, $J[\rho]$, and non-classical parts by defining

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)[1 + h(\mathbf{r}_1, \mathbf{r}_2)] \tag{2.30}$$

where $h(\mathbf{r}_1\mathbf{r}_2)$ is the pair correlation function. Slater [76] looked at this in a slightly different way, defining the exchange-correlation hole by

$$\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_2)h(\mathbf{r}_1, \mathbf{r}_2). \tag{2.31}$$

Using the spinless equivalent of equation (2.23) we find the condition

$$\int \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_2 = -1 \tag{2.32}$$

which must hold for all values of $\mathbf{r}_1$. The electron repulsion term can then be written

$$V_{ee} = J[\rho] + \frac{1}{2}\iint \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2, \tag{2.33}$$

where the non-classical part has been expressed as a repulsion between the density and the exchange correlation hole, a distribution of unit positive charge centered around $\mathbf{r}_1$. The coulomb potential due to the non-classical part has been shown to have the asymptotic behaviour [77]

$$\lim_{r_1\to\infty} v_{xc}(\mathbf{r}_1) = \lim_{r_1\to\infty} \int \frac{1}{r_{12}}\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_2 = -\frac{1}{\mathbf{r}_1} \tag{2.34}$$

2.6 The Uniform Electron Gas

There is no systematic way to find or improve a density functional. The most appealing way forward is to find the exact solution for a model system, and then assume that the system of interest behaves similarly to the model. The first density functionals were due to Thomas [78], Fermi [79–81] and Dirac [82], all of which used the uniform electron gas as their model.

The uniform electron gas is defined as a large number of electrons N in a cube of volume V, throughout which there is a uniform spread of positive charge sufficient to make the system neutral. The uniform gas is then defined as the limit $N \to \infty$, $V \to \infty$, with the density $\rho = N/V$ remaining finite. Although it does bear some resemblance to electrons in

metals, its widespread use is due to its simplicity — it is completely defined by one variable, the electron density ρ.

Using the uniform electron gas, an expression for the kinetic energy (the Thomas-Fermi kinetic functional) can be derived [83]

$$T^{TF27}[\rho_\sigma] = \frac{3}{10}(6\pi^2)^{2/3} \int \rho_\sigma^{5/3}(\mathbf{r})\, d\mathbf{r}, \tag{2.35}$$

where σ can take the values of α or β. When applied to atoms and molecules the Thomas-Fermi functional yields kinetic energies that are about 10% too small.

Similarly, an expression for the exchange energy of the uniform electron gas can be calculated (the Dirac exchange functional) [83]

$$E_x^{D30}[\rho_\sigma] = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3} \int \rho_\sigma^{4/3}(\mathbf{r})\, d\mathbf{r}. \tag{2.36}$$

The Dirac functional also gives exchange energies that are roughly 10% smaller than those from HF theory [84]. More worrying is that the spurious self-interaction of electrons is not exactly canceled.

A closed shell functional for the correlation energy of the uniform electron was determined by Vosko, Wilk and Nusair [85], who combined analytic information about the high and low density limits with the quantum Monte-Carlo simulation results of Ceperly and Alder [86]. The VWN functional usually overestimates the correlation energy of atoms and molecules by approximately a factor of two [87]. The uniform electron gas is obviously a better reference system for exchange energies than it is for correlation energies.

2.7 The Almost Uniform Electron Gas

The electron densities of atoms and molecules are often far from uniform, so functionals based on systems which include an inhomogeneous density should perform better. In 1935 von Weizsacker [88] placed infinitesimally small ripples on the uniform electron gas and calculated the second order correction to the kinetic energy

$$T^{W35}[\rho_\sigma] = T^{TF27}[\rho_\sigma] + \frac{1}{8} \int \rho_\sigma^{5/3} x_\sigma^2\, d\mathbf{r} \tag{2.37}$$

where $x(\mathbf{r})$ is a dimensionless quantity, the reduced density gradient

$$x(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}. \tag{2.38}$$

Unfortunately the original derivation was flawed and the above functional is too large by a factor of nine [83]. The corrected functional is a large improvement on $T^{TF27}[\rho]$, yielding kinetic energies typically within 1% of HF theory. The fourth order [89] and sixth order [90] corrections have subsequently been computed; however, the series is divergent due to the extremely large values of $x(\mathbf{r})$ in the Rydberg regions, and it is advantageous to stop after second order.

A similar correction was made to the Dirac exchange functional by Sham [91]. Kleinman [92] later showed that the Sham derivation was too small by 10/7. The second order correction to the exchange energy is

$$E_x^{SK71}[\rho_\sigma] = E_x^{D30}[\rho_\sigma] - \frac{5}{(36\pi)^{5/3}} \int \rho_\sigma^{4/3} x_\sigma^2 \, d\mathbf{r}. \tag{2.39}$$

The corrected functional gives exchange energies that are typically within 3% of HF; however, it is not seen as an improvement over the Dirac functional, as the potential is unbounded in the Rydberg regions of atoms and molecules.

2.8 Kohn-Sham Theory

The kinetic energy has a large contribution to the total energy. Therefore even the 1% error in the kinetic energy of the Thomas-Fermi-Weizsacker model prevented DFT from being used as a quantitative predictive tool. Thus DFT was largely ignored until 1965 when Kohn and Sham [93] introduced a method which treated the majority of the kinetic energy exactly.

The theory begins by considering the noninteracting reference system: N noninteracting electrons, each in one of N orbitals, ψ_i. Such a system will be defined by the Hamiltonian

$$\hat{H}_s = \sum_i^N (-\frac{1}{2}\nabla_i^2) + \sum_i^N v_s(\mathbf{r}_i) \tag{2.40}$$

which has an exact eigenfunction that is the single determinant constructed from the N lowest eigenstates of the one-electron equations

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right] \psi_i = \varepsilon_i \psi_i. \tag{2.41}$$

The corresponding Euler-Lagrange equation is

$$\mu = v_s(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})}. \tag{2.42}$$

For this system the kinetic energy and electron density are given *exactly* by

$$T_s[\rho] = \sum_i^N \langle \psi_i | -\frac{1}{2}\nabla_i^2 | \psi_i \rangle \tag{2.43}$$

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{2.44}$$

and the total energy is given by

$$E[\rho] = T_s[\rho] + \int v_s(\mathbf{r})\rho(\mathbf{r})\, d\mathbf{r} \tag{2.45}$$

The quantity $T_s[\rho]$ is well-defined, but not the exact kinetic energy, $T[\rho]$, defined in equation (2.7). Kohn and Sham reformulated the interacting problem so that its kinetic component is defined to be $T_s[\rho]$ and rearranging equation (2.7) to give

$$E[\rho] = T_s[\rho] + J[\rho] + \int v(\mathbf{r})\rho(\mathbf{r}) + E_{xc}[\rho], \tag{2.46}$$

where $E_{xc}[\rho]$ is the exchange-correlation energy, made up of the non-classical electron-electron repulsion and also the difference between the exact and noninteracting kinetic energy

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]. \tag{2.47}$$

The Euler-Lagrange equation (2.12) now becomes

$$\mu = v_{\text{eff}}(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \tag{2.48}$$

where the Kohn-Sham (KS) effective potential, v_{eff} is defined as

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' + v_{xc}(\mathbf{r}) \tag{2.49}$$

and the exchange-correlation potential, v_{xc} is

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \tag{2.50}$$

Kohn and Sham noticed that equation (2.48) is the same as that for a non-interacting system moving in the potential $v_{\text{eff}}(\mathbf{r})$. Thus, the exact density can be obtained by solving the N one-electron equations (the restricted KS equations)

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right]\psi_i = \varepsilon_i \psi_i. \tag{2.51}$$

Notice that v_{eff} depends on $\rho(\mathbf{r})$, via equation (2.50), hence the KS equations must be solved iteratively.

The KS equations are very similar to the Hartree-Fock equations. In fact, setting the exchange-correlation potential to the HF exchange potential,

$$v_{xc}(\mathbf{r}) = -\int \sum_j \frac{\psi_j(\mathbf{r}')\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\hat{P}_{\mathbf{r}\mathbf{r}'}\, d\mathbf{r}', \tag{2.52}$$

yields the HF equations. Drawing too many similarities to HF is dangerous, however. Firstly, the KS orbitals are simply a way of representing the density; they are not (as in HF) an approximation of the wavefunction. In particular, Koopmans' theorem [94] — that the ionization potentials and electron affinities are approximated by the negative of the HF occupied and virtual orbital eigenvalues respectively — is invalid for KS orbitals. The highest occupied KS eigenvalue has been shown to be the negative of the first ionization potential, though [95]. Also, HF theory is variational, providing an upper bound to the exact energy, yet DFT is only variational if the exact energy functional is used.

The above analysis is only appropriate for closed shell molecules. Because the KS equations so closely follow the restricted HF equations, both the restricted open shell and unrestricted methodologies are readily available. However, the KS equations are formally exact (given the exact $E_{xc}[\rho]$), so it must be able to produce an excess of β electron density at points in the molecule [96], and therefore only the unrestricted formalism is appropriate. The unrestricted KS equations are

$$\left[-\frac{1}{2}\nabla^2 + v^{\alpha}_{\text{eff}}(\mathbf{r})\right]\psi^{\alpha}_i = \varepsilon^{\alpha}_i\psi^{\alpha}_i \tag{2.53}$$

$$\left[-\frac{1}{2}\nabla^2 + v^{\beta}_{\text{eff}}(\mathbf{r})\right]\psi^{\beta}_i = \varepsilon^{\beta}_i\psi^{\beta}_i \tag{2.54}$$

where

$$v^{\alpha}_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho_\alpha,\rho_\beta]}{\delta\rho_\alpha(\mathbf{r})} \tag{2.55}$$

$$v^{\beta}_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho_\alpha,\rho_\beta]}{\delta\rho_\beta(\mathbf{r})}. \tag{2.56}$$

One problem with the above derivation of the KS equations is that the density must be non-interacting v-representable. That is, there must exist a potential v_s that will produce the same density as the exact wavefunction. If the density is not non-interacting v-representable, the determinant formed from the KS orbitals will be an excited state [97]. The criteria that make a density non-interacting v-representable are unknown.

Just as in HF theory, the KS equations are solved by expanding the orbitals over a basis set. The major advantage of DFT is that the basis set requirements are far more modest

than the more conventional correlated methods [98, 99]. In DFT the basis set only needs to represent the one electron density — the inter-electron cusp is accounted for by the effective potential, v_{eff}. In the more traditional methods the basis set describes the entire N-electron wavefunction, requiring an accurate description of the cusp which is sensitive to the basis set.

2.9 Exchange-Correlation Functionals

KS Theory allows the kinetic energy to be computed to a chemical accuracy, so 'all' that remains is an accurate form for the exchange-correlation energy functional, $E_{xc}[\rho]$. The exact form is obviously unknown, and with the accuracy of DFT determined mainly by the functional used, it is no surprise that finding new functionals is the focus of much modern research.

The simplest form for $E_{xc}[\rho]$ is the Dirac exchange term, forming Hartree-Fock-Slater (HFS) theory [76]. Slater also pointed out that E_x^{D30} systematically underestimated the exchange energy by about 10% and proposed multiplying the Dirac coefficient by 1.1, resulting in the semi-empirical X_α theory. While the HFS total energies are not as accurate as HF theory, for thermochemistry, HFS theory is a big improvement over HF theory [100]. This is due to a convenient cancellation of errors arising from HFS systematically underestimating the total energy.

The natural extension of HFS theory is to add the VWN functional for the correlation energy, thus using the uniform electron gas to model exchange and correlation effects. The resulting theory is termed the Local Spin Density approximation (LSDA). The LSDA is an improvement over HFS theory, yet VWN makes no account for the correction of the kinetic energy, $T_c[\rho]$, where

$$T_c[\rho] = T[\rho] - T_s[\rho] \tag{2.57}$$

which can reach the magnitude of the correlation energy itself [77]. As VWN usually overestimates E_c by a factor of two, inclusion of T_c could have a dramatic effect on the accuracy of the VWN functional.

2.9.1 The Becke Exchange Functional

The exchange energy is an order of magnitude larger than correlation energy, therefore, the 10% error of E_x^{D30} is the major problem of the LSDA. One reason for this could be the

incorrect asymptotic behaviour of E_x^{D30}.

The exchange energy density $\varepsilon_x(\mathbf{r}_1)$ may be defined as

$$E_x[\rho] = \int \rho(\mathbf{r}_1)\varepsilon_x(\mathbf{r}_1)\, d\mathbf{r}_1 \tag{2.58}$$

$$\varepsilon_x(\mathbf{r}_1) = \frac{1}{2}\int \frac{\rho_x(\mathbf{r}_1, \mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_2. \tag{2.59}$$

Using equation (2.32) the following constraint for $E_x[\rho]$ is obtained:

$$\lim_{r_1\to\infty} \varepsilon_x(\mathbf{r}_1) = -\frac{1}{2r_1}. \tag{2.60}$$

The long range behaviour of the electron density is

$$\lim_{r\to\infty} \rho(\mathbf{r}) = \exp\left[-2\sqrt{2I_{\min}}r\right]. \tag{2.61}$$

where $I_{\min}$ is the exact first ionization potential [101]. Therefore the LSDA ε_x will have the asymptotic form

$$\lim_{r\to\infty} \varepsilon_x^{LSDA}(\mathbf{r}_1) = \exp\left[-\frac{2}{3}\sqrt{2I_{\min}}r\right] \tag{2.62}$$

In 1988 Becke [102] introduced a correction to the Dirac exchange functional which gives the exchange energy density the correct asymptotic behaviour. The functional form is

$$E_x^{B88}[\rho_\sigma] = E_x^{D30}[\rho_\sigma] - b\int \rho_\sigma^{4/3}\frac{x_\sigma^2}{1 + 6bx_\sigma \sinh^{-1} x_\sigma}\, d\mathbf{r}, \tag{2.63}$$

with the parameter $b = 0.0042$ determined by fitting the exchange energies of the first six noble gas atoms. One deficiency of Becke's functional is that the potential decays asymptotically as [103, 104]

$$\lim_{r\to\infty} v_x^{B88}(\mathbf{r}) = \frac{1}{r^2}. \tag{2.64}$$

instead of the correct [104, 105]

$$\lim_{r\to\infty} v_x(\mathbf{r}) = \frac{1}{r} \tag{2.65}$$

Despite this drawback, E_x^{B88} is an extremely accurate density functional. For predicting atomic exchange energies it is 1-2 orders of magnitude better than the Sham-Kleinman functional [106].

2.9.2 The Perdew-Wang 91 Functional

In 1991 Perdew [107, 108] modified Becke's functional, constraining it to become like the Sham-Kleinman functional as $x \to 0$, and tend towards zero as $x \to \infty$, while trying to retain the E_x^{B88} behaviour for medium range x. The functional form is

$$E_x^{PW91}[\rho_\sigma] = E_x^{D30}[\rho_\sigma] - \int \frac{bx_\sigma^2 - (b - 5(36\pi)^{-5/3})x_\sigma^2 \exp(-1.6455x_\sigma^2) - 10^{-6}x_\sigma^4}{1 + 6bx_\sigma \sinh^{-1} x_\sigma - 10^{-6}x_\sigma^4/\alpha}\, d\mathbf{r} \quad (2.66)$$

where α is the Dirac coefficient of $\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}$. However, despite the added complexity, energies obtained from E_x^{PW91} are seldom an improvement over E_x^{B88} [106]. The functional has also been shown to violate the original condition upon which E_x^{B88} was developed [109].

2.9.3 The Gill Functional

With the knowledge that satisfying limiting constraints had not helped in the E_x^{PW91} functional and that the E_x^{B88} functional was a (successful) attempt to reduce the x^2 behaviour of the Sham-Kleinman functional, Gill [106] introduced a simple exchange functional

$$E_x^{G96}[\rho_\sigma] = E_x^{D30}[\rho_\sigma] - b \int \rho_\sigma^{4/3} x_\sigma^{3/2}\, d\mathbf{r} \quad (2.67)$$

where the parameter $b = 1/137$ was chosen by fitting the exchange energy of the Argon atom. The E_x^{G96} functional has incorrect behaviour for both high and low x, yet it is similar to E_x^{B88} for mid-range x values. The fact that this far simpler functional performs comparably to E_x^{B88} [106] shows again that the limiting behaviour of functionals is of less importance than the behaviour for mid-range values of x.

2.9.4 The Lee-Yang-Parr Functional

All of the above functionals use, in some way, the uniform electron gas. This approach was abandoned by Lee, Yang and Parr in 1988, who turned to the Helium atom instead. Colle and Salvetti [110] had already presented an approximate correlation energy formula for the Helium atom in terms of the second order HF density matrix. Lee, Yang and Parr [111] turned this into a functional of the density, gradient and Laplacian. Miehlich, Savin, Stoll and Press [112] later eliminated the Laplacian terms using integration by parts. For closed shell systems the functional is

$$E_c^{LYP}[\rho] = -a \int \frac{\rho}{1 + d\rho^{-1/3}} - \frac{11b}{24}\rho^2 |\nabla\rho|^2 + b\left[\frac{3}{10}(3\pi^2)^{2/3}\rho^{8/3} + |\nabla\rho|^2\left(\frac{5}{12} - \frac{7\delta}{72}\right)\right]\omega\rho^2\, d\mathbf{r} \quad (2.68)$$

where

$$\omega = \frac{\exp(-c\rho^{-1/3})}{1 + d\rho^{-1/3}}\rho^{-11/3} \tag{2.69}$$

$$\delta = c\rho^{-1/3} + \frac{d\rho^{-1/3}}{1 + d\rho^{-1/3}} \tag{2.70}$$

with the parameters $a = 0.04918$, $b = 0.132$, $c = 0.2533$ and $d = 0.349$ derived from the Colle-Salvetti fit to the Helium atom.

E_c^{LYP} is one of the most accurate correlation functionals at the moment, especially when combined with E_x^{B88}, forming BLYP theory. BLYP geometries are of a standard comparable to HF theory (however, BLYP tends to overestimate bond lengths, whereas HF underestimates), and atomization energies, ionization energies, electron affinities and proton affinities are usually accurate to within 20 kJ mol^{-1} of experiment [100]. BLYP predicts harmonic frequencies that are of an accuracy similar to that of MP2 theory — a noticeable improvement over HF.

BLYP can fail quite spectacularly though, especially in systems with stretched bonds. An example is the barrier height calculation for the reaction [113]

$$\mathrm{H} + \mathrm{H_2} \rightarrow \mathrm{H_2} + \mathrm{H} \tag{2.71}$$

While HF, MP2 and CCSD(T) all do quite well in predicting the barrier height, BLYP predicts a height only one quarter of that measured experimentally (however, LSDA predicts H_3 to be more stable than H + H_2!). The lack of an exact self-interaction correction was shown to be important.

2.9.5 The Wigner Functional

An extremely simple form for the correlation energy was suggested by Wigner in 1938 [114], which has subsequently been shown to perform better than E_c^{VWN} [115]. The functional form contains two parameters, and there have been a number of reparametrizations [116–118], each with its own strengths. The functional occurs as a term in E_c^{LYP}, and simply sticking with these parametisations (and pairing with E_x^{B88}) yields an accuracy surprisingly close to BLYP [118]. The spin-polarized functional form is

$$E_c^{W38}[\rho_\alpha, \rho_\beta] = -4a \int \frac{1}{1 + d(\rho_\alpha + \rho_\beta)^{-1/3}} \frac{\rho_\alpha \rho_\beta}{\rho_\alpha + \rho_\beta}\, d\mathbf{r} \tag{2.72}$$

with the parameters $a = 0.04918$ and $d = 0.349$. The Wigner form does satisfy limiting conditions [115], but it is perhaps of most importance due to its simplicity, showing that current correlation functionals (such as E_c^{LYP}) may be unnecessarily complicated.

2.9.6 Hybrid Density Functionals

A familiar theme when accessing the accuracy of exchange-correlation functionals is that density based functionals will overestimate a quantity which HF theory will underestimate (for example, bond lengths [100]). With this in mind, Becke [119] has argued that the exact exchange-correlation functional must include a fraction of HF exchange. Initially Becke [120] proposed a functional that consisted of 50% HF exchange and 50% density-based exchange-correlation. This functional was quickly replaced by the three parameter mix denoted B3P

$$E_{xc}^{B3P} = E_{xc}^{LSDA} + a_0(E_x^{HF} - E_x^{D30}) + a_x(E_x^{B88} - E_x^{D30}) + a_c E_c^{PW91} \tag{2.73}$$

where E_c^{PW91} is the gradient correction of the Perdew-Wang correlation functional [107, 108] (often used in conjunction with E_x^{PW91}). The three parameters $a_0 = 0.20$, $a_x = 0.72$ and $a_c = 0.81$ were determined by minimizing the atomization energies, ionization energies, electron affinities and proton affinities of the G2 dataset.

An alternative functional has been proposed, using E_c^{LYP} instead of E_c^{PW91} [121]. B3LYP has the form

$$E_{xc}^{B3LYP} = E_{xc}^{LSDA} + a_0(E_x^{HF} - E_x^{D30}) + a_x(E_x^{B88} - E_x^{D30}) + a_c(E_c^{LYP} - E_c^{VWN}) \tag{2.74}$$

with the same parameters as in B3P. B3LYP shows surprising accuracy for thermochemistry, structures and spectroscopic properties of first row molecules [122]. The high accuracy of B3LYP has made it perhaps the most popular functional of modern density functional theory.

2.10 Numerical Evaluation of Exchange-Correlation Integrals

The exchange-correlation functions are sufficiently complicated that the integrals required cannot be carried out analytically. Even the simplest of the family (E_x^{D30}) is too complex if Gaussian basis functions are used to represent the density. One way forward is to expand the exchange-correlation potential in terms of auxiliary Gaussian functions [123]. The disadvantage of this scheme is that it is hard to reduce the extent of approximation towards zero. A more popular way is to calculate the integrals numerically using quadrature schemes.

2.10.1 Voronoi Polyhedra

The electron density contains cusps at the nuclei and decays exponentially away from the nuclei. This suggests that the placement of quadrature points should be dependent

on the nuclear positions. This can be achieved by partitioning space into regions (Voronoi Polyhedra), each containing only one atom. The total integral is the sum of single atom integrals. There are a number of such partitioning schemes [123, 124]; however, the scheme used by the package produced in our research group, Q-CHEM [125], is that developed by Becke [126].

The integral

$$E_{xc}[\rho] = \int F(\mathbf{r})\, d\mathbf{r} \tag{2.75}$$

is split up via weight functions $w_A(\mathbf{r})$

$$E_{xc}[\rho] = \sum_A \int w_A(\mathbf{r}) F(\mathbf{r})\, d\mathbf{r} \tag{2.76}$$

where the weight functions obey

$$w_A(\mathbf{r}) \geq 0 \qquad \text{and} \qquad \sum_A w_A(\mathbf{r}) = 1. \tag{2.77}$$

These weights are constructed to be almost unity if A is the closest atom, and almost zero in the vicinity of other atoms. This is achieved through the variable μ, defined by

$$\mu_{AB} = \frac{r_A - r_B}{R_{AB}} \tag{2.78}$$

where r_A and r_B are the distance from atoms A and B, while R_{AB} is the interatomic distance between atoms A and B. The weight function is then described by

$$w_A(\mathbf{r}) = \frac{P_A(\mathbf{r})}{\sum\limits_{A \neq B} P_B(\mathbf{r})} \tag{2.79}$$

$$P_A(\mathbf{r}) = \prod_{A \neq B} s(\mu_{AB}) \tag{2.80}$$

where the function $s(\mu_{AB})$ is defined as

$$s(\mu_{AB}) = \begin{cases} 0 & \text{if } 0 < \mu_{AB} \leq 1 \\ 1 & \text{if } -1 \leq \mu_{AB} < 0. \end{cases} \tag{2.81}$$

Becke then removed the discontinuity at $\mu_{AB} = 0$ by redefining $s(\mu)$ as a function with

$$s(-1) = 1, \qquad s(+1) = 1, \qquad \text{and} \qquad \left.\frac{ds}{d\mu}\right|_{\pm 1} = 0. \tag{2.82}$$

Above is the implementation in Q-CHEM. Becke's original implementation also includes a correction for atomic size. This is accomplished by a change in variable, working with $s(\nu)$ instead of $s(\mu)$, where

$$\nu_{AB} = \mu_{AB} + a_{AB}(1 - \mu_{AB}^2) \tag{2.83}$$

with a_{AB} defined by

$$a_{AB} = \frac{u_{AB}}{u_{AB}^2 - 1} \tag{2.84}$$

$$u_{AB} = \frac{\chi - 1}{\chi + 1} \tag{2.85}$$

$$\chi = \frac{R_A}{R_B} \tag{2.86}$$

where R_A and R_B are the Bragg-Slater radii [127, 128].

Each of these single-center integrals is then calculated with a spherical polar quadrature grid (requiring the insertion of more weights, w_i), making the final expression for E_{xc}:

$$E_{xc} = \iiint F(r, \theta, \phi) r^2 \sin\theta \, dr d\theta d\phi = \sum_i \sum_A w_A(\mathbf{r}_i) w_i F(\mathbf{r}_i). \tag{2.87}$$

2.10.2 Radial Integration

Q-Chem, like most modern DFT packages, supports more than one radial quadrature scheme. With the time and accuracy of a calculation directly linked to the type of grid used it is not surprising that this is an area of much research, and new, improved grids are constantly being published [126, 129–132]. The work in this thesis, however, uses only the Euler-Maclaurin quadrature scheme.

The implementation follows that developed by Murray *et al.* [133]. The Euler-Maclaurin scheme integrates a function over the interval $[0, 1]$, using equally spaced points with equal weights. Thus a transformation from the interval $[0, \infty)$ is required. The convergence over the sum of quadrature points is more rapid if the integrand and its low derivatives are small at the end points. The nature of exchange-correlation functionals ensures that all derivatives vanish at $r = \infty$ and the Jacobian factor forces the integrand and its first derivative to vanish at $r = 0$. Handy and Boys [134, 135] thus introduced the transformation to the variable q,

$$r = \alpha \left(\frac{q}{1 - q} \right)^2 \tag{2.88}$$

$$\frac{dr}{dq} = 2\alpha \frac{q}{(1 - q)^3} \tag{2.89}$$

where α is taken as the Bragg-Slater radius. This transformation also has the benefit of forcing the integrand and all derivatives up to fifth order to vanish at $q = 0$, and all derivatives at $q = \infty$.

2.10.3 Angular Integration

There are two common types of angular quadrature schemes, differing in whether or not the θ and ϕ integrations are treated together (Lebedev) or separately (Gauss-Legendre).

Lebedev Quadrature

A large amount of work has been done on quadrature schemes to exactly integrate spherical harmonics over a sphere [136–139]. The most popular are those by Lebedev [140–143] which use quadratures based on the octahedral group. A Lebedev grid of degree L exactly integrates all spherical harmonics of degree L or less. The number of grid points, $\mathrm{NP}(L)$ required is approximately

$$\mathrm{NP}(L) \approx \frac{(L+1)^2}{3}. \tag{2.90}$$

Lebedev grids are currently available up to degree $L = 53$ (974 grid points), although Q-Chem has only up to degree $L = 29$ (302 grid points).

Gauss-Legendre Quadrature

While the Lebedev grids are very efficient they cannot be used for arbitrary precision. For this reason it is sometimes desirable to perform the θ and ϕ integration separately, even though it is likely to be less efficient.

The ϕ integration is performed with equally spaced points, while the θ integration uses Gauss-Legendre quadrature, which is designed to exactly integrate all polynomials up to degree $2\mathrm{NP} - 1$, where NP is the number of points. To exactly integrate all spherical harmonics up to degree L, the θ quadrature requires $(L+1)/2$ points and the ϕ quadrature $L+1$ points, thus

$$\mathrm{NP}(L) = \frac{(L+1)^2}{2} \tag{2.91}$$

which is 2/3 the efficiency of the Lebedev scheme.

2.10.4 Translational and Rotational Invariance

One of the problems that arises when integration is performed numerically with a finite grid is that translational and rotational invariance can be lost. Luckily, using the grid points as defined above ensures translational invariance. The grid points are linked to atomic centers and therefore translate with nuclear movement. However a rotation of the molecule

leaving the grid axes unrotated will cause a slight change in energy. This effect is even more pronounced with derivative calculations, in particular the calculation of harmonic frequencies [144].

An early attempt to solve rotational invariance was the use of randomly rotated angular grids, with the hope of averaging out any error [145]. A more rigorous solution was given by Johnson *et al.* [146] where the grid points are defined by

$$\mathbf{r}_i = \mathbf{R}_A + \mathbf{O}\mathbf{s}_i, \tag{2.92}$$

with $\mathbf{R}_A$ the atomic position, $\mathbf{s}_i$ the quadrature grid points, and $\mathbf{O}$ is the matrix formed from the eigenvectors, $\mathbf{M}$, of the charge moment tensor

$$\mathbf{M} = \sum_A Z_A[|\mathbf{R}_A - \mathbf{T}|^2\mathbf{I} - (\mathbf{R}_A - \mathbf{T})(\mathbf{R}_A - \mathbf{T})^T] \tag{2.93}$$

with Z_A the charge at nucleus A and

$$\mathbf{T} = \frac{\sum_A Z_A \mathbf{R}_A}{\sum_A Z_A}. \tag{2.94}$$

With this new definition the grid points are defined in terms of the molecule, thus removing the problem of rotational invariance. This is only a problem when finite grids are used. If the grid is made exhaustively large the problem begins to disappear, though this will usually be computationally infeasible.

2.10.5 Standard Quadrature Grid

Following the idea of standard basis sets, Gill *et al.* [147] have introduced the idea of a standard grid. The SG-1 grid was designed to give numerical integration errors of about 0.2 kcal mol^{-1} for medium sized molecules, while using as few grid points as possible. The grid is derived from the EML-(50,194) grid, which has 50 radial points, given by the Euler-Maclaurin rules, and 194 angular points positioned by the Lebedev rules (a Lebedev grid of degree 23).

As the nucleus is approached from the the valence region in a molecule the electron density becomes spherically symmetrical, and therefore less sophisticated angular grids should be required. The SG-1 grid partitions space into five spherical regions around the atom and then uses Lebedev grids with 6, 38, 86, 194 and 86 points respectively. The size of each region depends on the central atom. This produces about 2500 points per atom, roughly a quarter the size of the EML-(50,194) grid, yet yielding similar (within a few μ-Hartree) accuracy.

2.10.6 Standard Electronic Orientation

The previous two sections show how to 'standardize' the DFT energy of a molecule. However, it ignores the effect of a rotation of a degenerate orbital. For example, the occupied p-orbital of the Boron atom. In HF theory rotation of this orbital does not effect the energy. In DFT rotation will introduce grid effects, introducing an error dependent on the size of the grid. Obviously, with an infinite grid this problem does not exist.

This can be overcome by removal of degeneracy, forcing the orbital into a particular orientation. In Q-CHEM this is achieved by introducing a small quadrupole field [147], with the components

$$x^2 = +1\text{x}10^{-10} \tag{2.95}$$

$$y^2 = +2\text{x}10^{-10} \tag{2.96}$$

$$z^2 = -3\text{x}10^{-10} \tag{2.97}$$

The effect of the field on the total energy is negligible, but 'standardizes' the DFT energy for degenerate molecules.

2.10.7 XC Linear Scaling

A simple implementation of the above procedure will produce an algorithm which scales as $O(N^3)$. In 1993 Johnson [148] showed that the integration could be performed in only linear work, with a few modifications to the above procedure taking advantage of the fast decay of the basis functions.

For each basis function ϕ_α, centered at $\mathbf{R}_\alpha$, a sphere is defined beyond which its influence is deemed negligible. The radius of the sphere, λ_α, is found from choosing a threshold ϵ and requiring that $|\phi_\alpha(\mathbf{r})| \leq \epsilon$ for every point outside the sphere. For any grid point $\mathbf{r}_g$ all significant basis functions can be found by selecting those that fulfill

$$|\mathbf{r}_g - \mathbf{R}_\alpha| \leq \lambda_\alpha. \tag{2.98}$$

Therefore a list of all significant basis functions is created for every grid point. The important property of the list is that the number of significant basis functions at each grid point becomes independent of size for sufficiently large molecules. The construction of this list is an $O(N^2)$ process — (number of grid points) x (number of basis functions), but Stratman *et al.* [149] have noticed that the computational time involved is insignificant for even

very large molecules ($C_{384}H_{48}$). If the formation of this list becomes a problem in the future, Pérez-Jordá and Yang [150] have reformulated the algorithm into one which scales as $O(N \log N)$.

The density at any grid point is given by

$$\rho(\mathbf{r}_g) = \sum_{\alpha,\beta} P_{\alpha\beta}\phi_\alpha(\mathbf{r}_g)\phi_\beta(\mathbf{r}_g) \tag{2.99}$$

which forces the evaluation of the density to scale as $O(N^3)$. However, if the summation is restricted to only those basis functions which are significant at the grid point, the computational effort per grid point becomes independent of size (for a sufficiently large molecule) and the density evaluation should scale linearly with the molecular size.

2.11 The Kohn-Sham matrix

To minimize the energy within a basis set requires construction of the Kohn-Sham matrix, the matrix representation of the operator in equation (2.51). Writing the exchange-correlation energy as

$$E_{xc} = \int f_{xc}(\rho, \nabla\rho)\, d\mathbf{r}, \tag{2.100}$$

the exchange-correlation elements of the Kohn-Sham matrix are given by

$$F^{xc}_{\mu\nu} = \int \phi_\mu v_{xc} \phi_\nu \, d\mathbf{r} \tag{2.101}$$

with the potential defined by equation (2.50). By using calculus of variations [83] the potential can be expressed as

$$v_{xc} = \frac{\partial f_{xc}}{\partial \rho} - \nabla . \left(\frac{\partial f_{xc}}{\partial \nabla \rho} \right). \tag{2.102}$$

However, Pople *et al.* [151] pointed out that the calculus of variations procedure involves integration by parts in the $\nabla\rho$ contribution. Thus the numerical integration of equation (2.101) will have an increased error. A more consistent approach is to obtain the exchange-correlation part of the Kohn-Sham matrix from the direct minimization of the energy with respect to orbital variations, that is

$$F^{xc}_{\mu\nu} = \frac{\partial E_{xc}}{\partial P_{\mu\nu}} = \int \phi_\mu \frac{\partial f_{xc}}{\partial \rho} \phi_\nu \, d\mathbf{r} + \int \frac{\partial f_{xc}}{\partial \nabla \rho} . \nabla(\phi_\mu \phi_\nu) \, d\mathbf{r}. \tag{2.103}$$

Another advantage of this formulation is that the second-derivative of the density is no longer required, a major computational saving.

Chapter 3

Density Functional Partitions[1]

3.1 Introduction

The major disadvantage of DFT is that there is no systematic way to improve a density functional. Thus, quality information on the performance of a density functional is essential in determining how to improve functionals. It is common to compare the exchange-correlation energies from a density functional with those determined by configuration interaction (or the related perturbation or coupled-cluster theories) on a single determinant reference wavefunction. Of greater use in the design of new functionals would be the ability to examine a functional's performance for various subsets of electrons in a molecule. These partitions are common in conventional theory. For example, the correlation energy associated with inner-shell electrons is often separated and ignored (the frozen-core approximation). Also, it is often possible (particularly at the simplest MP2 level) to separate correlation between electrons of parallel ($\alpha\alpha$ or $\beta\beta$) and antiparallel ($\alpha\beta$) spin. With MP2 theory the total correlation energy can be expressed as a sum of electron-pair components, making such partitions straightforward. Another useful partition is the splitting of density into core and valence densities, allowing the examination of core-electron/core-electron, core-electron/valence-electron and valence-electron/valence-electron contributions to the correlation energy.

This chapter, following earlier work of Stoll *et al.* [152, 153] and Perdew *et al.* [154], presents similar partitions of DFT exchange-correlation energies to their conventional counterparts. This is carried out by examining electron correlation relative to the Kohn-Sham single determinant reference wavefunction. Correlation energies are obtained from conventional theory and DFT. Each may be partitioned by dividing the occupied spin orbitals into

[1] The work described in this chapter has been carried out in collaboration with Prof. John Pople

non-overlapping sets, corresponding to a division of the density into two parts, allowing the resulting energy components to be compared.

3.2 General Theory

Writing the full spin-dependent density as

$$\gamma(\mathbf{x}) = \rho_\alpha(\mathbf{r})|\alpha(s)|^2 + \rho_\beta(\mathbf{r})|\beta(s)|^2 \tag{3.1}$$

shows that electron density is readily separable into two parts. Integration over the spin coordinate s gives the regular density

$$\rho(\mathbf{r}) = \int \gamma(\mathbf{x})\, ds = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r}). \tag{3.2}$$

Remembering from Chapter 2 that the KS energy can be written as

$$E[\gamma] = T_s[\gamma] + V[\gamma] + J[\gamma] + E_{xc}[\gamma], \tag{3.3}$$

where an approximate functional is used for the exchange-correlation functional $E_{xc}[\gamma]$.

The KS treatment begins by writing γ in terms of a set of orthonormal spin orbitals, χ_i:

$$\gamma(\mathbf{x}) = \sum_{i=1}^{occ} |\chi_i|^2 \tag{3.4}$$

leading to the KS equations. The great advantage of this is that it allows the Hartree-Fock procedure to be written as a special case of KS density functional theory, simply by defining the Fock exchange-only functional for $E_{xc}[\gamma]$ as

$$E_x[\gamma] = -\frac{1}{2}\sum_{i,j}^{occ} \iint \frac{\chi_i^*(\mathbf{x}_1)\chi_j(\mathbf{x}_1)\chi_j^*(\mathbf{x}_2)\chi_i(\mathbf{x}_2)}{r_{12}}\, d\mathbf{x}_1 d\mathbf{x}_2. \tag{3.5}$$

Note that this exchange energy is defined for any appropriately normalized spin orbitals χ_i, and hence for any appropriate density $\gamma(\mathbf{x})$. Thus, the accuracy of a density produced by an exchange-correlation functional under the KS formalism can be examined by comparing its value for E_x above with the true Hartree-Fock energy. Thus, we define

$$E_{KS}[\gamma_{KS}] = T_s[\gamma_{KS}] + V[\gamma_{KS}] + J[\gamma_{KS}] + E_x[\gamma_{KS}]. \tag{3.6}$$

Also, it should be apparent that

$$E_{KS} \geq E_{HF} \tag{3.7}$$

as the HF energy is the lowest expectation energy that can be obtained from a single-determinant wavefunction. This also allows the correlation energy to be defined as the difference

$$E_c[\gamma_{KS}] = E_{xc}[\gamma_{KS}] - E_x[\gamma_{KS}]. \tag{3.8}$$

For the conventional treatment of electron correlation a set of orthonormal virtual spin orbitals, χ_a^{KS} needs to be introduced. A Fock matrix for the KS determinant can then be constructed

$$F_{pq}^{KS} = T_{pq} + V_{pq} + \sum_{i=1}^{occ}(pi||qi) \tag{3.9}$$

where $(pi||qi)$ is an antisymmetrized two-electron integral

$$(pq||rs) = \iint \chi_p(\mathbf{x}_1)\chi_r(\mathbf{x}_2)\frac{1}{r_{12}}\left[\chi_q(\mathbf{x}_1)\chi_s(\mathbf{x}_2) - \chi_s(\mathbf{x}_1)\chi_q(\mathbf{x}_2)\right]\, d\mathbf{x}_1 d\mathbf{x}_2. \tag{3.10}$$

Unless the Fock functional is used, the Fock matrix will not be diagonal. In particular, there will be nonzero elements F_{ia} connecting the occupied and virtual spin orbitals. To simply diagonalise this matrix would allow the occupied and virtual orbitals to mix, altering the density. This can be avoided by separating the matrix into two parts,

$$F^{KS} = F^{KS}(\mathrm{OO} + \mathrm{VV}) + F^{KS}(\mathrm{OV}) \tag{3.11}$$

corresponding to nonzero occupied-occupied, virtual-virtual blocks in the first part and nonzero occupied-virtual blocks in the second. The matrix $F^{KS}(\mathrm{OO} + \mathrm{VV})$ can then be diagonalised, providing a new set of spin orbitals that could be described as the canonical Fock orbitals for the constrained KS determinant. These new occupied orbitals will be an orthogonal transformation of the KS occupied spin orbitals and will yield the same density, that is

$$\gamma = \sum_{i=1}^{occ}|\chi_i|^2 = \sum_{i=1}^{occ}|\chi_i^{KS}|^2 = \gamma^{KS}. \tag{3.12}$$

The single determinant wavefunction formed from the occupied χ_i, written here as Ψ_0, will be equal to Ψ_{KS}.

Second-order Møller-Plesset theory has been used to determine the conventional correlation energy, starting from the Ψ_0 wavefunction. At first order, the MP energy is simply E_{KS} and the second-order correction is

$$E^{(2)} = -\sum_{ia}\frac{F_{ia}^2}{\epsilon_a - \epsilon_i} - \frac{1}{4}\sum_{ijab}\frac{(ij||ab)^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}. \tag{3.13}$$

Note that F_{ia} appears in equation (3.13) as the off-diagonal elements are part of the perturbation Hamiltonian. If the Fock functional is used to form the KS density all F_{ia} will vanish and equation (3.13) will reduce to the more familiar form of equation (1.75). This new term again essentially allows for a mixing of the occupied and virtual orbitals leading to a tendency for the KS orbitals to move toward HF, therefore modifying the density. Also, the F_{ia} contribution to $E^{(2)}$ is much smaller than the second part of equation (3.13), and have therefore been omitted here.

3.3 Energy Partitions

Suppose that the set of occupied spin orbitals χ_i are split into two non-overlapping subsets χ_i^A and χ_i^B. Since the density γ is the sum of the squares of the spin orbitals, it will consequently be partitioned into two parts γ^A and γ^B with

$$\gamma(\mathbf{x}) = \gamma^A(\mathbf{x}) + \gamma^B(\mathbf{x}). \tag{3.14}$$

Any energy functional $E[\gamma^A, \gamma^B]$ can then be split into 'pure A', 'pure B' and 'interacting AB' parts by the partition

$$E^A = E[\gamma^A, 0] \tag{3.15}$$

$$E^B = E[0, \gamma^B] \tag{3.16}$$

$$E^{AB} = E[\gamma^A, \gamma^B] - E^A - E^B. \tag{3.17}$$

Such a partition has been proposed by Stoll *et al.* [152, 153] for the spin components, but it can be applied elsewhere. A similar partition for the conventional energies can be achieved by treating the energies as functionals of the sets χ_i^A and χ_i^B

$$E^A = E[\chi_i^A, 0] \tag{3.18}$$

$$E^B = E[0, \chi_i^B] \tag{3.19}$$

$$E^{AB} = E[\chi_i^A, \chi_i^B] - E^A - E^B \tag{3.20}$$

where $E[\chi_i^A, 0]$ denotes that all integrals involving B spin orbitals are zeroed.

The above partition will cleanly split the MP2 correlation energy and also the exchange energy of the Fock functional. At higher levels of correlation treatment the partitions via equation (3.18) become questionable. There are numerous complicated many-body interactions involved which would be somewhat arbitrarily assigned. However, this is not addressed

here, as only MP2 computations are performed. It seems reasonable that the DFT and conventional partitions should be comparable at a coarse level, since MP2 usually accounts for the majority of correlation in simple pair terms.

3.4 Application

The partition schemes described above have been implemented within the Q-CHEM program. To demonstrate the use of partitioning, the LSDA functional has been tested for the first-row atoms and the molecules H_2, N_2, F_2, FH, OH_2, NH_3, and CH_4. The orbital basis used is 6-311+G(3df,2p), with the standard SG-1 quadrature grid. The geometries for the molecules are those of MP2 with the 6-31G(d) basis set.

The total energies are listed in Table 3.1. The first column is the energy obtained by solving the KS equations for the LSDA functional. The other columns are energies using this density. The second column gives the sum of the first three terms in equation (3.3) and the third is this with the Dirac exchange energy added. The final column lists E_{KS} as defined by equation (3.6).

The differences $E_{KS} - E_{HF}$, where E_{HF} are regular HF calculations with the same orbital basis and geometry, are listed in Table 3.2. They arise because of differences between the LSDA and HF densities. Note that, as expected, $E_{KS} \geq E_{HF}$. The magnitudes of these differences are small when compared with correlation energies, but clearly not negligible. This must be taken into account when comparing DFT correlation energies with conventional values based on an HF starting point.

The first partition results are presented in Table 3.3. The LSDA correlation functional has been partitioned with $A = \alpha$ and $B = \beta$. The MP2 correlation energies starting from the KS density have also been partitioned. Note that these are full correlation energies, taking account of all electron pairs.

Table 3.4 lists the DFT correlation energies, split into core-core, core-valence, and valence-valence partitions, along with the total DFT correlation energy. The valence-valence conventional correlation energy is also listed. The conventional core-core and core-valence partitions are not listed, as the orbital basis gives a very poor description of inner-shell correlation, limiting the significance of the results. The valence-valence results using QCISD(T) based on the HF starting point are also listed for comparative purposes. They have been incremented by the values in Table 3.2, thus allowing for the KS reference.

Table 3.1: Total Energies for Partition Molecules (hartrees)

	LSDA	NoXC	HFS	E_{KS}
H	-0.47835	-0.19921	-0.45663	-0.49894
He	-2.83256	-1.85619	-2.72081	-2.85821
Li	-7.34221	-5.67152	-7.19113	-7.43001
Be	-14.44484	-11.92421	-14.22024	-14.56933
B	-24.35242	-20.79262	-24.06338	-24.52534
C	-37.46510	-34.65274	-37.10778	-37.68297
N	-54.12937	-47.83831	-53.70141	-54.39175
O	-74.51998	-66.65201	-73.98786	-74.79881
F	-99.09882	-89.42861	-98.46121	-99.38934
Ne	-128.21133	-116.49864	-127.46811	-128.51382
H_2	-1.13674	-0.48254	-1.04256	-1.13229
N_2	-108.68189	-95.94700	-107.73911	-108.94796
F_2	-198.31988	-178.89857	-197.02225	-198.70752
FH	-99.83580	-89.70121	-99.13477	-100.04198
OH_2	-75.90141	-67.16846	-75.24027	-76.04380
NH_3	-56.10183	-48.59154	-55.47729	-56.20576
CH_4	-40.11538	-33.66071	-39.52360	-40.19840

Table 3.2: Energy Differences $E_{KS} - E_{HF}$ (mhartrees)

H	0.87	Ne	12.66
He	1.69	H_2	0.71
Li	2.02	N_2	24.59
Be	2.61	F_2	33.78
B	5.74	FH	13.35
C	7.28	OH_2	12.72
N	7.14	NH_3	12.40
O	10.53	CH_4	13.82
F	12.47		

Table 3.3: Spin Components of the Correlation Energy (mhartrees)

	DFT(LSDA)			conventional(MP2)		
	$\alpha\alpha$	$\beta\beta$	$\alpha\beta$	$\alpha\alpha$	$\beta\beta$	$\alpha\beta$
H	21.72	0	0	0	0	0
He	26.89	26.89	57.96	0	0	29.61
Li	47.12	32.36	71.59	0.16	0	13.43
Be	54.20	54.20	116.21	0.25	0.25	43.13
B	83.70	59.77	145.57	4.29	0.33	58.84
C	118.35	64.26	174.71	14.81	0.36	72.95
N	157.03	68.06	202.87	32.76	0.38	85.06
O	165.71	101.04	265.36	34.75	5.45	129.71
F	173.04	138.46	326.11	36.68	18.05	175.56
Ne	179.45	179.45	384.32	38.51	38.51	221.25
H_2	22.66	22.66	48.85	0	0	29.87
N_2	227.41	227.41	487.97	56.32	56.32	339.76
F_2	313.18	313.18	671.27	75.10	75.10	456.55
FH	169.16	169.16	362.71	38.91	38.91	230.33
OH_2	159.44	159.44	342.25	35.58	35.58	223.61
NH_3	150.54	150.54	323.46	29.30	29.30	204.97
CH_4	142.58	142.58	306.62	21.15	21.15	180.98

Table 3.4: Core-Valence Components of the Correlation Energy (mhartrees)

	DFT(LSDA)				conventional	
	total	core	core-val.	val.	val.(MP2)	val.(QCISD(T))
H	21.71	0	0	21.72	0	0
He	111.75	0	0	111.75	29.61	37.57
Li	151.08	134.28	2.61	14.18	0	0
Be	224.60	150.22	3.01	71.38	28.10	48.04
B	289.04	162.73	5.27	121.04	45.67	71.74
C	357.32	173.02	8.23	176.08	68.63	96.22
N	427.96	181.76	11.59	234.60	97.22	120.97
O	532.11	189.37	10.72	332.03	147.87	73.22
F	637.60	196.10	10.51	430.99	207.12	228.48
Ne	743.22	202.14	10.93	530.15	273.92	285.20
H_2	94.18	0	0	94.18	29.87	38.52
N_2	942.79	363.35	15.49	563.95	407.24	424.44
F_2	1297.62	392.17	19.70	885.75	559.73	582.74
FH	701.03	196.08	9.92	495.03	284.32	296.02
OH_2	661.14	189.34	8.87	462.93	271.61	288.76
NH_3	624.54	181.71	7.80	435.03	240.67	265.37
CH_4	591.77	172.90	6.74	412.13	200.65	234.61

With the use of partitioning twice, the $\alpha\alpha$, $\beta\beta$ and $\alpha\beta$ spin components of the valence-only correlation energy can be calculated. These, along with the corresponding MP2 partitions are listed in Table 3.5.

Table 3.5: Spin Components of the Valence-Valence Correlation Energy (mhartrees)

	DFT(LSDA)			conventional(MP2)		
	$\alpha\alpha$	$\beta\beta$	$\alpha\beta$	$\alpha\alpha$	$\beta\beta$	$\alpha\beta$
Li	14.18	0	0	0	0	0
Be	17.23	17.23	36.93	0	0	28.10
B	43.08	19.63	58.33	3.50	0	42.16
C	74.61	21.53	79.93	13.41	0	55.22
N	110.59	23.14	100.87	30.74	0	66.48
O	117.26	53.60	161.17	32.68	4.49	110.71
F	122.84	88.75	219.40	34.55	16.48	156.10
Ne	127.70	127.70	274.75	36.33	36.33	201.25
N_2	135.65	135.65	292.64	52.19	53.19	300.86
F_2	213.23	213.23	459.29	71.25	71.25	417.24
FH	119.17	119.17	256.69	36.89	36.89	210.54
OH_2	111.39	111.39	240.16	33.76	33.76	204.09
NH_3	104.64	104.64	225.75	27.62	27.62	185.44
CH_4	99.12	99.12	213.90	19.60	19.60	161.45

3.5 Discussion

The familiar overestimation of the LSDA functional is demonstrated by the total correlation energies in Table 3.4. These results are close to previously published LSDA correlation energies [155]. Correct total atomic correlation energies (relative to the HF reference) are known to range from 42 mhartrees for helium to 392 mhartrees for neon [156]. Thus, LSDA is overestimating by roughly a factor of two.

The work of Stoll *et al.* [152, 153] closely resembles Table 3.3. They partitioned the LSDA correlation energy in this manner, but then compared just the $\alpha\beta$ component to the *total* correlation energy from conventional theory. However, as can be seen from the first two columns of Table 3.3, the $\alpha\alpha$ and $\beta\beta$ components are too large to ignore. The spin-parallel

contributions are partly spurious, as indicated by the significant nonzero values for H, He and H_2. These represent a self-correlation for single electrons. These effects are undoubtedly present in the larger systems, contributing to their large $\alpha\alpha$ and $\beta\beta$ values — far greater than their corresponding conventional results, even accounting for the poor treatment of core electrons with MP2 for this basis. The conventional results do show that spin-parallel effects are significant, beyond that accounted for by the exchange term.

The spin-antiparallel components, shown in Table 3.3, are still significantly larger for LSDA than MP2. For the helium atom, where there is only $\alpha\beta$ correlation, the LSDA value of 58 mhartrees is considerably larger than the known accurate value of 42 mhartrees, even after accounting for the 1.7 mhartree correction of Table 3.2, but no longer by a factor of two. The larger systems show even greater overestimation, but note that the MP2 results will be underestimates, due to the crude description of the inner-shell conventional correlation.

If the LSDA error was roughly constant for each atom, a convenient cancellation of errors would occur when examining most chemical properties. To see if this is the case the contributions of parallel and antiparallel correlation to chemical binding energies are listed in Table 3.6, along with their conventional counterparts.

Table 3.6: Spin Components of the Correlation Binding Energy (mhartrees)

	DFT(LSDA)			conventional(MP2)		
	$\alpha\alpha+\beta\beta$	$\alpha\beta$	total	$\alpha\alpha+\beta\beta$	$\alpha\beta$	total
H_2	1.88	48.85	50.74	0	29.87	29.87
N_2	4.64	82.23	86.87	46.37	169.64	216.01
F_2	3.37	19.05	22.41	40.74	105.44	146.18
FH	5.10	36.60	41.70	23.09	54.77	77.86
OH_2	8.69	76.89	85.58	30.96	93.89	124.85
NH_3	10.83	120.59	131.42	25.46	119.91	145.37
CH_4	15.65	131.91	147.56	27.13	108.03	135.16

The conventional total correlation contribution shows that electron correlation does play a major role in binding. The DFT totals are mostly lower (with the exception of H_2 and CH_4), with F_2 being remarkably lower. When broken down into the spin components, the contributions of parallel spins are quite small by LSDA, even though the individual parallel values of Table 3.3 are large. The MP2 results indicate that LSDA underestimates the paral-

lel contribution (except for H_2, where there is none). The LSDA $\alpha\beta$ correlation contributions to binding show improvement when compared with the general overestimation of total correlation energies by a factor of two. However, there are wide variations with the type of bond. LSDA describes moderately well the $\alpha\beta$ contributions to bonds involving hydrogen. Yet the triple bond in N_2, where antiparallel correlation in the three pairs is a major stabilizing factor, shows an LSDA underestimation by more than a factor of two. This is particularly disturbing, remembering that the total correlation energies are overestimated by about this factor. The LSDA functional also badly underestimates the $\alpha\beta$ correlation contribution to binding in the F_2 molecule. F_2 is bound by LSDA (just), however 99 mhartrees of the 122 mhartrees come from the exchange part, when conventional HF does not bind F_2.

Turning to the core-valence separations of Table 3.4, the first thing to note is that the inter core-valence correlation energies are small. This is not surprising as the respective orbitals are principally located in different spatial regions. Comparison with good conventional numbers is not really possible, due to the deficiencies of the basis set. On the whole, the LSDA functional shows a good separation of correlation energy into the core and valence regions.

Although good MP2 core-core correlation numbers are unavailable, the atomic LSDA values should be close to those for the corresponding two-electron ions (He, Li^+, Be^{2+}, B^{3+}, ...). It is known that these remain fairly constant, approaching a limit of about 46 mhartrees [156]. The LSDA numbers here are larger by up to a factor of four, and show a pattern of increasing steadily. This is obviously a major contribution to the overestimation of total correlation energies by the LSDA functional.

The valence-valence LSDA correlation energies are also too large, by roughly a factor of two. Again, the inadequate MP2 basis set provides a worrying underestimation, yet the results are similar for the more sophisticated QCISD(T) numbers.

The LSDA functional provides roughly a constant error for the core-core correlation components in moving from atom to molecule. Thus the failure to accurately describe the inner-shell electrons is not related to any failures in the description of chemical bonding. This is also true of the core-valence component. This suggests that the valence-only behaviour of the LSDA functional is more important. The spin-component analysis of the valence correlation energies is summarized in Table 3.5. The orbital basis is of better quality in the valence region, allowing a more satisfactory comparison with conventional MP2 results.

The spin-parallel components of valence correlation energies are still too large, and again show spurious self-correlation effects (for example Li and Be, where the valence correlation

should be zero). The LSDA $\alpha\beta$ terms, however, do show a far better agreement with MP2 than the all-electron results. LSDA is too large again, but only by a factor of about 1.5, not two as before.

The partitions can be continued to examine the spin components of the binding energy using only the valence density, which are listed in Table 3.7.

Table 3.7: Spin Components of the Valence Correlation Binding Energy (mhartrees)

	DFT(LSDA)			conventional(MP2)		
	$\alpha\alpha + \beta\beta$	$\alpha\beta$	total	$\alpha\alpha + \beta\beta$	$\alpha\beta$	total
N_2	3.85	90.89	94.74	44.89	167.90	212.79
F_2	3.28	20.49	23.77	40.45	105.04	145.48
FH	5.02	37.29	42.31	22.76	54.44	77.20
OH_2	8.47	78.99	87.46	30.36	93.38	123.73
NH_3	10.39	124.88	135.27	24.49	118.95	143.44
CH_4	15.20	133.97	149.17	25.80	106.23	132.02

As expected, the performance of the valence-only theory (for binding) is similar to the all-electron results of Table 3.6. The correlation bindings are mostly too small and there is incorrect division between the spin-parallel and antiparallel components. This is consistent with a good cancellation of errors with the inner-shell contributions in moving from atoms to molecules.

Finally, in Table 3.8, the spin components of the correlation energy contributions to the ionization energies are presented. The total LSDA contributions of the five molecules studied are all about 60 mhartrees. This bears no resemblance to the large variations seen for the conventional values. The LSDA parallel and antiparallel components are roughly constant as well, with antiparallel providing the majority of the total. This, again is in contrast to the MP2 results, which show a roughly equal parallel/antiparallel contribution, with the components changing in similar ways to the total. Table 3.8 demonstrates a failure of the LSDA functional.

3.6 Conclusions

The work presented here shows that it is possible to compare the energy partitions of density functionals with more conventional theories. This is achieved by using the Kohn-

Table 3.8: Spin Components of the Correlation Ionization Potentials (mhartrees)

	DFT(LSDA)			conventional(MP2)		
	$\alpha\alpha+\beta\beta$	$\alpha\beta$	total	$\alpha\alpha+\beta\beta$	$\alpha\beta$	total
N_2	24.92	32.25	57.17	-12.32	-8.86	-21.18
FH	26.00	36.08	62.08	24.02	21.22	45.24
OH_2	24.27	34.31	58.58	21.43	19.78	41.21
NH_3	22.43	32.58	55.01	17.27	17.14	34.41
CH_4	23.33	32.20	55.53	9.45	11.49	20.93

Sham single determinant as a reference. It is possible to use the same procedure to determine the spin orbital or density partitions of core/valence or α spin/β spin (or indeed any partition into non-overlapping spin orbitals).

When applied to the LSDA functional, spin partition shows a strong overestimation of the parallel spin components (including spurious one-electron effects). Antiparallel correlation is also overestimated, but by a smaller factor. The core/valence partitioning reveals core correlation values which are much too large. These values also do not approach any limiting value as they should. One bonus of the LSDA functional, though, is that these core values remain fairly constant in moving from atom to molecule, allowing a systematic cancellation of errors for most chemical properties. The LSDA functional does, however, show better performance in the valence region, overestimating the antiparallel contributions by only a factor of two (there are still major errors for the parallel components here). However, the major failing of the LSDA functional is the relative contributions of $\alpha\alpha+\beta\beta$ and $\alpha\beta$ correlation to binding energies. The $\alpha\beta$ contribution to binding is usually too low, even though the LSDA functional generally overestimates the binding energy.

The partitioning method developed here is applicable to any energy functional, allowing a more detailed evaluation of a functional than was previously possible. Hopefully the study of the piecewise inadequacies of future functionals will allow the convergence towards an accurate, correct energy functional. The LYP functional, which is used to develop the empirical density functional of the following chapter, for example, does not have the spurious one-electron effects seen here for the LSDA functional.

Chapter 4

Empirical Density Functionals[1]

4.1 Introduction

Exchange-correlation functionals are often tested by systematic comparison of computed total energies with high-quality experimental data. The most common test set used for such studies [98, 100, 106, 155] is the G2 set of atomization energies, ionization potentials, electron affinities and proton affinities [52]. These tests showed a significantly improved agreement with experiment by the introduction of local density gradients into the functionals. Recently, a portion of the Fock exchange has been included with the exchange functional [119]. With adjustment of some of the parameters involved, the mean absolute deviations between calculation and experiment for chemical energetics begin to approach that for G2 theory.

This chapter uses the G2 experimental data to examine some of the common functionals in use today. This is carried out with three main objectives in mind. Firstly, the emphasis on empirical parameterization is increased. Most previous functionals have first been constrained to satisfy certain limiting conditions, for example the correct behaviour for the uniform electron gas and the scaling at large distances, and then use parametization for the middle ranges. By removing these constraints a functional of high practical value may be attainable. The results may also give some indication of the changes to functionals that are implied by the experimental data, and thus point the way for future functional improvements.

Previously it has been implicitly assumed that a functional designed to be optimal for large basis sets will be equally suitable for small basis set calculations. In fact, E. Bright Wilson's argument (see section 2.2) holds for the density from a small basis set, as well as an infinite one. Yet one would expect the ultimate functionals for the differing basis sets to be

[1] The work described in this chapter has been carried out in collaboration with Prof. John Pople

quite different. A functional designed for smaller systems would also have a high practical value as it would allow large systems (which require small basis sets) to be studied more accurately.

The third objective is to examine the necessity of including Fock exchange in order to obtain good agreement with experiment. Becke has alleged that "a small exact-exchange component is a natural and necessary constituent of any exchange-correlation approximation aiming for accurate molecular energetics" [119]. However, doing so introduces non-local effects and consequent computational complications [157, 158], and should therefore only be included if it is really needed.

4.2 General Theory

The orbital basis used throughout this chapter is 6-31+G*. This is small enough to allow the study of large molecules, yet also has enough complexity to contain the general features required of a basis set [100]. The G2 geometries are the MP2/6-31G* geometries, and zero-point vibrational corrections are calculated from Hartree-Fock harmonic frequencies. From the original G2 set the two excited states, N_2^+ and SH_2^+, have been removed. The atomization energy and proton affinity of H_2 and the ionization potentials of inert gas atoms have been added. Also the electron affinity of the H atom has been excluded, as the 6-31+G* basis set does not place a diffuse orbital on hydrogen.

The quality of a functional is judged by the root-mean-square (RMS) deviation of the computed results from the experimental values for the 129 data points. This requires the evaluation of self-consistent energies on 150 atoms and molecules. The RMS deviation is minimized with respect to parameters included in the functionals. Linear combinations of different functionals and combinations of the same functional with different parameters are considered. This optimization can be carried out in two ways. In the first, termed 'internal' optimization, the full functional is written as a linear combination of component functionals

$$E_{xc} = \int \sum_i c_i f_i(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)\, d\mathbf{r} \tag{4.1}$$

with adjustable coefficients c_i. This functional is then used to calculate the 150 self-consistent energies, leading to an RMS deviation from experiment, which is minimized with respect to the c_i. The Hartree-Fock exchange can be included by adding another term to this sum with an additional c coefficient. This is the conventional mixing method introduced by Becke [119].

The second method used is termed 'external' optimization. Here the 150 energies are calculated for each of the component functionals

$$E_{xc}^{i} = \int f_i(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)\, d\mathbf{r}. \tag{4.2}$$

Note that, if there are n functionals, this will require the $150n$ single point Kohn-Sham calculations. The energies for each functional can be arranged as n vectors $\mathbf{E}^i$ and then combined, using coefficients c_i, to give a single set of 150 energies as a vector

$$\mathbf{E}(\mathbf{c}) = \sum_{i}^{n} c_i \mathbf{E}^i. \tag{4.3}$$

This mixture is, in effect, a 'linear combination of model chemistries' defined by the coefficients c_i. As before, the RMS deviation from experiment is minimized with respect to the c_i.

While the second type could be used as a model, the internal optimization is clearly preferable as only one calculation is required (compared with n calculations). However optimization by the 'external' method is much faster as it only involves quadratic minimization of the coefficients. Thus the external optimization can be used first as a pointer to worthwhile candidates for internal optimization.

4.3 Results

The BLYP functional is known to perform quite well on the G2 set [98,100,106], making it an ideal starting point for the optimizations. The results of modifying the BLYP parameters are listed in Table 4.1.

Table 4.1: Modified BLYP RMS deviations (kcal/mol)

Functional	RMS deviation
B(0.0042)-LYP	5.290
B(0.0035)-LYP	5.069
B(0.0042) + 1.0431LYP	4.963
B-LYP(optimum)	4.848

The unmodified BLYP functional (with Becke's β parameter of 0.0042 in the exchange part) gives an RMS error of 5.290 kcal/mol. The corresponding mean absolute deviation is 4.11 kcal/mol. Next, the Becke β parameter was optimized, moving to 0.0035, giving an RMS

deviation of 5.069 kcal/mol. Third, an optimal linear combination of the original Becke and LYP parts is found, lowering the RMS deviation to 4.963 kcal/mol. A final, fuller optimization only lowers the RMS error to 4.848 kcal/mol. (This has a Dirac coefficient of 1.0072 times the original, a β value of 0.003705 and LYP parameters of $a, b, c, d = 0.049, 0.108, 0.24, 0.342$.)

The next step is to see the effect of the different β values through external mixing. By taking a combination of B(0.0035)-LYP, B(0.0042)-LYP, Hartree-Fock-Slater and Hartree-Fock-Becke theories the RMS deviation is lowered to 4.543 kcal/mol. Clearly the linear combination of two Becke functionals is a better exchange functional than either separately. The external optimization coefficients are

$$9.558\text{B}(0.0035)\text{LYP} - 7.727\text{B}(0.0042)\text{LYP} - 0.861\text{HFS} + 0.743\text{HFB}(0.0042). \tag{4.4}$$

These coefficients were then refined via an internal optimization, along with a redetermination of the LYP parameters. The complete functional is then

$$\begin{aligned} E_{xc} = {} & 1.030952 E_x^{D30} + 10.4017 \Delta E_x^{B88}(0.0035) - 8.44793 \Delta E_x^{B88}(0.0042) \\ & + E_c^{LYP}(0.055, 0.158, 0.25, 0.3505) \end{aligned} \tag{4.5}$$

where ΔE_x^{B88} represents the Becke correction, that is the B88 functional without the Dirac term. With this functional the RMS deviation is lowered to 4.237 kcal/mol and the mean absolute deviation is 3.215 kcal/mol. The non-LYP part of the above functional has been termed the double-Becke functional, while the complete functional is called 'Empirical Density Functional 1' or EDF1.

Table 4.2: External linear mixing of Hartree-Fock (kcal/mol)

Functional Combination	RMS deviation
BLYP + HFB(0.0042) + HFS	4.920
BLYP + HFB(0.0042) + HFS + HF	4.499
EDF1 + HFB(0.0042) + HFB(0.0035) + HFS + BLYP	4.211
EDF1 + HFB(0.0042) + HFB(0.0035) + HFS + BLYP + HF	4.211

The third step is to investigate the effect of adding a fraction of Fock exchange to EDF1. Table 4.2 gives the results of external mixing of Hartree-Fock with BLYP and EDF1 plus their components. Clearly, there is strong HF mixing with BLYP, but virtually none with EDF1. This is confirmed by trying to mix in HF with an internal optimization of the

EDF1 components and Fock exchange. The HF coefficient is less than 0.001 and there is no significant lowering of the RMS error.

To asses the usefulness of EDF1 as a density functional it has been used to obtain the thermochemistry of the molecules in the G2 set, using the 6-31G* basis. The results are listed in Table 4.3 and Table 4.4. In addition, the BLYP and B3LYP functionals have been included for comparison. Perhaps surprisingly, B3LYP performs quite poorly, being inferior to both BLYP and EDF1. However, the parametization of the B3LYP functional was carried out using what is, in effect, an infinite basis. It is therefore not unreasonable to expect performance to improve if it was reoptimized for the 6-31G* basis. An internal reoptimization along these lines shows a decrease in the Fock exchange coefficient to about 5%, with an RMS error of 4.65 kcal/mol, still considerably inferior to EDF1.

The overall improvement in moving from B3LYP (or BLYP) to EDF1 is largely due to the better atomization energies and proton affinities. The electron addition and removal energies are only slightly superior. It is interesting that the worst EDF1 results (atomization energies of SO_2 and O_2, ionization energy of O_2, electron affinity of Cl_2, and proton affinity of H_2) are also problematic cases for BLYP and B3LYP. This confirms the underlying similarity of each of the three functionals, that they are all made from essentially the same main components.

By writing density functionals in the form

$$E_{xc}[\rho] = \int \rho^{4/3}(\mathbf{r}) g(x)\, d\mathbf{r} \tag{4.6}$$

the double-Becke functional can be compared with the original B88 form. The B88 $g(x)$ is

$$g_{B88}(\beta, x) = C_0 - \frac{\beta x^2}{1 + 6\beta x \sinh^{-1}(x)} \tag{4.7}$$

with $\beta = 0.0042$ and C_0 is the coefficient of the Dirac functional. The new $g(x)$ is

$$g_{doubleB}(x) = 1.030952 C_0 + 10.4017 g_{B88}(0.0035, x) - 8.44793 g_{B88}(0.0042, x) \tag{4.8}$$

These two functions are plotted in Figure 4.1.

At $x = 0$ the double-Becke value is slightly below that of the uniform electron gas. The double-Becke curve is also much flatter at the origin. This can be seen from the initial term in the Taylor expansion

$$g(x) = g(0) + \frac{1}{2} g''(0) x^2 + \ldots \tag{4.9}$$

which is $g''(0) = -0.00184$, compared with -0.0084 from the original B88 form. The double-Becke value is much closer to the Sham-Kleinman value of -0.00387 . Also, the double-Becke curve is below B88 until $x = 4$ and the curves cross again at $x = 9.7$.

Table 4.3: Deviations from experiment for various functionals

	Exp.	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
atomization energies (kcal/mol)				
H_2	103.3	-0.1	-0.6	-3.1
LiH	56.0	1.0	1.0	1.3
BeH	46.9	-7.0	-7.9	-7.2
CH	79.9	-0.3	0.3	-0.4
$CH_2(^3B_1)$	179.6	1.9	0.1	-2.3
$CH_2(^1A_1)$	170.6	4.0	3.8	2.2
CH_3	289.2	1.5	-0.8	-3.7
CH_4	392.5	3.9	0.6	-3.7
NH	79.0	-3.5	-1.6	-2.7
NH_2	170.0	-2.6	-0.2	-2.6
NH_3	276.7	2.6	4.4	-0.2
OH	101.3	0.7	2.8	0.4
OH_2	219.3	7.6	10.3	4.6
FH	135.2	5.3	7.7	2.9
$SiH_2(^1A_1)$	144.4	1.5	0.2	-0.6
$SiH_2(^3B_1)$	144.4	1.5	0.2	-0.6
SiH_3	214.0	4.4	1.0	-0.3
SiH_4	302.8	6.0	1.0	0.7
PH_2	144.7	-0.7	-1.0	-2.8
PH_3	227.4	4.4	3.0	0.4
SH_2	173.2	7.5	7.1	3.4
ClH	102.2	6.5	6.4	3.5
Li_2	24.0	4.2	4.2	5.4
LiF	137.6	2.3	5.8	4.4
HCCH	388.9	8.4	11.3	4.1
H_2CCH_2	531.9	7.0	5.6	-1.3
H_3CCH_3	666.3	9.5	3.7	-3.0
CN	176.6	-3.7	7.8	-2.1
HCN	301.8	-2.1	6.6	-1.5

Table 4.3 (continued)

	Exp.	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
CO	256.2	2.5	10.5	2.4
HCO	270.3	-4.4	3.6	-6.3
H_2CO	357.2	-1.0	4.9	-4.5
H_3COH	480.8	6.9	7.8	-0.1
N_2	225.1	-4.5	8.8	0.2
H_2NNH_2	405.4	-0.1	5.2	-2.4
NO	150.1	-8.5	3.6	-6.7
O_2	118.0	-13.1	1.3	-13.9
HOOH	252.3	0.5	11.3	-0.2
F_2	36.9	-9.1	3.8	-7.3
CO_2	381.9	-2.8	11.9	-5.4
Na_2	16.6	-0.7	-0.1	1.3
Si_2	74.0	1.9	9.0	0.8
P_2	116.1	2.5	10.4	3.7
S_2	100.7	1.8	7.4	0.0
Cl_2	57.2	7.7	10.9	6.6
NaCl	97.5	7.4	6.7	6.7
SiO	190.5	3.6	12.7	6.2
SC	169.5	3.6	10.8	2.6
SO	123.5	-1.1	9.0	-1.4
ClO	63.3	-3.5	5.8	-2.9
ClF	60.3	-0.9	6.1	-0.7
CH_3Cl	371.0	7.1	5.1	-0.2
Si_2H_6	500.1	13.2	4.5	4.1
CH_3SH	445.1	11.1	8.4	2.1
HOCl	156.3	4.1	10.8	3.0
SO_2	254.0	16.3	35.9	15.3
ionization potentials (eV)				
H	13.60	0.12	0.08	0.06
He	24.59	-0.12	-0.19	-0.20

Table 4.3 (continued)

	Exp.	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
Li	5.39	-0.13	-0.15	-0.12
Be	9.32	0.33	0.29	0.36
B	8.30	-0.27	-0.30	-0.26
C	11.26	-0.13	-0.19	-0.18
N	14.54	-0.01	-0.09	-0.13
O	13.61	-0.54	-0.43	-0.31
F	17.42	-0.33	-0.26	-0.24
Ne	21.56	-0.22	-0.17	-0.28
Na	5.14	-0.19	-0.19	-0.07
Mg	7.65	0.02	0.00	0.14
Al	5.98	0.11	0.04	0.04
Si	8.15	0.20	0.11	0.12
P	10.49	0.29	0.17	0.16
S	10.36	-0.01	-0.05	0.01
Cl	12.97	0.07	0.00	0.04
Ar	15.76	0.11	0.01	0.00
CH_4	12.62	0.13	0.00	0.07
NH_3	10.18	0.19	0.20	0.13
OH	13.01	-0.11	-0.06	-0.04
OH_2	12.62	0.17	0.19	0.11
FH	16.04	0.08	0.11	0.00
SiH_4	11.00	0.21	0.02	0.18
PH	10.15	0.15	0.04	0.03
PH_2	9.82	0.05	-0.05	-0.05
PH_3	9.87	0.18	0.14	0.17
SH	10.37	0.09	0.03	0.08
SH_2	10.47	0.24	0.16	0.18
ClH	12.75	0.19	0.10	0.10
HCCH	11.40	0.40	0.36	0.32
H_2CCH_2	10.51	0.35	0.35	0.28

Table 4.3 (continued)

	Exp.	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
CO	14.01	0.01	-0.12	0.07
N_2	15.58	0.24	-0.15	0.21
O_2	12.07	-0.47	-0.79	-0.53
P_2	10.53	0.32	-0.36	0.20
S_2	9.36	-0.02	-0.26	-0.13
Cl_2	11.50	0.30	0.04	0.23
ClF	12.66	0.19	-0.02	0.16
SC	11.33	-0.06	-0.14	-0.05
electron affinities (eV)				
C	1.26	-0.07	-0.02	-0.06
CH	1.24	-0.07	-0.02	-0.07
CH_2	0.65	-0.07	0.05	0.05
CH_3	0.08	0.14	0.24	0.21
CN	3.82	-0.06	-0.18	0.02
NH	0.38	-0.07	0.09	0.07
NH_2	0.74	0.08	0.21	0.15
NO	0.02	-0.37	-0.36	-0.30
O	1.46	-0.26	-0.06	-0.11
OH	1.83	0.02	0.19	0.08
O_2	0.44	-0.14	-0.11	0.07
F	3.40	-0.22	-0.02	-0.17
Si	1.38	0.19	0.13	0.18
SiH	1.28	0.14	0.09	0.13
SiH_2	1.12	0.08	0.04	0.08
SiH_3	1.44	0.11	0.11	0.17
P	0.75	-0.10	-0.07	-0.02
PH	1.00	0.01	0.02	0.08
PH_2	1.26	0.12	0.11	0.16
PO	1.09	-0.05	-0.17	-0.06
S	2.08	-0.04	-0.03	0.02

Table 4.3 (continued)

	Exp.	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
SH	2.31	0.07	0.07	0.09
S_2	1.66	0.07	-0.03	0.12
Cl	3.62	0.01	-0.01	0.00
Cl_2	2.39	-0.69	-0.70	-0.56
proton affinities (kcal/mol)				
H_2	100.8	10.7	11.5	8.5
HCHH	152.3	0.5	-0.2	-2.5
NH_3	202.5	1.1	-0.2	-1.7
H_2O	165.1	5.3	4.1	2.9
SiH_4	154.0	4.0	5.7	2.9
PH_3	187.1	4.0	2.8	1.5
H_2S	168.8	3.4	3.5	1.0
HCl	133.6	5.6	6.5	3.4

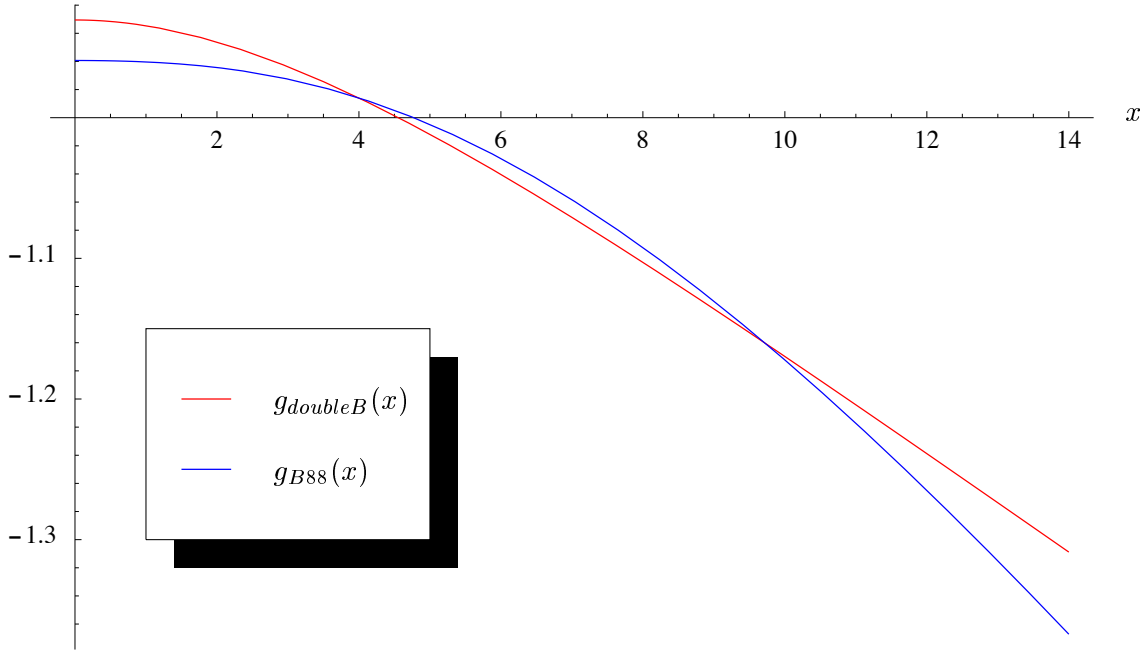
Table 4.4: RMS errors for functionals

RMS Errors (kcal/mol)	Exp.-BLYP	Exp.-B3LYP	Exp.-EDF1
atomization energies	5.75	8.10	4.41
ionization potentials	5.14	5.05	4.34
electron affinities	4.39	4.40	3.79
proton affinities	5.23	5.51	3.77

4.4 Conclusions

There are three principal conclusions from the work presented in this chapter. Firstly, the $g(x)$ from the B88 functional differs quite considerably from that obtained by optimizing with respect to experimental energies. A new functional, EDF1, has been found which, although more empirical than those before, has been designed with a small basis set in mind. Thus, it should be suitable for computations on large molecules. For the 6-31+G* basis set it appears to be more accurate than the extremely popular B3LYP functional. Finally, there is no significant improvement when the Fock exchange functional is added to the empirical mix. It should be emphasized that this result is specific to the small basis set used, and may

Figure 4.1: The $g(x)$ functions for the B88 and double-Becke exchange functionals



change if larger bases are used. However, it does show that the present accuracy attainable by density functionals can be achieved without the expensive Fock term.

Chapter 5

Faster Integral Calculation

5.1 Introduction

One thing that all the calculations in this thesis have in common is their reliance on the calculation of a number of two-electron repulsion integrals (ERIs) (for example, equation (1.24)) over contracted Gaussian-type basis functions. As mentioned in section (1.5.7), the number of ERIs grows as $O(N^2)$ and has become the computational bottleneck for most HF and DFT calculations. Because of this, the efficient calculation of ERIs has been the focus of much research [56,159–176]. Even the new $O(N)$ approaches for construction of the Coulomb matrix require the computation of ERIs for short-range interactions.

It should be noted that almost all ERIs calculated at present are over Gaussian type orbitals. This is entirely due to the Gaussian product rule: that the product of two Gaussian functions is another Gaussian, centered somewhere on the line connecting the first two centers. Thus, any four-center ERI involving GTOs can immediately be reduced to a two-center integral (over GTOs), a simplification not available with STOs. Thus the fundamental (contractionless and momentumless) integral

$$I = \iint e^{-\alpha|\mathbf{r}_1-\mathbf{A}|^2} e^{-\beta|\mathbf{r}_1-\mathbf{B}|^2} \frac{1}{r_{12}} e^{-\gamma|\mathbf{r}_2-\mathbf{C}|^2} e^{-\delta|\mathbf{r}_2-\mathbf{D}|^2} \, d\mathbf{r}_1 d\mathbf{r}_2 \tag{5.1}$$

becomes

$$I = U \iint e^{-\zeta|\mathbf{r}_1-\mathbf{P}|^2} \frac{1}{r_{12}} e^{-\eta|\mathbf{r}_2-\mathbf{Q}|^2} \, d\mathbf{r}_1 d\mathbf{r}_2, \tag{5.2}$$

where

$$U = \exp\left[\frac{-\alpha\beta}{\alpha+\beta}|\mathbf{A}-\mathbf{B}|^2\right] \exp\left[\frac{-\gamma\delta}{\gamma+\delta}|\mathbf{C}-\mathbf{D}|^2\right] \tag{5.3}$$

and

$$\zeta = \alpha + \beta \qquad \eta = \gamma + \delta \tag{5.4}$$

$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\alpha + \beta} \qquad \mathbf{Q} = \frac{\gamma\mathbf{C} + \delta\mathbf{D}}{\gamma + \delta}. \tag{5.5}$$

How equation (5.2) is then computed depends on the integral method used. Each method has a number of subtle peculiarities, but all involve four basic steps (although the order of action varies from method to method).

The operator step, O, generates the momentumless two-center integrals over the two-electron operator. In the PRISM method [160] (which includes the McMurchie-Davidson [163], Obara-Saika [164], Head-Gordon-Pople [165], and Ten-no [166] as special cases) this forms the $[0]^{(m)}$ integrals.

In the momentum step L, recursive identities are used to build the momentumless quantities into those of the required angular momenta. This is where the bulk of integral research has been focused over the last two decades [159–166, 169–175].

The contraction step gathers together the primitive (uncontracted) components to form fully contracted contributions. The main advance of the PRISM method was the execution of the contraction step at the most optimal time. Previously, the contraction step has always been after the operator step (in the case of PRISM, forming the $(0)^{(m)}$s). The work presented in this chapter describes how contraction can be carried out first, thus providing the prospect of massive computational savings, and also forming the CO path to the COLD PRISM [176].

In the density step D, the contracted quantities are multiplied by the density matrix elements $P_{\mu\nu}$. This step is usually termed 'digestion' of the integrals and was traditionally performed last. One of the discoveries of the COLD PRISM, however, is that large time savings can be achieved by performing this step earlier.

5.2 Traditional $(0)^{(m)}$ Generation

Momentum is built up by recurrence relations which use $(L+1)K_{Bra}K_{Ket}$ (where K_{Bra} is the number of Gaussians on center **A** times the number at center **B**, and similarly for K_{Ket} with centers **C** and **D**) $[0]^{(m)}$s (one m for each degree of momentum, L).

The simplest $[0]^{(m)}$ (that is, $[0]^{(0)}$) is the fundamental electron-repulsion integral, equation (5.2). The integration is performed by first replacing each of the three factors by its Fourier

representation, giving (after some re-ordering)

$$[0]^{(0)} = (2\pi)^{-3} U \left(\frac{\pi^2}{\zeta\eta}\right)^{3/2} \iiint e^{-k_1^2/4\zeta - k_3^2/4\eta} e^{-i\mathbf{k}_1\cdot\mathbf{P} - i\mathbf{k}_3\cdot\mathbf{Q}} \frac{4\pi}{k_2^2} \\ (2\pi)^{-3} \int e^{i\mathbf{r}_1\cdot(\mathbf{k}_1+\mathbf{k}_2)}\, d\mathbf{r}_1\ (2\pi)^{-3} \int e^{i\mathbf{r}_2\cdot(\mathbf{k}_3-\mathbf{k}_2)}\, d\mathbf{r}_2\, d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3. \tag{5.6}$$

The last two integrals above are Fourier representations of the three-dimensional Dirac delta function, allowing simplification to

$$[0]^{(0)} = \frac{U}{8(\zeta\eta)^{3/2}} \iiint e^{-k_1^2/4\zeta - k_3^2/4\eta - i\mathbf{k}_1\cdot\mathbf{P} - i\mathbf{k}_3\cdot\mathbf{Q}} \frac{4\pi}{k_2^2} \delta(\mathbf{k}_1+\mathbf{k}_2)\delta(\mathbf{k}_3-\mathbf{k}_2)\, d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3. \tag{5.7}$$

Remembering that the Dirac δ function has the sampling property of

$$\int \delta(\mathbf{r}_1 - \mathbf{r}_2) h(\mathbf{r}_1)\, d\mathbf{r}_1 = h(\mathbf{r}_2) \tag{5.8}$$

for any function $h(\mathbf{r})$ allows the triple integral to collapse to the single integral

$$[0]^{(0)} = \frac{2\pi^2 U}{(\zeta\eta)^{3/2} R} \int_0^\infty \frac{\sin u}{u} e^{-u^2/(4\mathrm{T})}\, du \tag{5.9}$$

where, for convenience and computational efficiency, the following variables have been introduced:

$$\mathbf{R} = \mathbf{Q} - \mathbf{P} \tag{5.10}$$

$$\mathrm{T} = \upsilon^2 R^2 \tag{5.11}$$

$$\upsilon^2 = \frac{\zeta\eta}{\zeta+\eta}. \tag{5.12}$$

The final integration is usually performed with the introduction of a special function $G_m(\mathrm{T})$,

$$G_m(\mathrm{T}) = \sqrt{\frac{2}{\pi}} \int_0^1 t^{2m} e^{-\mathrm{T}t^2}\, dt \tag{5.13}$$

giving

$$[0]^{(0)} = U\sqrt{2\upsilon^2} G_0(\mathrm{T}). \tag{5.14}$$

The higher momentum $[0]^{(m)}$ formulae can be found via the relation

$$[0]^{(m+1)} = \frac{-d[0]^{(m)}/dR}{R} \tag{5.15}$$

producing the general formula

$$[0]^{(m)} = U\left(2\upsilon^2\right)^{m+1/2} G_m(\mathrm{T}). \tag{5.16}$$

If the value of T is less some critical value, the function $G_L(\mathrm{T})$ is calculated using Chebyshev interpolation [167]. The $G_m(\mathrm{T})$, $0 \leq m < L$ are then calculated by downward recursion using

$$G_m(\mathrm{T}) = \frac{1}{2m+1}\left[2\mathrm{T}G_{m+1}(\mathrm{T}) + e^{-\mathrm{T}}\right]. \tag{5.17}$$

Note that this also requires an interpolation for $e^{-\mathrm{T}}$.

If T is greater than a critical value, the distributions overlap negligibly, and $G_m(\mathrm{T})$ can be approximated

$$G_m(\mathrm{T}) \approx \frac{(2m-1)!!}{(2\mathrm{T})^{m+1/2}} \tag{5.18}$$

and the $[0]^{(m)}$ reduces to the classical multipole formula

$$[0]^{(m)} = \left(\frac{U}{R}\right)\left(\frac{1}{R^2}\right)\left(\frac{3}{R^2}\right)\cdots\left(\frac{2m-1}{R^2}\right), \tag{5.19}$$

which can be computed extremely fast via recursion.

The $(0)^{(m)}$s are formed if contraction is carried out immediately after forming the primitives. Contraction can occur early or late, depending on the PRISM path; however, it is advantageous to contract early if the total degree of contraction ($K_{tot} = K_{Bra}K_{Ket}$) is high. With the new algorithms mainly reducing the cost of high momentum, the point is fast being approached where ERI generation is dominated by making and contracting the $[0]^{(m)}$s. During contraction, scalings are introduced which are required for the later transformations, so a $(0)^{(m)}$ is more accurately represented as:

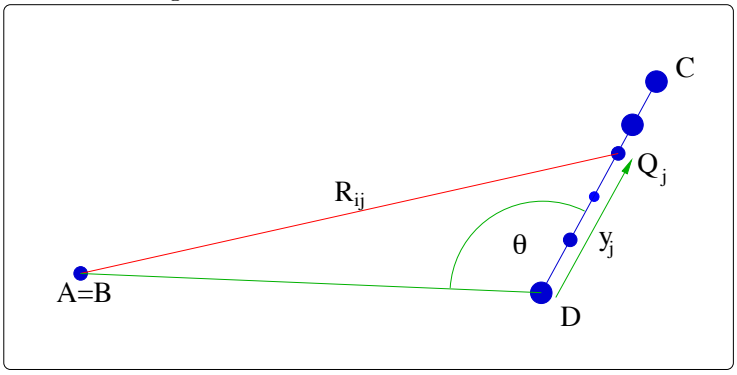
$${}_{a'b'p'}(0)^{(m)}_{c'd'q'} = \sum^{K_{Bra}}\sum^{K_{Ket}} \frac{(2\alpha)^{a'}(2\beta)^{b'}}{(2\alpha+2\beta)^{p'}}[0]^{(m)}\frac{(2\gamma)^{c'}(2\delta)^{d'}}{(2\gamma+2\delta)^{q'}}. \tag{5.20}$$

The double sum above is the problem. The amount of work required to calculate $(0)^{(m)}$ scales with $K_{Bra}K_{Ket}$. Note that this means two interpolations and a divide are calculated $K_{Bra}K_{Ket}$ times! The rest of this chapter presents a method which scales independently of K_{Bra} and K_{Ket}.

5.3 The Bra Concentric Case

For simplicity the special case of $\mathbf{A} = \mathbf{B}$ is considered first. As well as being conceptually simpler it also produces computational savings which are useful, so this special case has been coded into the Q-Chem program.

Figure 5.1: The Bra concentric case



The first (and most drastic) approximation is to assume that the Bra and Ket are well separated, something equivalent to assuming that all Gaussians on a shell-pair are point charges.

R_{ij} is generated from the length DB, which is the same for all elements of the contracted shell-pair and therefore only needs to be calculated once per $(0)^{(m)}$. Note that y_j is the fraction that Q_j is along DC. Computational time can be saved if the formula generates the reciprocal of R_{ij}, avoiding a costly divide. Now, by simple vector addition:

$$R_{ij} = DB - DCy_j \tag{5.21}$$

$$R_{ij}^2 = DB \cdot DB - 2DC \cdot DBy_j + DC \cdot DCy_j^2 \tag{5.22}$$

$$= |DB|^2 - 2|DC||DB| \cos\theta y_j + |DC|^2 y_j^2 \tag{5.23}$$

where

$$\cos\theta = \frac{DC \cdot DB}{|DC||DB|}, \tag{5.24}$$

therefore,

$$\frac{1}{R_{ij}} = \frac{1}{|DB|}\sqrt{1 - 2\frac{|DC|}{|DB|}y_j \cos\theta + \frac{|DC|^2}{|DB|^2}y_j^2}. \tag{5.25}$$

Using a binomial expansion and the Legendre polynomials P_k, the square root is replaced by an infinite series

$$\frac{1}{R_{ij}} = \frac{1}{|DB|}\sum_{k=0}^{\infty} P_k(\cos\theta)\left(\frac{|DC|}{|DB|}y_j\right)^k \tag{5.26}$$

$$= \frac{1}{|DB|}\sum_{k=0}^{\infty} a_{0k} y_j^k \tag{5.27}$$

with

$$a_{0k} = \left(\frac{|DC|}{|DB|}\right)^k P_k(\cos\theta). \tag{5.28}$$

The magnitude of the kth term is bounded by $\left(\frac{|DC|}{|DB|}\right)^k$ because $0 \leq y_j \leq 1$ and $-1 \leq P_k(\cos\theta) \leq 1$. Therefore the series will converge whenever the ratio of $|DC|$ to $|DB|$ is less than one. The coefficients a_{0k} can also be generated recursively (providing easier computation):

$$P_k(\cos\theta) = \left(\frac{2k-1}{k}\right)\cos\theta P_{k-1}(\cos\theta) - \left(\frac{k-1}{k}\right)P_{k-2}(\cos\theta) \tag{5.29}$$

$$\begin{aligned}\left(\frac{|DC|}{|DB|}\right)^k P_k(\cos\theta) = {} & \left(\frac{|DC|}{|DB|}\right)^k \left(\frac{2k-1}{k}\right)\frac{DC\cdot DB}{|DB||DB|}\left(\frac{|DB|}{|DC|}\right)P_{k-1}(\cos\theta) \\ & - \left(\frac{|DC|}{|DB|}\right)^k \left(\frac{k-1}{k}\right)P_{k-2}(\cos\theta)\end{aligned} \tag{5.30}$$

$$a_{0k} = \left(\frac{2k-1}{k}\right)\frac{DC\cdot DB}{|DB|^2}a_{0(k-1)} - \left(\frac{k-1}{k}\right)\frac{|DC|^2}{|DB|^2}a_{0(k-2)}. \tag{5.31}$$

The power of this group of formulae is that R_{ij}s can be calculated for the entire $(0)^{(m)}$ using only one measurement that involves both shell-pairs (the distance DB). The other information required (the distance DC and the fractional lengths y_j) is all within the ket shell-pair, and can therefore be constructed *before* pairing all shell-pairs, thus removing a large fraction of the $O(N^2)$ work, at the cost of a little extra $O(N)$ work. The general case is now presented.

5.4 The Non-Concentric Case

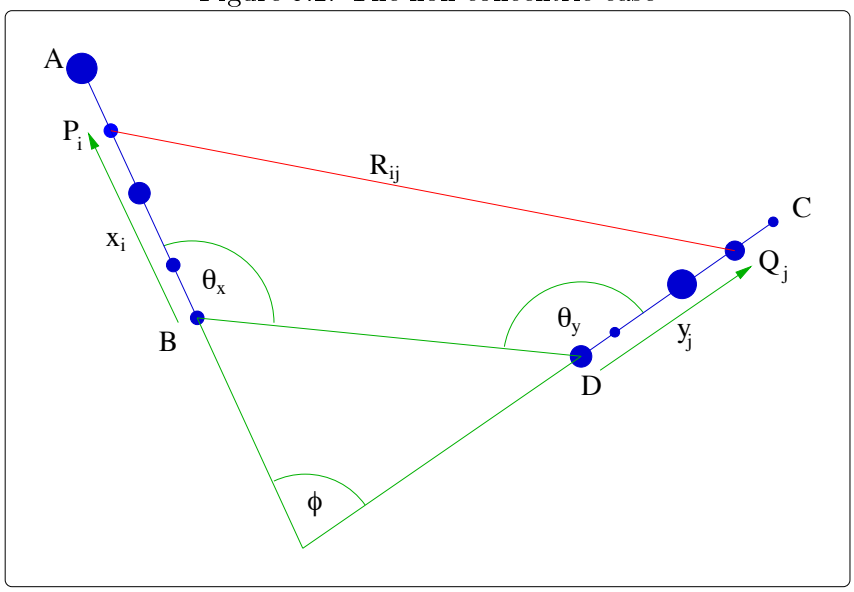
Proceeding in a similar manner to the concentric case, the well-separated approximation is invoked before developing a general formula for R_{ij}:

$$R_{ij} = ABx_i + DCy_j - DB \tag{5.32}$$

$$R_{ij}^2 = |DB|^2 - 2AB\cdot DBx_i - 2DC\cdot DBy_j + |AB|^2x_i^2 + 2AB\cdot DCx_iy_j + |DC|^2y_j^2 \tag{5.33}$$

$$\begin{aligned}= {} & |DB|^2 - 2|AB||DB|\cos\theta_x x_i - 2|DC||DB|\cos\theta_y y_j + |AB|^2x_i^2 \\ & + 2|AB||DC|\cos\phi x_iy_j + |DC|^2y_j^2;\end{aligned} \tag{5.34}$$

Figure 5.2: The non-concentric case



therefore, the reciprocal of R_{ij} is

$$\frac{1}{R_{ij}} = \frac{1}{|DB|}\left(1 - 2\frac{|AB|}{|DB|}x_i\cos\theta_x - 2\frac{|DC|}{|DB|}y_j\cos\theta_y + \frac{|AB|^2}{|DB|^2}x_i^2 + 2\frac{|AB||DC|}{|DB||DB|}x_i y_j\cos\phi + \frac{|DC|^2}{|DB|^2}y_j^2\right)^{-1/2} \tag{5.35}$$

$$= \frac{1}{|DB|}\left(1 - 2\alpha_i\cos\theta_x - 2\beta_j\cos\theta_y + \alpha_i^2 + 2\alpha_i\beta_j\cos\phi + \beta_j^2\right)^{-1/2} \tag{5.36}$$

where

$$\alpha_i = \frac{|AB|}{|DB|}x_i \qquad \beta_j = \frac{|DC|}{|DB|}y_j \tag{5.37}$$

which leads to:

$$\frac{1}{R_{ij}} = \frac{1}{|DB|}\sum_{k=0}^{\infty}\sum_{k=0}^{\infty}d_{kl}\alpha_i^k\beta_j^l \tag{5.38}$$

$$= \frac{1}{|DB|}\sum_{k=0}^{\infty}\sum_{k=0}^{\infty}a_{kl}x_i^k y_j^l \tag{5.39}$$

with

$$a_{kl} = \left(\frac{|AB|}{|DB|}\right)^k\left(\frac{|DC|}{|DB|}\right)^l d_{kl}. \tag{5.40}$$

Like the three-center case, the d_{kl} are recursively related, but now there are two recurrence

relations, one for increasing k, another for l:

$$d_{kl} = \left(\frac{2k+2l-1}{k}\right)\cos\theta_x d_{(k-1)l} - \left(\frac{k+2l-1}{k}\right)d_{(k-2)l} - \left(\frac{2l-1}{k}\right)\cos\phi d_{(k-1)(l-1)} + d_{k(l-2)} \tag{5.41}$$

$$d_{kl} = \left(\frac{2l+2k-1}{l}\right)\cos\theta_y d_{k(l-1)} - \left(\frac{l+2k-1}{l}\right)d_{k(l-2)} - \left(\frac{2k-1}{l}\right)\cos\phi d_{(k-1)(l-1)} + d_{(k-2)l}. \tag{5.42}$$

The d_{kl} relations are easily converted to relations for the more convenient variable a_{kl}. For example, equation (5.41) can become

$$\begin{aligned}\left(\frac{|AB|}{|DB|}\right)^k\left(\frac{|DC|}{|DB|}\right)^l d_{kl} = {} & \left(\frac{2k+2l-1}{k}\right)\left(\frac{|AB|}{|DB|}\right)^{k-1}\left(\frac{|DC|}{|DB|}\right)^l\frac{AB\cdot DB}{|DB|^2}d_{(k-1)l} \\ & - \left(\frac{k+2l-1}{k}\right)\frac{|AB|^2}{|DB|^2}\left(\frac{|AB|}{|DB|}\right)^{k-2}\left(\frac{|DC|}{|DB|}\right)^l d_{(k-2)l} - \left(\frac{2l-1}{k}\right)\frac{AB\cdot DC}{|DB|^2} \\ & \left(\frac{|AB|}{|DB|}\right)^{k-1}\left(\frac{|DC|}{|DB|}\right)^{l-1} d_{(k-1)(l-1)} + \frac{|DC|^2}{|DB|^2}\left(\frac{|AB|}{|DB|}\right)^k\left(\frac{|DC|}{|DB|}\right)^{l-2} d_{k(l-2)}\end{aligned} \tag{5.43}$$

which leads to the recurrence relations

$$\begin{aligned}a_{kl} = {} & \left(\frac{2k+2l-1}{k}\right)\frac{AB\cdot DB}{|DB|^2}a_{(k-1)l} - \left(\frac{k+2l-1}{k}\right)\frac{|AB|^2}{|DB|^2}a_{(k-2)l} \\ & - \left(\frac{2l-1}{k}\right)\frac{AB\cdot DC}{|DB|^2}a_{(k-1)(l-1)} + \frac{|DC|^2}{|DB|^2}a_{k(l-2)}\end{aligned} \tag{5.44}$$

$$\begin{aligned}a_{kl} = {} & \left(\frac{2l+2k-1}{l}\right)\frac{DC\cdot DB}{|DB|^2}a_{k(l-1)} - \left(\frac{l+2k-1}{l}\right)\frac{|DC|^2}{|DB|^2}a_{k(l-2)} \\ & - \left(\frac{2k-1}{l}\right)\frac{AB\cdot DC}{|DB|^2}a_{(k-1)(l-1)} + \frac{|AB|^2}{|DB|^2}a_{(k-2)l}.\end{aligned} \tag{5.45}$$

This general formula is more expensive than the concentric case, but it uses no extra shell-pair information. Also (as a nice check) if $AB = 0$ the relation reduces to the concentric case. The above recurrence relations can also be extended to deal with higher powers of R_{ij} (which are needed for higher momentum $(0)^{(m)}$s). The argument below shows that only the coefficients are changed:

$$\left(1 - 2\alpha_i\cos\theta_x - 2\beta_j\cos\theta_y + \alpha_i^2 + 2\alpha_i\beta_j\cos\phi + \beta_j^2\right)^{-1-m/2} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} d_{kl}^{(m)}\alpha_i^k\beta_j^l \tag{5.46}$$

$$d_{kl}^{(m)} = \left(\frac{2k+2l+m}{k}\right)\cos\theta_x d_{(k-1)l}^{(m)} - \left(\frac{k+2l+m}{k}\right) d_{(k-2)l}^{(m)} - \left(\frac{2l+m}{k}\right)\cos\phi d_{(k-1)(l-1)}^{(m)} + d_{k(l-2)}^{(m)} \tag{5.47}$$

$$d_{kl}^{(m)} = \left(\frac{2l+2k+m}{l}\right)\cos\theta_y d_{k(l-1)}^{(m)} - \left(\frac{l+2k+m}{l}\right) d_{k(l-2)}^{(m)} - \left(\frac{2k+m}{l}\right)\cos\phi d_{(k-1)(l-1)}^{(m)} + d_{(k-2)l}^{(m)}. \tag{5.48}$$

For example, the first relation in the R_{ij}^{-3} expansion ($m = 1$) is

$$a_{kl} = \left(\frac{2k+2l+1}{k}\right)\frac{AB\cdot DB}{|DB|^2}a_{(k-1)l} - \left(\frac{k+2l+1}{k}\right)\frac{|AB|^2}{|DB|^2}a_{(k-2)l} - \left(\frac{2l+1}{k}\right)\frac{AB\cdot DC}{|DB|^2}a_{(k-1)(l-1)} + \frac{|DC|^2}{|DB|^2}a_{k(l-2)}. \tag{5.49}$$

5.5 Convergence Test

The resulting integral will only be accurate if the binomial expansion converges fast enough. For the concentric case we have a rigorous error bound for the kth term. The error introduced is of the order of the kth-plus-one term, where k is the number of terms in the expansion. Increasing k will obviously increase the accuracy of an integral, allowing for more of the shell-pair interactions to be calculated via this new method (provided that the two shell-pairs are far enough apart that the point-charge approximation still holds). But this will also increase the cost of a calculation. Using all terms up to, and including, ninth order has been found to be a good compromise. This will give an error of

$$\frac{|DC|}{|DB|} = \sqrt[10]{\text{ErrorBound}}. \tag{5.50}$$

For the non-concentric case, the rigorous error bound is too complex to use in a computation, so an error bound is estimated by

$$\text{ErrorBound} = \left(\frac{|AB|+|CD|}{|DB|}\right)^{10}. \tag{5.51}$$

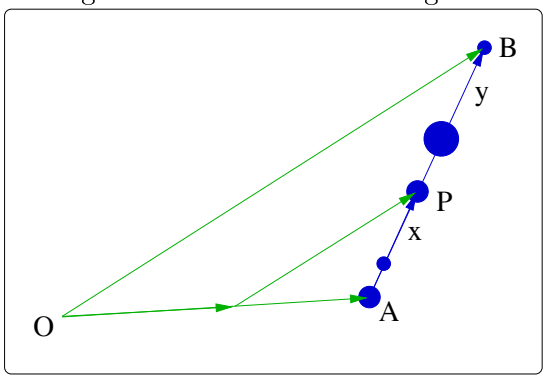
Even if the shell-pair passes the above error-bound, $|DB|$ must still be large enough for the point charge approximation to be valid. This is solved with a final restriction, that $|DB|$ must be greater than some distance, chosen with the most diffuse functions and the desired accuracy in mind.

5.6 Constructing Moments

It is now clear that $(0)^{(m)}$s can be calculated without the expensive $K_{Bra}K_{Ket}$ scaling. 'Moments' along the shell-pairs (the amplitude of the Gaussian multiplied by the fraction along AB raised to each power up to the length of the expansion) need to be constructed. These moments can then be paired together, along with appropriate coefficients for the power of R_{ij} to generate the $(0)^{(0)}$. For higher momentum the moments need to be scaled by certain exponents first.

The fraction along AB can be generated from the shell-pair data which store the exponents of the constituent Gaussians (α and β).

Figure 5.3: The fraction along AB



$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\alpha + \beta} \tag{5.52}$$

$$= \mathbf{B} + y\mathbf{BA} \tag{5.53}$$

$$y\mathbf{BA}(\alpha + \beta) = \alpha(\mathbf{A} - \mathbf{B}) \tag{5.54}$$

$$y = \frac{\alpha}{\alpha + \beta}. \tag{5.55}$$

Note that the fraction from **B** is calculated (not **A**). Allowing the moments to be calculated via

$$u_i = \sum_{p=1}^{K_{Bra}} \frac{(2\alpha_p)^{a'+i}(2\beta_p)^{b'}}{(2(\alpha_p + \beta_p))^{p'+i}} \exp\left[\frac{-\alpha\beta}{\alpha + \beta}|\mathbf{A} - \mathbf{B}|^2\right] \tag{5.56}$$

$$v_j = \sum_{q=1}^{K_{ket}} \frac{(2\gamma_q)^{c'+j}(2\delta q)^{d'}}{(2(\gamma q + \delta q))^{q'+j}} \exp\left[\frac{-\gamma\delta}{\gamma + \delta}|\mathbf{C} - \mathbf{D}|^2\right] \tag{5.57}$$

where i and j vary from zero to nine. The implementation in Q-Chem generates moments only for the present class of integrals (those of the same momentum and contraction). This saves memory as not all moments need to be stored and only those scalings required for the

immediate class need to be included. The one down-side is that two sets of moments have to be calculated.

5.7 Matrix Equations

The two sets of moments above form two vectors (each of length ten) for each $(0)^{(m)}$. With these, and a matrix **H**, whose elements are the coefficients from the earlier expression for R_{ij}, we can form a much simpler matrix form for the $(0)^{(m)}$s:

$$_{a'b'p'}(0)^{(m)}_{c'd'q'} = \frac{(2m-1)!!}{|DB|^{2m+1}}\mathbf{u}^t\mathbf{H}\mathbf{v}. \tag{5.58}$$

Further efficiency is gained by forming **Hv** (which depends only on the ket and m) and then contracting the product with the necessary **u**. The advantage of a matrix form is that the $(0)^{(m)}$s can now be generated using BLAS-2 and BLAS-3 constructs, which are linear algebra packages highly optimized for a variety of computer platforms.

5.8 General Algorithm

It is helpful to present the entire final algorithm. The six steps involved are:

1. *Construct moments.* Loop over all the shell-pairs, generating moments up to and including ninth order, with a separate set for each type of scaling required for that type of shell-pair. Note that this part of the algorithm is only $O(N)$ in cost.

2. *Convergence test.* Loop over all kets for the present bra, finding for which shell-pairs the expansion will converge, and construct lists of those that pass and fail. This section (and those that follow) is $O(N^2)$.

3. *Find geometric parameters.* Calculate the values of $\frac{AB\cdot DB}{DB\cdot DB}$, etc.

4. *Form* **H** from the recurrence relation. This is performed once per m value of the class. There is a significant time saving if $|AB|$=0.

5. *Form* **Hv**. Required once per unique **Hv** product. Again there is a significant time saving if $|AB|$=0.

6. *Generate all* $(0)^{(m)}s$, until a new **Hv** is required.

5.9 Theoretical Performance Analysis

A common way to measure the performance of integral techniques is to count the number of floating point operations (flops) required. Table 5.1 lists the flop count for each part of the new algorithm. Only those flops in the $O(N^2)$ parts of the code are counted; for any large molecule the rest are insignificant. It is also assumed that $\mathbf{A} \neq \mathbf{B}$ and that the bra and ket are far apart (again this is true for the majority of shell-quartets in any large system).

Table 5.1: Flop costs for each part of the $(0)^{(m)}$ construction

Task	When Called	Cost	$n = 10$ cost
Geometric constants	Once	33	33
Form $\mathbf{H}$	Once per m value	$5n^2 - 9n - 4 + 2m$	$406 + 2m$
Generate $\mathbf{Hv}$	For each unique $\mathbf{Hv}$	n^2	100
Generate $\mathbf{u}^t\mathbf{Hv}$	For each unique $(0)^{(m)}$	$2n + 2$	22
Copy duplicate $(0)^{(m)}$s	For each duplicate	1	1

Note that most of the routines above have a cost depending on the length of the expansion, n. Decreasing this length will mean fewer shell-quartets pass the convergence tests, forcing more of the work to be done the traditional way. Thus there is a trade-off here and $n = 10$ has proven to be a good compromise.

Using the above table, and how often each routine is called, a cost per class can be calculated. This cost is usually described in terms of the primitive, half-contracted and contracted contributions,

$$\text{Cost} = xK^4 + yK^2 + z, \tag{5.59}$$

where K is the degree of contraction of each of the four shells. Table 5.2 compares the contraction-first method (CO) described above with the fastest operator-first method (OC) of the COLD PRISM for the low momentum integrals.

In addition to the numbers presented in Table 5.2, the OC calculations require 1 square root and 2 divides as x-type work and CO needs 1 square root and 1 divide as z-type work. By making reasonable assumptions for the flop cost of these functions, the point at which CO becomes faster than OC can be estimated. These cross-over points are listed in the final column. The momentumless class shows the lowest CO/OC crossover. Here CO is faster whenever $K \geq 2$, making the new path extremely useful. The CO costs do grow rather fast

Table 5.2: Comparison of CO and OC flop counts

	CO			OC			
Class	x	y	z	x	y	z	K_{tot} CO win
$(ss\|ss)$	0	0	561	8	0	0	12
$(ps\|ss)$	0	0	1235	14	8	0	21
$(ds\|ss)$	0	0	2121	22	24	0	31
$(pp\|ss)$	0	0	2189	22	30	0	31
$(ps\|ps)$	0	0	2368	26	28	0	32
$(dp\|ss)$	0	0	3514	32	76	0	40
$(ds\|ps)$	0	0	4019	42	76	0	41
$(pp\|ps)$	0	0	4228	42	94	0	42
$(dd\|ss)$	0	0	5554	44	176	0	48
$(ds\|ds)$	0	0	7729	82	198	0	50
$(dp\|ps)$	0	0	7035	62	224	0	50
$(ds\|pp)$	0	0	8245	82	242	0	51
$(pp\|pp)$	0	0	9493	94	296	0	52
$(dp\|dp)$	0	0	32578	254	1482	0	63
$(dd\|dd)$	0	0	109201	586	5962	0	67

with momentum though, indicating that it will only rarely be required for classes with a total momentum greater than 4. An exception may be calculations with transition metals and heavy main-group elements, where highly contracted d functions are used.

While flop-counts are a good guide to the performance of an algorithm, they are not the whole story. The number of memory operations, for example, is also an important consideration. Use of the computer's architecture is also relevant. For these reasons it is always important to perform a timings analysis.

5.10 Empirical Performance Analysis

The CO path is designed for large systems. Therefore the test 'molecule' is C_{340} arranged as a 4x85 rectangular lattice in which adjacent atoms are 1.25 Å apart. The basis set used is 6-31G*, which gives 5100 basis functions. The timings were carried out on an IBM model 43P Power PC workstation.

Table 5.3 presents the timings data for a selection of typical classes that arise. The actual timings are of little interest; instead, the ratios between the CO and OC times are presented.

Table 5.3: Timing ratios for the CO and OC Paths

Class	(K_{Bra}, K_{Ket})	OC/CO Ratio
$(ss\|ss)$	(12,12)	2.83
$(ss\|ss)$	(6,6)	1.52
$(ps\|ss)$	(8,12)	2.48
$(ps\|ss)$	(8,6)	1.64
$(pp\|ss)$	(8,12)	2.26
$(pp\|ss)$	(8,6)	1.44
$(ds\|ss)$	(6,12)	2.07
$(ds\|ss)$	(6,6)	1.38
$(dp\|ss)$	(3,12)	1.07
$(dp\|ss)$	(3,6)	1.00
$(dd\|ss)$	(1,12)	1.00
$(dd\|ss)$	(1,6)	0.98
$(ps\|ps)$	(8,8)	2.03
$(pp\|ps)$	(8,8)	1.76
$(pp\|pp)$	(8,8)	1.76
$(ds\|ps)$	(6,8)	1.85
$(ds\|pp)$	(6,8)	1.08
$(dp\|ps)$	(3,8)	0.96
$(ds\|ds)$	(6,6)	1.01

The timings confirm most of the trends seen in the flop-counts. CO is obviously faster for integrals of high contraction and low momentum. For $(ss|ss)$ with $K_{Bra} = K_{Ket} = 12$ it is nearly three times faster. However, as the momentum increases and the contraction decreases, the CO advantage diminishes, until, by $(dd|ss)$ with $K_{Bra} = 1$ and $K_{Ket} = 6$, CO is slightly slower. The timings overall show a slight improvement over that predicted by flop-counts.

5.11 Conclusion

This chapter presents a new path for the construction of $(0)^{(m)}$s. By performing contraction earlier than was previously possible, large savings in time can be made for highly contracted integrals. This has the effect of moving previously $O(N^2)$ work dependent on K into an $O(N)$ part of the code. The algorithm also uses the BLAS matrix algebra, adding to the efficiency of the implementation. The CO path has been implemented in the Q-Chem program and shows a three-fold speed up for only moderately contracted $(ss|ss)$ integrals. However, the new algorithm shows a steeper increase in cost with momentum than the more traditional methods, and is seldom faster when the total momentum is greater than four.

Chapter 6

The CASE Approximation

"The electrons on the nose of Professor Karplus do not interact with the electrons on the nose of Professor Eyring" - Enrico Clementi [177]

6.1 Introduction to Linear Methods

As mentioned in the previous chapter, the calculation of the ERIs is the bottleneck for modern day SCF calculations. The central problem is that, no matter how fast these interactions are computed, there are still $O(N^2)$ Coulomb and Exchange interactions, which grows just too fast for large molecules to be studied. It was shown in Chapter 2 that the DFT exchange (and correlation) terms can be computed in $O(N)$ work. Therefore $O(N)$ DFT calculations are possible if the Coulomb problem (to determine the Coulomb energy of a system of N localized distributions of charge in $O(N)$ work) can be solved.

6.1.1 The Fast Multipole Method

The breakthrough was made in 1985 when Greengard and Rokhlin [178–180] showed how to compute the Coulomb energy of point charges in only linear work. Greengard's Fast Multipole Method (FMM) belongs to the family of algorithms called tree codes. Tree codes [181] acquire their speed by transforming the information about a cluster of charge into a simpler representation which is used to compute the influence of the cluster on objects at large distances.

The Fast Multipole Method begins by scaling all particles into a box with coordinate ranges $[0\ldots1, 0\ldots1, 0\ldots1]$ to ensure numerical stability of subsequent operations. The parent box is then divided in half along each Cartesian axis. Each child box is then further

subdivided, forming a 'computational family tree'. The number of subdivisions is chosen so that the number of particles in the lowest level boxes is approximately independent of the number of particles (this is required to achieve linear scaling).

Each particle is then placed within a box on the lowest level of the tree. Any empty boxes are removed to allow efficiency for inhomogeneous systems. The charges of particles within each lowest level box are then expanded in multipoles about the center of its box. The multipole expansion of each lowest level box is then translated to the center of its parent box via one of three special FMM operators.

The second of the FMMs operators is used for each box to transform the multipole expansions of all well-separated boxes (those that are not nearest neighbours) into Taylor expansions about the center of the current box. However, only those multipole expansions from boxes which are well-separated at the present level and not well-separated at the parent level are interacted. The multipole expansions are also translated to Taylor expansions in the parent box. This allows each transformation to be performed as high up the tree as possible. In practice this pass is the bottleneck of the algorithm, yet, like all the other passes, it is $O(N)$.

The third pass transforms the parent Taylor expansions down the tree to the child boxes, so that each low-level box contains the Taylor expansion representing all well-separated particles. The fourth pass calculates the far-field potential for each particle via the Taylor representation in the particles box. A final pass calculates the interactions between particles that are not well-separated at any level in the tree.

6.1.2 The Continuous Fast Multipole Method

The FMM has subsequently been applied to problems in astrophysics, plasma physics, molecular dynamics, fluid dynamics, partial differential equations and numerical complex analysis. The FMM was generalized to handle continuous distributions (forming the Continuous FMM, CFMM) in 1994 by White *et al.* [157, 182] after making several improvements to the original FMM [180, 183, 184]. The main change was the introduction of a well-separated index, describing the distance required before interactions can be calculated via multipoles. This index depends on the diffuseness of the charge distributions involved. Over the last few years CFMM has become a very mature algorithm, and the Q-Chem implementation is now a highly optimized, efficient code.

6.1.3 ONX: Order N Exchange

The first method to be able to calculate the Hartree-Fock exchange energy in only $O(N)$ work was developed by Schwegler and Challacombe in 1996 [185]. It can be shown that the one particle density matrix $\rho_1(\mathbf{r}_1, \mathbf{r}_2)$ decays exponentially with $|\mathbf{r}_1 - \mathbf{r}_2|$ for one-dimensional insulators [186]. A similar behaviour for three-dimensional insulators has also been speculated. Kohn [187] conjectured that, for disordered molecular systems, the density matrix behaves as

$$\rho(\mathbf{r}_1, \mathbf{r}_2) \sim e^{-|\mathbf{r}_1 - \mathbf{r}_2|/l}, \qquad (|\mathbf{r}_1 - \mathbf{r}_2| \to \infty) \tag{6.1}$$

where l is a system-specific parameter. This fast decay should allow the long-range exchange matrix elements to be screened out.

This screening had not been possible previously as the off-diagonal elements of the density matrix were shown to decay only slowly. Schwegler and Challacombe have pointed out that this slow decay is simply an artifact of the incomplete basis set. Thus, it was thought reasonable to assume

$$K_{ab} = \iint \frac{\rho_{ab}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} e^{-|\mathbf{r}_1 - \mathbf{r}_2|/l} \, d\mathbf{r}_1 d\mathbf{r}_2 \tag{6.2}$$

which, with the introduction of some thresholding criteria, leads to an $O(N)$ algorithm for computing the exchange matrix.

The ONX method is not without its problems however. Most importantly, it is only linear for insulators. For non-insulators the computational time becomes the traditional $O(N^2)$. The algorithm is also extremely expensive compared to the CFMM (which is calculating the Coulomb matrix). This shows just how much harder an $O(N)$ exchange algorithm is compared to the Coulomb problem.

6.1.4 The KWIK Algorithm

While CFMM is a linear method, the authors consider it an expensive $O(N)$ technique and believe that there will someday be a faster method. In 1996 Dombroski *et al.* [188] introduced the KWIK algorithm, which may become an alternative to the CFMM.

The KWIK algorithm begins by examining the Coulomb operator, which contains a singularity at the origin and exhibits very slowly decaying long range behaviour. It was thought that separating these two characteristics would allow easier approximation. Thus

the separation into short- and long-range operators,

$$\frac{1}{r} \equiv \frac{f(r)}{r} + \frac{1-f(r)}{r} \tag{6.3}$$

was introduced. The short-range operator should have only $O(N)$ significant interactions. The long-range operator is transformed into Fourier Space where it decays rapidly, and interactions are calculated in $O(N)$ work via summation over Fourier coefficients and particles.

A separator function $f(r)$, which produces a rapidly decaying short-range function and a long-range function whose Fourier Transform decays rapidly (so that the summation required converges quickly) is desired. Dombroski tried several forms for the separator function [189] including $\exp(-\omega r)$, $\exp(-\omega r^2)$, $\tanh(\omega r)$ and $\mathrm{erfc}(\omega r)$. The decay parameter ω was introduced so that the amount of short- to long-range work could be tuned. Of these functions only $\mathrm{erfc}(\omega r)$ decays Gaussianly and has a Fourier transform which decays Gaussianly.

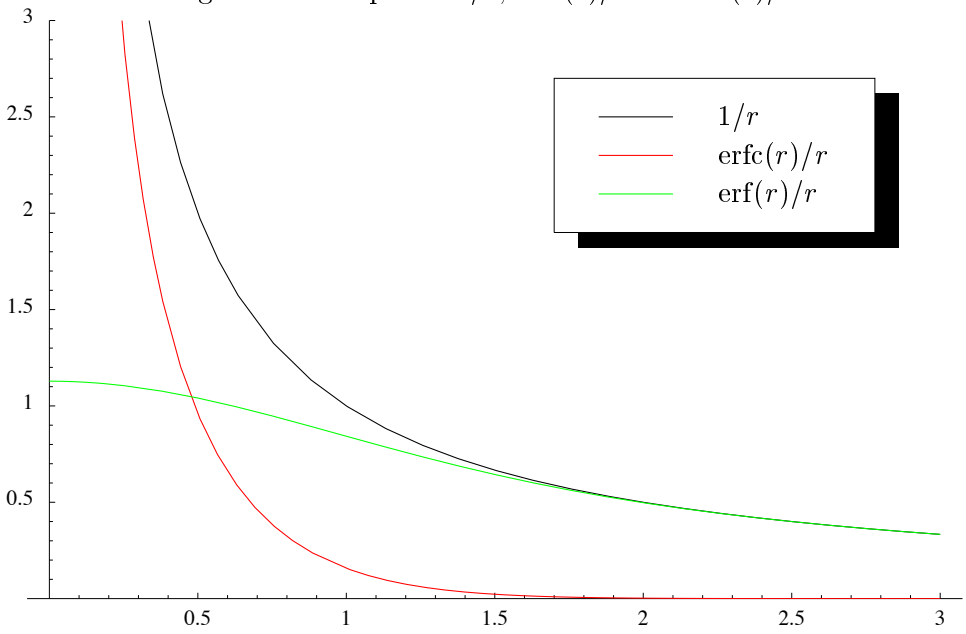
In an attempt to find a more appropriate separator function, Lee *et al.* [190] defined the ultimate separator as the one which minimizes the decay in real and Fourier space. The solution is a rather complex function, expressed as either modified Bessel, Hermite or parabolic cylinder functions. This function decays asymptotically as $r^{-1/2}\exp(-r^2/2)$, slightly faster than the Gaussian decay of $\mathrm{erfc}(\omega r)$. For practical reasons however, the $\mathrm{erfc}(\omega r)$ separator is preferred over the Lee separator as the decay speed difference between the two is small and the error function is far simpler to implement and optimize [191].

The KWIK algorithm has shown promising behaviour for point charges, yet the extension to continuous charges is still under development. Problems exist in calculating the long range energy accurately; however, it should be stressed that KWIK is still a young algorithm, requiring much more research.

6.2 The CASE Approximation

Examining the KWIK $\mathrm{erfc}(\omega r)/r$ separator in figure 6.1 shows two distinctly contrasting curves. The first is singular at the origin and decays very rapidly. The second, accounting for the long-range behaviour of $1/r$, is exceptionally smooth (that is, lacking in high-frequency components) and finite at the origin. If the long range operator was a flat line it would have no effect on the wavefunction, only altering the total energy. Given the flatness of the long-range operator it is reasonable to assume that it is of less importance to the wavefunction than the short-range function.

Figure 6.1: Graphs of $1/r$, $\mathrm{erfc}(r)/r$ and $\mathrm{erf}(r)/r$



The CASE (Coulomb-Attenuated-Schrödinger-Equation) approximation is to take this idea to its logical conclusion and completely neglect the long-range component. This seems a drastic step to take, but it is not as preposterous as it may first appear. At worst it could be viewed as a test for the short-sightedness of electrons [192]. The CASE results may be strictly only applicable to a universe where the Coulomb operator decays rapidly, yet it is possible that they could be a useful guide to chemistry. It is well-known to chemists that molecules are essentially non-polar over large distance scales. Therefore it is not unreasonable to expect the attractive (nuclear-electron) and repulsive (nuclear-nuclear and electron-electron) Coulomb interactions between widely separated regions of a molecule to approximately cancel. What is required, therefore, is a way of smoothly cutting off the long-range interaction — something that the $\mathrm{erfc}(\omega r)/r$ function provides.

If the approximation proves to be a disaster all is not lost, as CASE can be employed as a zeroth-order approximation upon which higher-order corrections can be constructed. The value of CASE in its own right should be examined first, however.

6.3 The Hydrogen Atom

Whenever introducing a new approximation it is often useful to examine its performance on the simplest model systems. Here the effect of CASE is investigated on the H atom, the

attenuated Schrödinger equation for which is

$$\frac{1}{2}\frac{d^2\Psi}{dr^2} + \frac{1}{r}\frac{d\Psi}{dr} + \frac{\mathrm{erfc}(\omega r)}{r}\Psi + E\Psi = 0. \tag{6.4}$$

With $\omega = 0$ CASE becomes the traditional Schrödinger equation giving $E_0 = -1/2$ hartrees and $\Psi_0(r) = \sqrt{\pi}^{-1/2}\exp(-r)$. As ω is increased the attractive nuclear-electron contribution to the energy rapidly decreases and the total energy rises. The effect on the wavefunction, however, is far less intuitive. A quantitative investigation of Ψ is possible through first-order perturbation theory.

The CASE energy and wavefunction can be expanded as

$$E = E_0 + E_1 + E_2 + \cdots \tag{6.5}$$

$$\Psi = \Psi_0 + \Psi^1 + \Psi^2 + \cdots \tag{6.6}$$

$$\Psi^1 = c_2\Psi_0^2 + c_3\Psi_0^3 + \cdots \tag{6.7}$$

where Ψ_0^2, Ψ_0^3, $\cdots$ are the $\omega = 0$ excited-state wavefunctions. This allows the first-order correction to be expressed as the expectation value of the long-range operator $\hat{L}$ (the 'background')

$$E_1 = \int \Psi_0\hat{L}\Psi_0 \tag{6.8}$$

$$= \int_0^\infty 4r^2\exp(-r)\frac{\mathrm{erf}(\omega r)}{r}\exp(-r)\,dr \tag{6.9}$$

$$= \frac{2}{\omega\sqrt{\pi}} + (1 - 2\omega^{-2})\exp(\omega^{-2})\,\mathrm{erfc}(\omega^{-1}) \tag{6.10}$$

and the coefficients of the first-order correction to the wavefunction for any ω are given by the $\omega = 0$ components

$$c_k = \frac{\int \Psi_0\hat{L}\Psi_0^k}{E_0 - E_0^k}. \tag{6.11}$$

The first-order energy $E_0 + E_1$, along with the exact energy, which has been obtained using the mathematics package Mathematica [193] by solving equation (6.4) numerically, are listed in Table 6.1 for a selection of ω values. The table shows that the energy does rise with increasing ω. What is more interesting, however, is the surprisingly close agreement between the exact and first-order energies. This suggests that, at least for small ω values, Ψ_0 is an excellent approximation to Ψ, or conversely, that attenuation has only a small effect on the wavefunction Ψ. This is supported by the first two coefficients c_2 and c_3, presented in Table 6.1, being exceedingly small, indicating that the first-order correction to the wavefunction is negligible.

Table 6.1: CASE First-order perturbation theory on the H atom

ω	E	$E_0 + E_1$	c_2	c_3
0	-0.500000	-0.500000	0	0
0.001	-0.498872	-0.498872	$3\text{x}10^{-9}$	$1\text{x}10^{-9}$
0.01	-0.488717	-0.488717	$3\text{x}10^{-6}$	$1\text{x}10^{-6}$
0.1	-0.388270	-0.388266	$3\text{x}10^{-3}$	$1\text{x}10^{-3}$
0.2	-0.282838	-0.282631	$2\text{x}10^{-2}$	$6\text{x}10^{-3}$
0.5	-0.051071	-0.031011	$1\text{x}10^{-1}$	$5\text{x}10^{-2}$

This simple example above does show that CASE chemistry is not the disaster that might have been expected. A good wavefunction is produced, thanks to the smoothness of the background. The large magnitude of the background does, however, produce large effects on the total energy under attenuation. This is in contrast to variational calculations which usually produce a good energy from a poor wavefunction. CASE is clearly not that useful if total energies are the goal of a calculation, yet if relative energies are the goal, it is conceivable that there may be a systematic cancellation of error and, by virtue of an accurate wavefunction, relative energies could be largely unaffected by the neglect of the background. This is investigated in the next section.

6.4 Molecular Results

The necessary integrals to perform CASE calculations have been coded into the PRISM algorithm [176] of the Q-Chem [125] program (see Chapter 7). In light on the results of the previous section an attenuation parameter of $\omega = 0.1\, a_0^{-1}$ is chosen.

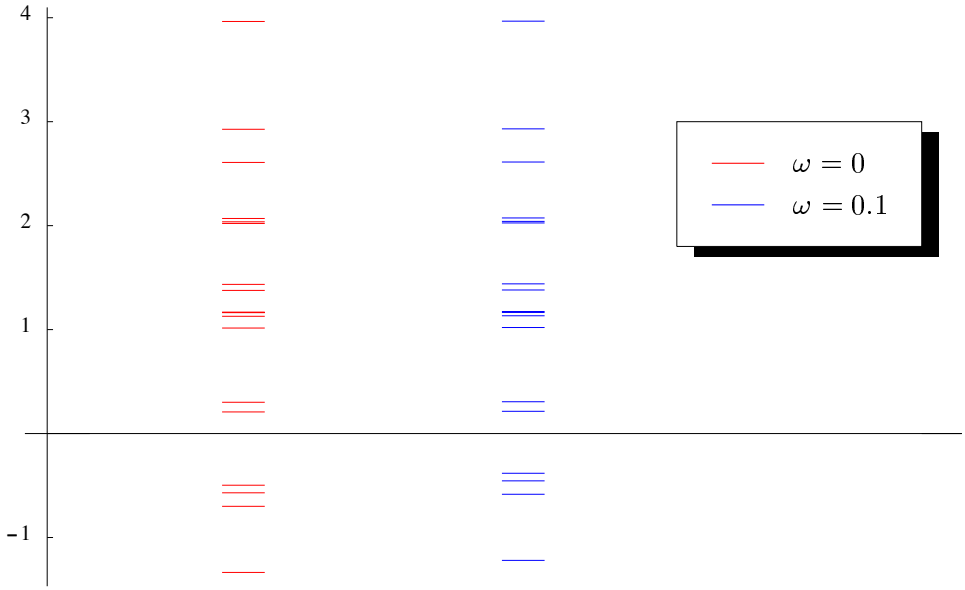
To begin, the effect of CASE on the molecular orbitals of H_2O is examined. The MP2/6-31G* geometry is used. The HF/6-31G* calculations using $\omega = 0$ and $\omega = 0.1$ give total energies of -76.00981 and -71.72696 respectively. Thus, as expected, there is a huge effect on the total energy. The MO energies are presented in Figure 6.2. This figure reveals the interesting effect that Coulomb attenuation raises the occupied MO energies, but has very little effect on the unoccupied MOs.

The MO coefficients of the highest-occupied (HOMO) and lowest-unoccupied (LUMO) molecular orbitals and their energies are listed in Table 6.2. These MOs typify the changes found for all the molecular orbitals on introducing attenuation. It is clear that the wavefunc-

Table 6.2: MO coefficients and energies for the HOMO and LUMO of H_2O using HF/6-31G*

	HOMO (B_1)		LUMO (A_1)	
	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$
O (1s)	0	0	0.10002	0.09992
O (2s)	0	0	-0.05859	-0.05873
O ($2p_x$)	0	0	0	0
O ($2p_y$)	0.63998	0.63962	0	0
O ($2p_z$)	0	0	-0.22243	-0.22247
O (3s)	0	0	-1.38818	-1.38818
O ($3p_x$)	0	0	0	0
O ($3p_y$)	0.51110	0.51148	0	0
O ($3p_z$)	0	0	-0.51217	-0.51233
O ($3d_{xx}$)	0	0	0.05532	0.05519
O ($3d_{xy}$)	0	0	0	0
O ($3d_{yy}$)	0	0	0.07027	0.07016
O ($3d_{xz}$)	0	0	0	0
O ($3d_{yz}$)	0.03347	0.03334	0	0
O ($3d_{zz}$)	0	0	0.04191	0.04180
H_1 (1s)	0	0	0.05745	0.05744
H_1 (2s)	0	0	1.03009	1.02992
H_2 (1s)	0	0	0.05745	0.05744
H_2 (2s)	0	0	1.03009	1.02992
MO energy	-0.49736	-0.38174	+0.20821	+0.21353

Figure 6.2: Molecular orbital energies (a.u.) for H_2O using CASE HF/6-31G* (core MO not shown)



tion is little effected by a small attenuation.

Given the very large changes to the total energies for the hydrogen atom and the water molecule, it might seem unlikely that bond dissociation energies could be reproduced well. To test this, the UHF/6-31G** potential curve of H_2 was scanned with and without Coulomb attenuation. The total energies are listed in Table 6.3. Despite the fact that the total energies are 224 millihartrees higher than their $\omega = 0$ counterparts, the difference is so constant over a wide range of bond distances that the spectroscopic parameters remain effectively unchanged. For example the equilibrium bond distance, r_e, increases from 0.7326 to 0.7338 Å, D_e falls from 354.1 to 353.5 kJ mol^{-1} and ν_e falls from 4635 to 4616 cm^{-1}. For this simple case at least, the neglect of the background has meant no deterioration in bond dissociation.

With the realization that attenuation has little effect on the bonding of H_2 the next step is to test CASE on a range of more complicated electronic structures. Table 6.4 presents the atomization energies for 32 first and second row molecules. Zero-point vibrational corrections are not included. Again, the total energies are raised by large amounts. But there is almost imperceptible movement in the relative (that is, atomization) energies which are typically 0-3 kJ mol^{-1}. The largest error in the set (6.4 kJ mol^{-1}) occurs for LiF where the attenuated calculation fails to capture all of the Coulombic stabilization. This is the most ionic species of the 32 molecules, so it is not surprising that CASE 'fails' here as the non-polar assumption is least valid.

Table 6.3: UHF/6-31G** energies (hartrees) of H_2 as a function of bond length (Å)

	$E(R)$		$E(R) - E(\infty)$	
R	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$
0.4	-0.93620	-0.71184	+0.06027	+0.06107
0.5	-1.06148	-0.83730	-0.06502	-0.06439
0.6	-1.11393	-0.88993	-0.11747	-0.11702
0.7	-1.13050	-0.90667	-0.13403	-0.13376
0.8	-1.12843	-0.90479	-0.13197	-0.13188
0.9	-1.11652	-0.89305	-0.12005	-0.12014
1.0	-1.09947	-0.87619	-0.10301	-0.10329
1.1	-1.07994	-0.85685	-0.08348	-0.08394
1.2	-1.05940	-0.83649	-0.06293	-0.06359
1.3	-1.04147	-0.81847	-0.04501	-0.04556
1.4	-1.02864	-0.80553	-0.03217	-0.03262
1.5	-1.01950	-0.79631	-0.02303	-0.02340
1.6	-1.01299	-0.78974	-0.01653	-0.01683
1.7	-1.00836	-0.78505	-0.01189	-0.01214
1.8	-1.00505	-0.78169	-0.00858	-0.00878
1.9	-1.00267	-0.77928	-0.00621	-0.00637
2.0	-1.00097	-0.77754	-0.00450	-0.00464
∞	-0.99647	-0.77291	-0.00000	-0.00000

Table 6.4: HF/6-31G* total energies (hartrees), atomization energies (kJ mol^{-1} and MP2/6-31G* correlation energies (millihartrees) of various molecules

	Total Energy		Atomiz. Energy		Correlation Energy	
Molecule	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$
H_2	-1.12679	-0.90305	342.2	341.7	-17.4	-17.3
H_2CCH_2	-78.03107	-72.85255	1775.9	1776.1	-263.2	-261.7
H_2CO	-113.86372	-107.21602	1056.7	1058.9	-311.2	-309.9
H_2NNH_2	-111.16800	-104.40912	1061.2	1062.4	-336.4	-335.2
H_3CCH_3	-79.22854	-73.82606	2303.7	2302.8	-275.4	-274.0
H_3COH	-115.03419	-108.16217	1513.5	1513.7	-319.1	-318.0
HCCH	-76.81560	-71.86094	1200.9	1201.9	-260.6	-258.9
HCN	-92.87019	-87.23785	802.5	805.9	-296.8	-295.0
HCO	-113.24518	-106.70926	740.8	743.0	-295.2	-293.9
HOOH	-150.76012	-142.41897	514.0	516.4	-374.8	-373.7
Li_2	-14.86689	-13.52642	10.9	13.6	-20.0	-18.3
LiF	-106.93418	-101.18581	361.9	355.4	-195.3	-194.9
LiH	-7.98087	-7.19726	134.6	131.7	-15.6	-15.4
BeH	-15.14731	-13.91139	215.6	211.2	-24.1	-23.9
CH	-38.26485	-35.78739	225.1	225.3	-77.6	-77.2
$CH_2(^1A_1)$	-38.87219	-36.28291	511.6	511.7	-101.8	-101.4
$CH_2(^3B_1)$	-38.92142	-36.33159	640.9	639.5	-86.0	-85.7
CH_3	-39.55892	-36.85718	1006.5	1004.8	-114.1	-113.6
CH_4	-40.19507	-37.38157	1368.6	1366.9	-142.0	-141.3
CN	-92.17398	-86.65255	282.7	283.8	-220.8	-219.7
CO	-112.73448	-106.31014	708.1	709.7	-293.7	-292.4
CO_2	-187.62841	-177.14614	996.9	1002.0	-490.0	-487.9
N_2	-108.93540	-102.62432	431.9	434.9	-326.2	-324.4
NH	-54.95924	-51.69139	198.4	198.5	-102.2	-101.9
NH_2	-55.55731	-52.17765	460.5	460.5	-136.4	-136.0
NH_3	-56.18384	-52.69221	797.4	796.9	-173.5	-173.0
NO	-129.24730	-122.03288	204.6	207.0	-317.2	-315.9
O_2	-149.60681	-141.48969	102.3	105.8	-347.5	-346.2

Table 6.4 (continued)

	Total Energy		Atomiz. Energy		Correlation Energy	
Molecule	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$
OH	-75.38186	-71.21088	261.7	261.8	-141.3	-141.0
OH_2	-76.00981	-71.72696	602.3	602.2	-189.4	-189.0
F_2	-198.67283	-188.52364	-149.9	-146.9	-366.0	-365.1
FH	-100.00229	-94.81532	365.2	365.1	-181.9	-181.5

While CASE was designed with Hartree-Fock in mind, given the good cancellation of error, it may be possible to obtain useful correlation energies with the CASE integrals. The simplest correlation method to try this with is second-order Møller-Plesset perturbation theory (equation (1.75)). To investigate the effect of Coulomb attenuation on the integrals required, orbital energies from an $\omega = 0$ Hartree-Fock calculation have been used. The MP2/6-31G* correlation energy has then been calculated for each of the 32 molecules above, and is presented in the last two columns of Table 6.4. The numerators of equation (1.75) are the difference between two integrals which allows a cancellation of error, producing good agreement with unattenuated calculations. The attenuated correlation calculations are systematically lower than traditional values, but the difference is only slight in all cases. It seems that neglect of the background is of little importance in the calculation of correlation energies.

With the use of localized orbitals [194] and a Laplace transform [195, 196] to remove the denominators of equation (1.75), the introduction of CASE here may provide a way to lower the scaling of MP2; however, this has not been investigated here.

There are some properties for which the CASE approximation performs poorly, though. The most severe failure is for atomic ionization energies, which are listed in Table 6.5. The UHF/6-31G* ionization energies show a marked decrease with the introduction of attenuation. The background is important here because, for the first time in this chapter, there is a distant interaction that does not cancel — an electron is lost from the system, which interacts with the cation that remains.

A simple estimate for this effect can be computed by calculating the effect of the first correction for the background on ionization. Taking just the first term of the Taylor expansion

Table 6.5: UHF/6-31G* total energies (hartrees) and ionization energies (eV)

	Total Energy (atom)		Total Energy (cation)		Ionization Energy		
Atom	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$	$\omega = 0$	$\omega = 0.1$	Δ
H	-0.49823	-0.38645	-0.00000	-0.00000	13.56	10.52	3.04
He	-2.85516	-2.51752	-1.99362	-1.76853	23.44	20.38	3.06
Li	-7.43137	-6.76063	-7.23554	-6.67201	5.33	2.41	2.92
Be	-14.56694	-13.44452	-14.27552	-13.26483	7.93	4.89	3.04
B	-24.52204	-22.83410	-24.23406	-22.66026	7.84	4.73	3.11
C	-37.68086	-35.31512	-37.28708	-35.03592	10.72	7.60	3.12
N	-54.38544	-51.22933	-53.87220	-50.83083	13.97	10.84	3.12
O	-74.78393	-70.72469	-74.34264	-70.39817	12.01	12.46	3.13
F	-99.36496	-94.28980	-98.79206	-93.83177	15.59	12.46	3.13
Ne	-128.47441	-122.27059	-127.75171	-121.66270	19.67	16.54	3.12

to represent the background,

$$\frac{\mathrm{erf}(\omega r)}{r} \approx \frac{2\omega}{\sqrt{\pi}}. \tag{6.12}$$

Then the effect of the background on the ionization energy is that of the departing electron interacting (via the background operator) with the charge that remains:

$$I_L = \langle \rho_{cation} | \frac{2\omega}{\sqrt{\pi}} | \rho_{electron} \rangle \tag{6.13}$$

$$= \frac{2\omega}{\sqrt{\pi}} \langle \rho_{cation} | \rho_{electron} \rangle \tag{6.14}$$

$$= \frac{2\omega}{\sqrt{\pi}}, \tag{6.15}$$

which has a value of 3.12 eV for $\omega = 0.1$, in very good agreement with the values listed in Table 6.5.

Note that this will not be the case for electron affinities, as the molecule that interacts with the approaching electron is neutral, rather than charged.

6.5 Conclusions

This chapter presents an introduction to linear methods and the CASE approximation, which neglects long-range Coulomb interactions by replacing the Coulomb operator with the

$\mathrm{erf}(\omega r)/r$ function. This leaves only $O(N)$ significant Coulomb interactions (how to calculate these interactions in only $O(N)$ work is covered in Chapter 7).

While Coulomb attenuation does have a major effect on the total energy of a system, the wavefunction is largely unaffected. Most relative energies are also reproduced quite well, as long as the number of particles is conserved in the reaction. Correlation energies, being the difference between two integrals, are also largely unaffected, something which may allow inexpensive correlation treatments in the future.

All of these properties can be ascribed to the blandness of the excluded background term, giving it little chemical significance. As well as being computationally useful, the CASE approximation is a comfortable idea for chemists as it affirms the well known idea that the behaviour of an atom within a molecule is principally governed by its immediate vicinity.

Chapter 7

Efficient Short-Range Integrals

"There is nothing so useless as doing efficiently that which should not be done at all" - Drucker

7.1 Introduction

With the realization that CASE may be a useful theory the next task is to improve the speed of CASE integral calculations. As mentioned in the previous chapter, by attenuation of the Coulomb operator to $\mathrm{erfc}(\omega r)/r$, the number of significant Coulomb interactions grows as only $O(N)$ (compared with the previous $O(N^2)$) for a large enough system. The task of finding which of the $O(N^4)$ integrals are significant still needs to be accomplished with an algorithm that is itself, at most, $O(N)$.

Before that is achieved the calculation of each single integral must be optimized. This reduces to finding the most efficient way to build the $(0)^{(m)}$s, as the various recurrence relations described in Chapter 5 are the same for all two-electron integrals, regardless of the operator used (to change an operator all that is required is to change the generation of $(0)^{(m)}$s).

This chapter presents how to calculate the short-range energy as efficiently as possible. It should also be pointed out that the work described in this chapter is not only useful for CASE, but is essential for all calculations using the KWIK family of methods, as each one requires the computation of some short-range elements.

7.2 Integral Generation

Traditional ERIs require the evaluation of the integral

$$I = \left(\frac{\zeta\eta}{\pi^2}\right)^{3/2} \iint e^{-\zeta r_1^2} \frac{1}{r_{12}} e^{-\eta|\mathbf{r}_2-\mathbf{R}|^2}\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.1}$$

$$= \frac{1}{2\pi^2 R^2} \int_0^\infty u \sin(u) e^{-u^2/4\mathrm{T}} \Im(u/R)\, du \tag{7.2}$$

where $\Im(k)$ is the Fourier transform of the operator. In this case

$$\Im(k) = \frac{4\pi}{k^2} \tag{7.3}$$

which, proceeding in the same way as section 5.2, leads to

$$I = R^{-1} \operatorname{erf}\left(R/\sqrt{\frac{1}{\zeta}+\frac{1}{\eta}}\right). \tag{7.4}$$

Note that this is the same result as equation (5.14), as the error function and $G_0(\mathrm{T})$ are closely related:

$$G_0(\mathrm{T}) = \frac{\operatorname{erf}(\sqrt{\mathrm{T}})}{\sqrt{2\mathrm{T}}}. \tag{7.5}$$

However the $\operatorname{erfc}(\omega r)/r$ operator requires the evaluation of

$$I_{CASE} = \left(\frac{\zeta\eta}{\pi^2}\right)^{3/2} \iint e^{-\zeta r_1^2} \frac{\operatorname{erfc}(\omega r_{12})}{r_{12}} e^{-\eta|\mathbf{r}_2-\mathbf{R}|^2}\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.6}$$

$$= \frac{1}{2\pi^2 R^2} \int_0^\infty u \sin(u) e^{-u^2/4\mathrm{T}} \Im(u/R)\, du \tag{7.7}$$

with

$$\Im(k) = \frac{4\pi}{k^2}\left(1 - e^{-k^2/4\omega^2}\right) \tag{7.8}$$

which produces, in effect, two integrals, one of which is the same as before

$$I_{CASE} = \frac{1}{2\pi^2 R^2}\left(\int_0^\infty u \sin(u) e^{-u^2/4\mathrm{T}} 4\pi \frac{R^2}{u^2}\, du - \int_0^\infty u \sin(u) e^{-u^2/4\mathrm{T}} 4\pi \frac{R^2}{u^2} e^{-R^2/4u^2\omega^2}\, du\right) \tag{7.9}$$

$$= R^{-1}\left\{\operatorname{erf}\left(R/\sqrt{\frac{1}{\zeta}+\frac{1}{\eta}}\right) - \operatorname{erf}\left(R/\sqrt{\frac{1}{\zeta}+\frac{1}{\eta}+\frac{1}{\omega^2}}\right)\right\}. \tag{7.10}$$

Rearranging the above formula into a more computationally useful form gives

$$[0]^{(m)}_{\mathrm{CASE}} = U\left\{\left(2v^2\right)^{m+1/2} G_m(\mathrm{T}) - \left(2v_\omega^2\right)^{m+1/2} G_m(\mathrm{T}_\omega)\right\}, \tag{7.11}$$

which is very similar to the traditional $[0]^{(m)}$ formula of

$$[0]^{(m)} = U\left(2v^2\right)^{m+1/2} G_m(\mathrm{T}) \tag{7.12}$$

except for the addition of two new variables:

$$\mathrm{T}_\omega = \upsilon_\omega^2 R^2 \tag{7.13}$$

$$\upsilon_\omega^2 = 1/\left[\frac{1}{\zeta} + \frac{1}{\eta} + \frac{1}{\omega^2}\right]. \tag{7.14}$$

Thus, computation of a CASE integral is very similar to a traditional integral. All the traditional work of before must still be done, and in addition, υ_ω and T_ω must be formed and the interpolation $G_m(\mathrm{T}_\omega)$ must be calculated. Luckily, T and T_ω are always greater than zero and, as ω is real,

$$\mathrm{T}_\omega \leq \mathrm{T}, \tag{7.15}$$

This is useful as it means that the interpolation table built for $G_m(\mathrm{T})$ does not need to be extended to include $G_m(\mathrm{T}_\omega)$. With all of these features in mind it should not be surprising if the computation of CASE integrals takes double the amount of time of the traditional integrals. This is time that needs to be made up by the removal of insignificant integrals.

There is a difference between CASE and traditional methods, however, when the values of T and T_ω become large. If T is large enough to use the classical multipole expression mentioned in Chapter 5 (and T_ω is not large enough), ie:

$$\mathrm{T}_{\mathrm{crit}} < \mathrm{T} \tag{7.16}$$

a computational saving can be made for one of the two parts of the CASE integral. However, if

$$\mathrm{T}_{\mathrm{crit}} < \mathrm{T}_\omega \quad (< \mathrm{T}) \tag{7.17}$$

then, applying the $G_m(\mathrm{T})$ expansion for large T

$$G_m(\mathrm{T}) \approx \frac{(2m-1)!!}{(2\mathrm{T})^{m+1/2}} \tag{7.18}$$

to both parts of the integral yields

$$[0]_{\mathrm{CASE}}^{(m)} \approx U\frac{(2m-1)!!}{2^{m+1/2}}\left\{\frac{\left(2\upsilon^2\right)^{m+1/2}}{\mathrm{T}^{m+1/2}} - \frac{\left(2\upsilon_\omega^2\right)^{m+1/2}}{\mathrm{T}_\omega^{m+1/2}}\right\} \tag{7.19}$$

$$= U(2m-1)!!\left\{\frac{\left(\upsilon^2\right)^{m+1/2}}{\left(\upsilon^2 R^2\right)^{m+1/2}} - \frac{\left(\upsilon_\omega^2\right)^{m+1/2}}{\left(\upsilon_\omega^2 R^2\right)^{m+1/2}}\right\} \tag{7.20}$$

$$= 0. \tag{7.21}$$

Therefore the size of T_ω is a convenient way to check the significance of a CASE integral. All integrals for which $T_{crit} < T_\omega$ do not need to be computed. The value of T_ω cannot be used to find the significant integrals unfortunately, as it requires the computation of R for each integral, which is $O(N^2)$ work. Another way of screening out integrals is needed, and this is the topic of the next section.

7.3 Boxing Scheme

There exist a number of $O(N)$ algorithms for calculating the energy due to a sum of pair-wise short-range interactions [197], all of which involve dividing the system into cubical boxes (or some other space-filling shape [198]) and then interacting the contents of each box with only those in neighbouring boxes. As the number of significant interactions per particle, M becomes $O(1)$ when $M \ll N$, the algorithm scales as $O(N)$. The algorithm used here is the most common in the field — the standard linked-cell method [199]. The additional complexity of alternative methods is unlikely to offer significant benefit, and the CFMM implementation in the Q-CHEM program, which uses the linked-cell method, has already been highly optimized.

The first step is to determine the number of boxes in each Cartesian direction. The sidelength of each box is taken to be the minimum distance over which two point charges, interacting via the CASE operator is negligible, that is, the value of x when

$$\frac{\mathrm{erfc}(\omega x)}{x} = \text{Desired integral accuracy}, \tag{7.22}$$

which can easily be found (given ω and the integral accuracy required) using the Newton-Raphson method [200]. The number of boxes in a direction is then the length of the molecule in that direction, divided by the sidelength.

Following the CFMM, each contracted shell-pair is assigned to a box. If the contracted shell-pair contains primitives with centers in different boxes, then the contracted shell-pair is split into two (or more) shell-pairs, one for each box that the original shell-pair centers straddle.

The box size above is fine for point charges, but the CASE integrals deal with continuous distributions, which can be much larger than the box itself. To solve this problem the CFMM notion of a well-separated index (WS) [157], which determines the distance above which a distribution behaves like a point charge (to within the integral accuracy required) was introduced. The Coulombic interaction between two spherical Gaussian charge distributions

can be represented (using the notation of the previous section) as

$$I = R^{-1}\,\mathrm{erf}\left(R/\sqrt{\frac{1}{\zeta}+\frac{1}{\eta}}\right). \tag{7.23}$$

The erf factor rapidly approaches 1 as R grows, and the two distributions then interact as classical point charges. Thus, the extent of a distribution is defined as

$$r_{\mathrm{ext}} = \sqrt{\frac{2}{\zeta}}\,\mathrm{erf}^{-1}(1-\epsilon) \tag{7.24}$$

where ϵ is the desired precision. A Taylor expansion is made for erf, allowing the extent to take the simple form

$$r_{\mathrm{ext}} = \frac{1}{2}\sqrt{\frac{2}{\zeta}\ln(\epsilon)}. \tag{7.25}$$

In molecular calculations the charge is represented as a product of Gaussians, which contain a prefactor U, dependent on the distance AB between the original shells. This can be included in the definition of extent, allowing for shorter interaction distances. The extent is thus given by

$$r_{\mathrm{ext}} = \frac{1}{2}\sqrt{\frac{2}{\zeta}\ln(\epsilon) - AB^2}. \tag{7.26}$$

The WS index is then defined as

$$WS = 2\left\lceil\frac{r_{\mathrm{ext}}}{l}\right\rceil \tag{7.27}$$

where l is the box size.

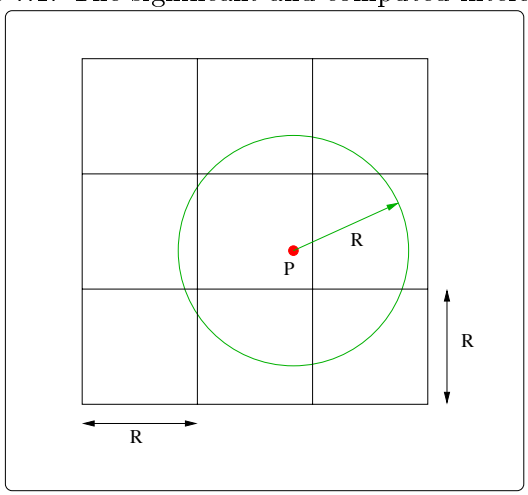
The boxes are then looped over, and every shell-pair of a box is interacted with the contents of all boxes within the range $WS+1$ multiples of the box length. This is exactly the same as the near-field part of the CFMM algorithm.

7.4 Integral Screening

The above boxing scheme is $O(N)$, yet it is still inefficient. Figure 7.1 compares all those interactions considered with those that are actually significant.

With an interaction distance of R (which is $WS+1$ for the above implementation) a shell-pair in a box will interact with 26 neighbouring interaction-boxes and the current interaction-box (where an interaction-box is defined as the cube of boxes with a total sidelength of $WS+1$). Thus each shell-pair will interact with all others over a volume of $27R^3$. Yet

Figure 7.1: The significant and computed interactions



only those shell-pairs within a distance R of the current point need to be calculated. This produces a sphere of volume $\frac{4}{3}\pi R^3$.

If the efficiency of a method is defined to be

$$\epsilon = \frac{\text{Number of significant interactions}}{\text{Number of interactions computed}}, \tag{7.28}$$

then the efficiency of the algorithm described above is only 15%! (It should be noted that in lower dimensions the efficiency is greater, 35% for 2D and 67% for 1D). This is an inherent deficiency of the linked-cell method.

There are several ways that the linked-cell method can be improved. One way is to order all particles by an increasing co-ordinate. Then, once the inter-shell-pair difference in this coordinate becomes greater than the interaction distance, interactions can be ignored. This requires sorting work and work to check the inter-shell-pair difference along this coordinate. This dramatically increases the efficiency in lower dimensions (100% and 52% for 1D and 2D respectively), but only yields a modest 23% efficiency for three-dimensional systems.

A second method is to change the basic cell structure. It has been suggested that the rhombic dodecahedron and truncated octahedron be used to tessellate space [201]. This allows a better approximation of a sphere than the cubes, but the cube is almost always used due to its geometric simplicity.

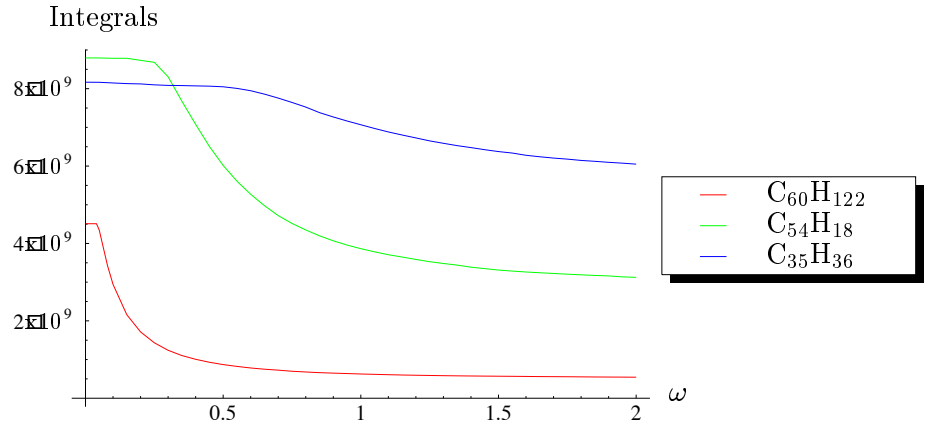
A third technique for increasing the efficiency is to reduce the basic cell size, increasing WS by a corresponding amount. This will allow better approximation of the sphere as the outlying corners of the interaction cube can be completely neglected. The trouble with this technique is that, in addition to a small amount of extra computational work, the vector loop

lengths are drastically reduced.

A more useful compromise is to introduce a pre-screening routine. Just before looping over all shell-pairs to be interacted with a given shell-pair, the T_ω values for all primitive shell-quartets are calculated, and if all are greater than T_{crit} the contracted shell-quartet does not need to be included, and so is removed from the interaction list. Thus, after the screening, only significant interactions will remain. This does introduce extra work (the calculated T_ω values are not stored for the later integral calculation), but the possibility of removing 85% of the integrals makes this well worthwhile.

7.5 Results and Discussion

Figure 7.2: number of significant ERIs as a function of ω

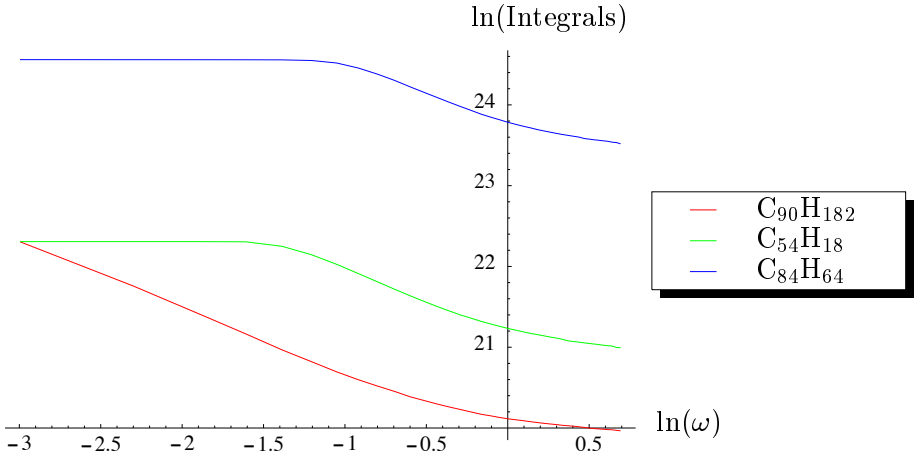


The first aspect to examine is just how many integrals Coulomb attenuation removes. Figure 7.2 shows the number of significant integrals (defined here as greater than 10^{-9}) as a function of ω for a linear alkane ($C_{60}H_{122}$), a hydrogen terminated graphite sheet ($C_{54}H_{18}$) and a hydrogen terminated diamond chunk ($C_{35}H_{36}$). The 6-31+G* basis is used in all cases. The three molecules may be characterized as one-dimensional, two-dimensional and three-dimensional, respectively.

Examination of the alkane curve shows that extremely small attenuation has no effect on the number of integrals, but once attenuation starts to remove integrals, the removal of integrals is very efficient and, beyond mild attenuation ($\omega = 0.2\ a_0^{-1}$), most of the integrals are insignificant. The graphene is more compact than the the alkane and, as a result, much larger attenuation is required before the pruning begins to take effect. Once the pruning has

begun, the curve shows a similar shape to the linear alkane, with a quite rapid removal of integrals. By $\omega = 1.0\ a_0^{-1}$ only half of the original integrals remain. The effect of CASE on this diamond chunk, however, is less spectacular. $C_{35}H_{36}$ is very compact and a significant reduction in the number of integrals does not begin until large ω values. Once pruning of integrals begins, though, a similar rapid decrease in the number of integrals is observed. Clearly, the efficiency of CASE depends on the shape of the system, as well as the number of atoms present.

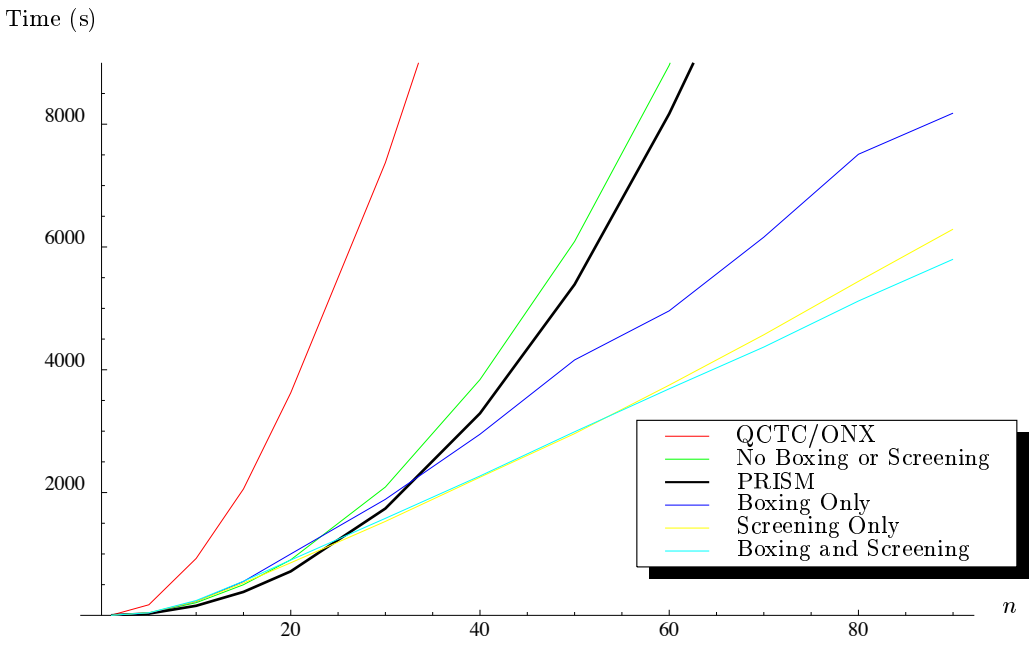
Figure 7.3: Log-Log plot of significant ERIs against ω



The rate of decrease of the three systems is presented in Figure 7.3 through a log-log plot. An integral accuracy of 10^{-6} has been used, with some slightly larger systems. The graph shows that, once integral removal has begun, the three-dimensional system ($C_{84}H_{64}$) shows a faster rate of decrease than the two- and one-dimensional systems. This is as expected, as increasing ω decreases the interaction distance, which, in a one-dimensional system, produces a corresponding decrease in the number of integrals. However in a three-dimensional system, the same decrease in interaction distance will remove the the number of integrals removed for the one-dimensional system, raised to the third power, simply because there are now three dimensions within which particles interact.

Figure 7.4 presents SGI Power-Challenge CPU times to calculate all the two-electron integrals required for a HF/6-31G* Q-Chem calculation on a series of linear alkanes, using $\omega = 0.25\ a_0^{-1}$ and an integral cutoff of 10^{-9}. The solid black curve shows the quadratic behaviour typical of the conventional $O(N^2)$ algorithm. Just above this lies the time for a CASE calculation using the traditional algorithm, clearly showing the cost of an extra interpolation. Included with these curves is the QCTC/ONX which shows an expensive

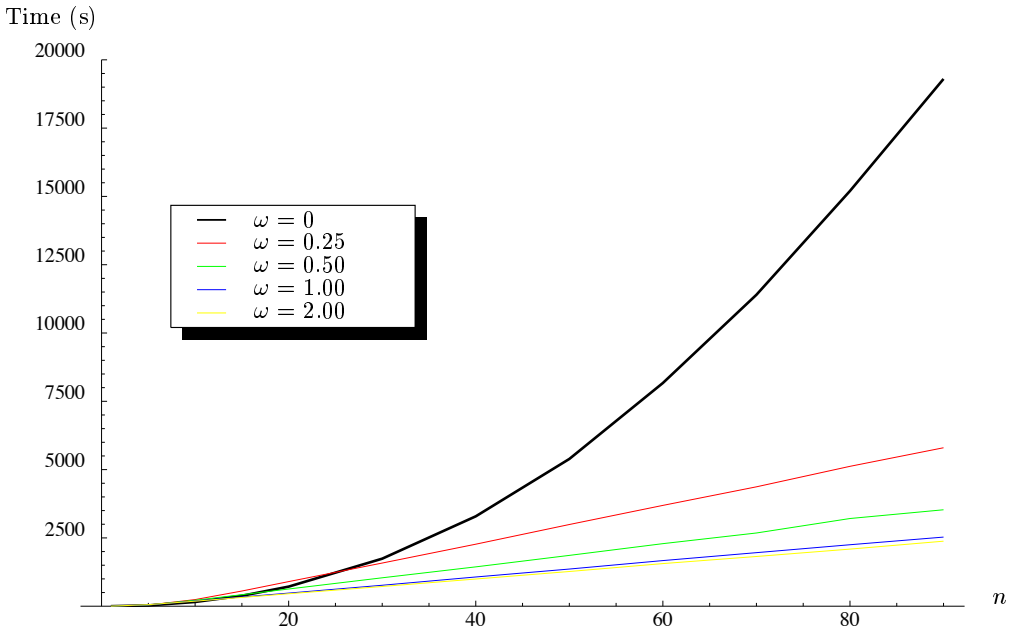
Figure 7.4: HF/6-31G* timings of linear alkanes C_nH_{2n+2} with $\omega = 0.25$



$O(N^2)$ behaviour (QCTC is a linear Coulomb method, similar to the CFMM). A direct comparison with ONX can not really be made as the two algorithms use different integral code and different screening strategies, yet the overall nature of the curve is still useful. The poor performance of ONX is due to the high integral accuracy used. As shown later in this section, ONX becomes more competitive at lower integral accuracy. The two near-linear curves are for CASE using integral screening, and boxing with integral screening. The remaining curve is for boxing only. The fluctuation of this final curve is due to the additional computational overhead required when previously unsplit shell-pairs now cover two boxes, forcing the shell-pairs to be split for no computational gain. It is perhaps surprising that integral screening, although formally $O(N^2)$, performs better than boxing alone for these systems. Yet the quadratic nature of this curve can just be seen, and screening will become more expensive than boxing alone for extremely large systems. The fastest timings for large systems are, not surprisingly, when both boxing and integral screening are used. It should be noted though that there is a small window where screening alone is the fastest method.

Figure 7.5 shows the effect on CPU times of altering ω for this same series of alkanes. $\omega = 0$ is the traditional $1/r$ code, exhibiting the expected quadratic trait, while higher ω values all show a linear behaviour. Notice that the time gained by doubling ω decreases as ω

Figure 7.5: HF/6-31G* CASE Boxing and Screening timings of linear alkanes C_nH_{2n+2}



increases. For these systems there seems little to be gained by increasing ω past $1.0\,a_0^{-1}$. By $\omega = 1.0\,a_0^{-1}$ the integrals that remain involve diffuse orbitals, requiring very large ω values (much larger than $2\,a_0^{-1}$) to render them insignificant.

The effect of changing the integral accuracy (significance level) on the time to calculate the integrals for $C_{50}H_{102}$ is presented in Figure 7.6. The ratio of the conventional CPU time to each method's time reveals that all of these methods become relatively faster when the integral threshold is lowered. This is due to the extra chance (at the shell-quartet level) to screen out computational work based on the integral accuracy required. The greatest increase is seen for ONX, yet with these high accuracy computations it is still poor.

Boxing and screening behave in similar ways with changes to the basis set. Figure 7.7 shows the the ratio of CPU times for the new CASE implementations with that of the traditional algorithm. The more uncontracted the basis set, the faster the method relative to the unattenuated calculation. This is a reflection of the time taken to test each *primitive* shell-pair for its T_ω value. If a shell-pair has only a small degree of contraction, the amount of testing required is small. ONX shows a similar type of behaviour, which may just be a product of the different integral technology, as well as a significant slow-down with the addition of d functions, which is consistent with a delocalization of charge removing the

Figure 7.6: Ratio of PRISM time to Method time for $C_{50}H_{102}$ at various integral thresholds

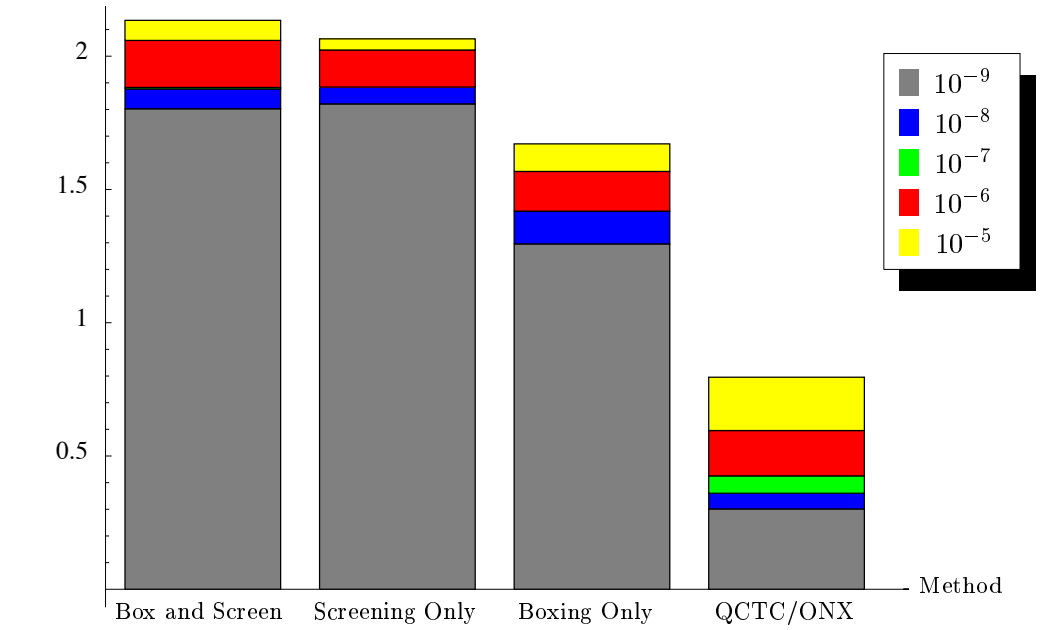
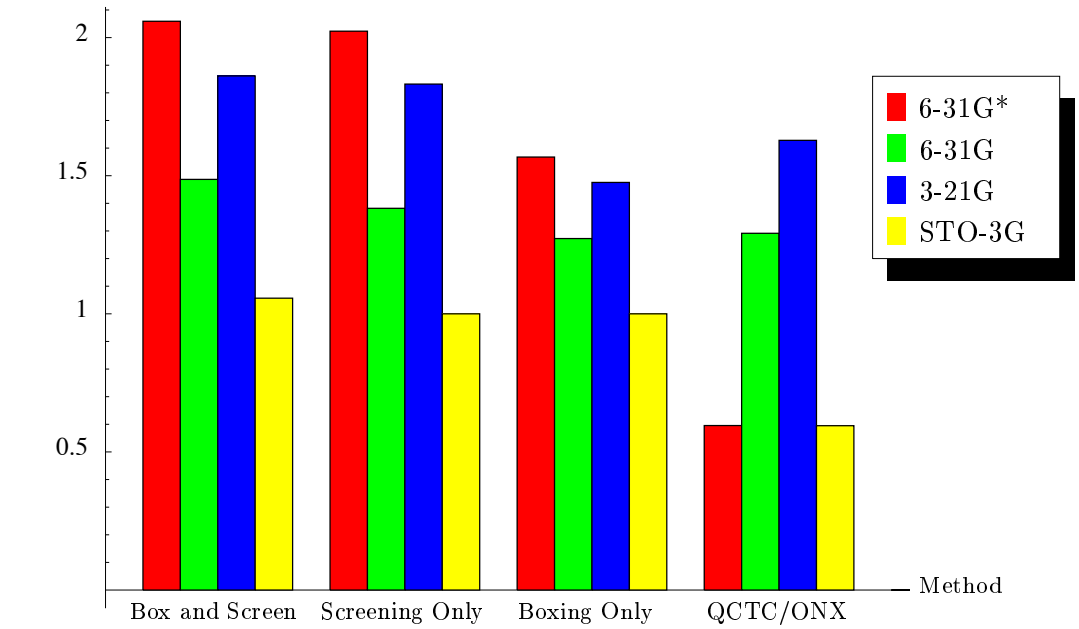
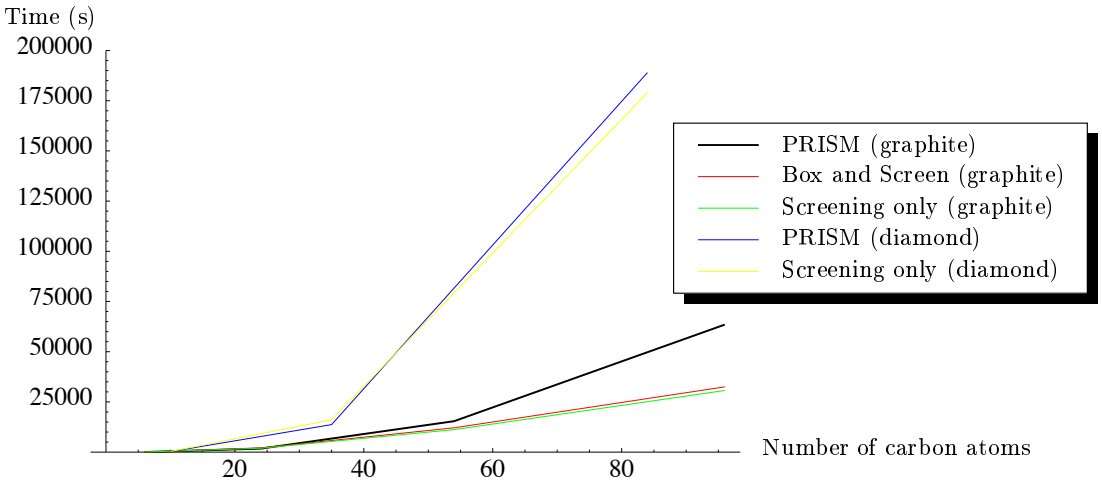


Figure 7.7: Ratio of PRISM time to method time for $C_{50}H_{102}$ with various basis sets



benefits of density screening.

Figure 7.8: HF/6-31G* timings for graphite and diamond chunks with $\omega = 0.5\ a_0^{-1}$



The performance of CASE on two- and three-dimensional systems can be seen in Figure 7.8. The black curve is the timings for the traditional PRISM method on a series of graphite sheets. Below this curve lie the boxing with screening and screening only curves. These CASE curves show linear behaviour, but are not as efficient as those for the linear alkanes. The CASE timings for the diamond chunks however are disappointing. It is only for $C_{84}H_{64}$ that CASE is faster, and this is just the screening, without boxing. The answer for the poor performance lies in the physical size of the system. The largest three-dimensional structure included ($C_{84}H_{64}$) is only 23 a_0 long on its largest side, compared with 36 a_0 for the largest graphite sheet used ($C_{96}H_{24}$) and 195 a_0 for the largest linear alkane ($C_{90}H_{182}$).

7.6 Concluding Remarks

This chapter has presented an efficient way to calculate the short-range Coulomb energy. This is the first $O(N)$ implementation of the CASE approximation, allowing very fast computation of CASE energies for large systems. This has been achieved through the introduction of boxing code, similar to that of the CFMM. A further speed-up has been achieved by introducing a screening for significant integrals. The increase in speed is heavily dependent on the shape of the system. For modest ω values, CASE represents a useful increase in speed even for moderately-sized systems (eg: $C_{50}H_{102}$).

The efficient implementation of the short-range integral code is crucial for the success of KWIK and other methods which can be viewed as providing corrections for the neglected part of CASE theory. These corrections will allow much larger values for ω thus enabling the CASE code to run even faster.

Chapter 8

Coulomb-Attenuated Exchange Energy Density Functionals[1]

8.1 Introduction

The CASE approximation presented in the previous two chapters was introduced to overcome the two-electron bottleneck in Hartree-Fock theory. However DFT also suffers from this bottleneck and so CASE DFT will also be useful. The attenuation must be done in a consistent manner, though. The Coulomb, exchange and correlation energies all stem from the Coulomb operator in the Schödinger equation and so must all be attenuated in a systematic way.

In HF theory this balance is achieved automatically as these three energies are calculated via the two-electron integrals. In DFT the exchange and correlation energies are determined via density functionals. Consequently, in order to perform CASE DFT calculations, appropriately Coulomb-attenuated density functionals need to be found.

Following the derivation of the exchange energy in Chapter 6 of Parr and Yang [83], the Coulomb-attenuated exchange energy may be written in terms of the first-order spinless density matrix

$$K(\omega) = \frac{1}{4}\iint \frac{\operatorname{erfc}(\omega r_{12})}{r_{12}} |\rho_1(\mathbf{r}_1, \mathbf{r}_2)|^2 \, d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.1}$$

$$= \pi \int_0^\infty s \operatorname{erfc}(\omega s) \int |\rho_1(\mathbf{r}, s)|^2 \, d\mathbf{r} ds \tag{8.2}$$

[1] The work described in this chapter has been carried out in collaboration with Prof. John Pople

where the co-ordinates

$$\mathbf{r} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2) \qquad \mathbf{s} = \mathbf{r}_1 - \mathbf{r}_2 \tag{8.3}$$

have been introduced.

8.2 Coulomb-Attenuated Dirac Exchange Functional

As mentioned in section 2.6, the Dirac exchange functional is derived from the uniform electron gas. Again following Parr and Yang, the orbitals of the UEG can be represented as plane waves, giving the first-order spinless density matrix

$$\rho_1(\mathbf{r}, s) = 3\rho(\mathbf{r}) \left[\frac{\sin t - t\cos t}{t^3} \right] \tag{8.4}$$

where

$$t = k_F(\mathbf{r})s \tag{8.5}$$

$$k_F(\mathbf{r}) = \left[3\pi^2\rho(\mathbf{r})\right]^{1/3}. \tag{8.6}$$

Substitution of the UEG density matrix into equation (8.1) provides a formula for the attenuated exchange energy of the uniform electron gas:

$$K^{LSDA}(\omega) = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int \rho^{4/3}(\mathbf{r}) \int_0^\infty \operatorname{erfc}\left(\frac{\omega}{k_F(\mathbf{r})}\right) \frac{4(\sin t - t\cos t)^2}{t^5}\, dt d\mathbf{r}. \tag{8.7}$$

Performing the integration over t (by parts) gives the spin-compensated exchange functional

$$K^{LSDA}(\omega) = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int \rho^{4/3}(\mathbf{r}) F\left(\frac{\omega}{k_F}\right) d\mathbf{r} \tag{8.8}$$

with the function

$$F(\lambda) = 1 - \frac{2\lambda}{3}\left[2\sqrt{\pi}\operatorname{erf}\left(\frac{1}{\lambda}\right) - 3\lambda + \lambda^3 + (2\lambda - \lambda^3)\exp\left(\frac{-1}{\lambda^2}\right)\right] \tag{8.9}$$

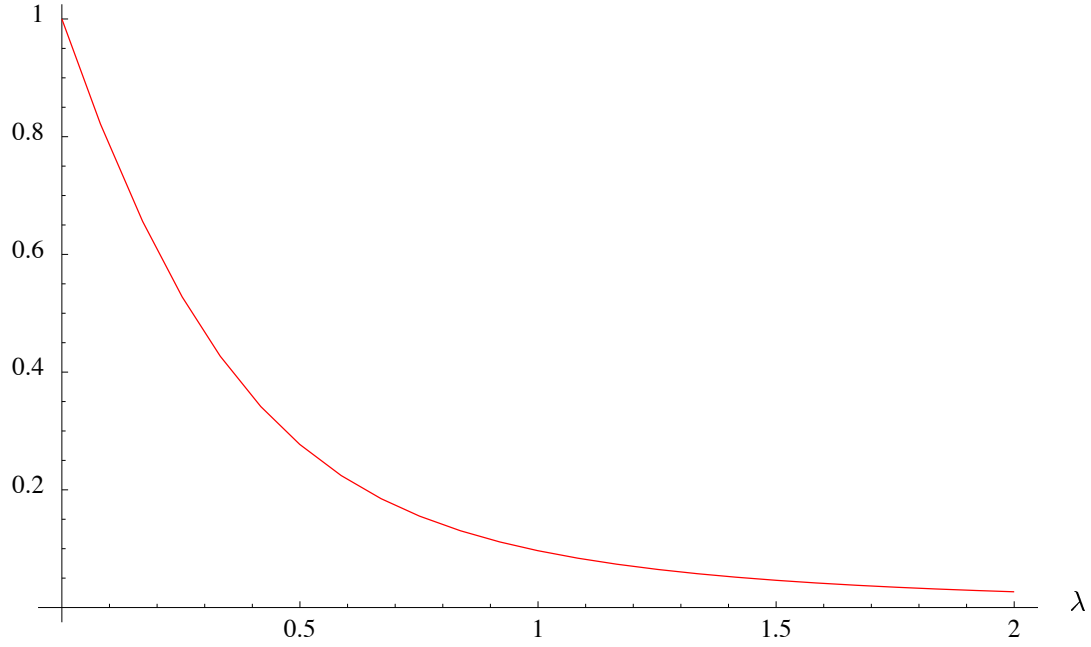
which is plotted in Figure 8.1.

The spin-polarized exchange functional can easily be derived from the spin-compensated version, and is given by

$$K_\sigma^{LSDA}(\omega) = \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} \int \rho_\sigma^{4/3}(\mathbf{r}) F\left(\frac{\omega}{k_F^\sigma}\right) d\mathbf{r}. \tag{8.10}$$

The above formula and its derivative with respect to ρ have been implemented into the Q-Chem program, allowing the self-consistent energy can be obtained.

Figure 8.1: Function $F(\lambda)$ as defined by equation (8.9)



By expanding the erfc function of equation (8.7) as a Taylor series in ω about $\omega = 0$ and integrating term by term produces the small-ω expansion

$$K_\sigma^{LSDA}(\omega) \approx \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} \int \rho_\sigma^{4/3}(\mathbf{r})\, d\mathbf{r} - \frac{\omega N_\sigma}{\sqrt{\pi}} + \ldots \tag{8.11}$$

where N_σ is the number of electrons of that spin. This also confirms that the attenuated function reduces to the Dirac functional in the limit $\omega \to 0$. The large-ω expansion is

$$K_\sigma^{LSDA}(\omega) \approx \frac{\pi}{2\omega^2} \int \rho_\sigma^2(\mathbf{r})\, d\mathbf{r} - \frac{6^{5/3}\pi^{7/3}}{80\omega^4} \int \rho_\sigma^{8/3}\, d\mathbf{r} + \ldots \tag{8.12}$$

Note that the leading term here involves the integral of ρ^2, instead of the traditional $\rho^{4/3}$.

8.3 Coulomb-Attenuated Exact Exchange Functional

Berkowitz [202] has shown that the density matrix for any homogeneous system can be expanded as

$$\rho_1^2(\mathbf{r}, s) = \rho^2(\mathbf{r}) - [4\rho(\mathbf{r})\tau(\mathbf{r}) - \rho(\mathbf{r})\nabla^2\rho(\mathbf{r})]\frac{s^2}{12} + \ldots \tag{8.13}$$

where τ is a generated from the molecular spin orbitals $\chi_i(\mathbf{r})$,

$$\tau(\mathbf{r}) = \sum_i^{occ} |\nabla\chi_i(\mathbf{r})|^2. \tag{8.14}$$

Substituting this density matrix into equation (8.1) yields an expansion for the spin-polarized exact exchange functional

$$K_\sigma(\omega) \approx \frac{\pi}{2\omega^2}\int \rho_\sigma^2(\mathbf{r})\,d\mathbf{r} - \frac{\pi}{32\omega^4}\int [4\rho_\sigma\tau_\sigma + |\nabla\rho_\sigma|^2]\,d\mathbf{r} + \ldots \tag{8.15}$$

Because of the τ term this is not a density functional in the purest sense, however, density functionals containing $\tau(\mathbf{r})$ are becoming more popular, for example a recent functional proposed by Becke [203].

Perhaps the most surprising thing to note about this new functional is that the leading term is the same as that for the large-ω expansion of the attenuated Dirac functional. Thus the LSDA exchange approximation becomes exact when the Coulomb operator is strongly attenuated — even for highly inhomogeneous systems. This is a reflection of the fact that for extremely strong attenuation the short-sightedness of the Coulomb operator allows the density to appear roughly constant.

A further interesting point about this new functional and the large-ω expansion of the attenuated Dirac functional is that the leading term only involves integrals of ρ^2. This is very significant computationally because, if a Gaussian basis set is used, the integral can be evaluated analytically, without resorting to the numerical quadrature that is required for the more traditional $\rho^{4/3}$. Furthermore, all the terms of the exact expansion involve powers of the density that can be computed analytically. It should be noted that this is not the case for the Dirac exchange expressions of the previous section.

8.4 The Hydrogen Atom

The simplicity of the H atom allows its exchange energy to be written in closed form, making it the ideal model to examine the performance of the two new attenuated functionals. The exact H atom density is

$$\rho(r) = \frac{e^{-2r}}{\pi}. \tag{8.16}$$

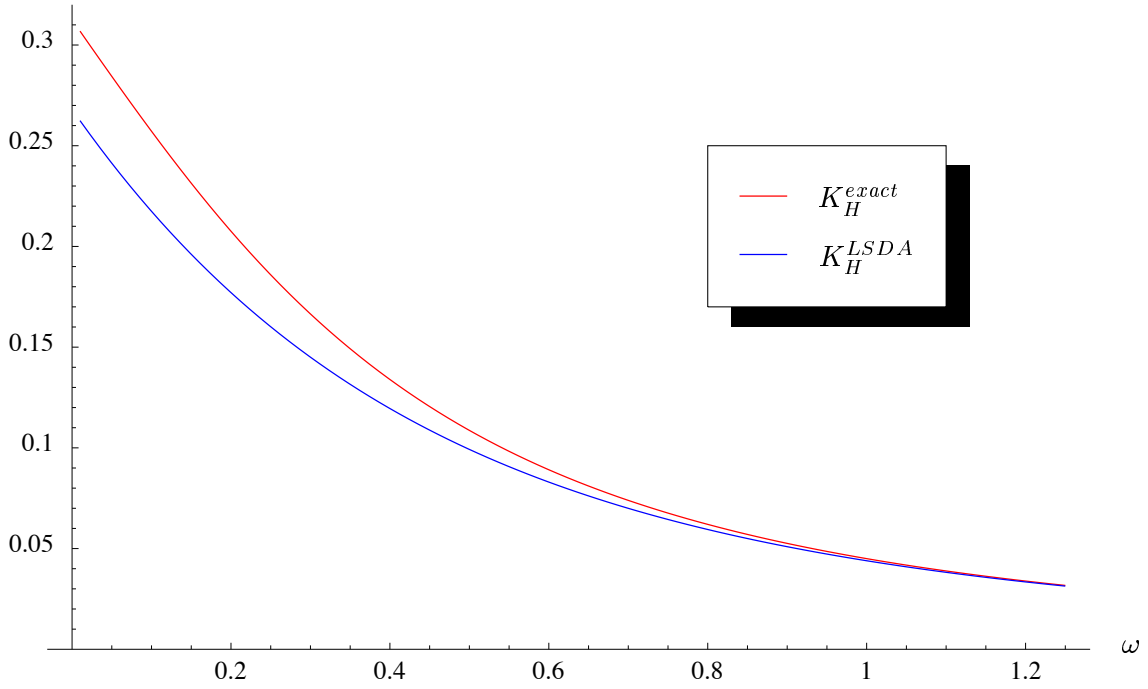
Substituting this into the exact exchange energy,

$$K(\omega) = \frac{1}{2}\iint \rho(r_1)\frac{\operatorname{erfc}(\omega r_{12})}{r_{12}}\rho(r_2)\,d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.17}$$

and integrating out $\mathbf{r}_1$ and $\mathbf{r}_2$ gives

$$K(\omega) = \frac{5}{16} - \left(\frac{5}{16} - \frac{3}{8\omega^2} + \frac{1}{4\omega^4} - \frac{1}{6\omega^6}\right)\exp\left(\frac{1}{\omega^2}\right)\operatorname{erfc}\left(\frac{1}{\omega}\right) - \left(\frac{5}{8\omega} - \frac{1}{3\omega^3} + \frac{1}{6\omega^5}\right)\frac{1}{\sqrt{\pi}}. \tag{8.18}$$

Figure 8.2: The Coulomb-attenuated exact and Dirac exchange energies for the H atom



This is the upper curve of Figure 8.2. As would be expected, the energy decays from the unattenuated value of $5/16$ with increasing attenuation. The Coulomb-attenuated Dirac exchange energy has also been calculated via equation (8.10) and using Mathematica [193] to perform the $\mathbf{r}$ integration numerically. This is the lower curve of Figure 8.2, which, like the exact functional, decays from its $\omega = 0$ value of 0.2680, agreeing with the rule of thumb that the Dirac functional underestimates exchange energies by roughly 10%. The two curves approach each other very rapidly, suggesting that the leading terms of the large-ω expansion are particularly important. This can be seen by examining the two series with the H atom density

$$K_H^{LSDA} \sim \frac{1}{16\omega^2} - \frac{1}{35.698\omega^4} + \ldots \tag{8.19}$$

$$K_H \sim \frac{1}{16\omega^2} - \frac{1}{32\omega^4} + \ldots \tag{8.20}$$

confirming that the LSDA is a surprisingly good exchange approximation for this highly inhomogeneous system, even for rather modest values of ω.

8.5 Conclusion

The extension of CASE to density functional theory has been presented with the introduction of two attenuated density functionals. By following the original derivation of the Dirac exchange functional, except with the CASE operator, an attenuated LSDA exchange functional has been found.

When the Coulomb-attenuated exact and Dirac exchange functionals are expressed as asymptotic series in ω, the leading terms are identical, showing that the LSDA exchange functional becomes exact in the high attenuation limit. Finally, if a Gaussian basis is used, the terms of the exact expansion can be evaluated without resorting to quadrature on a grid.

Chapter 9

A Family of Attenuated Coulomb Operators

"There are many ways of going forward, but only one way of standing still." - Franklin D. Roosevelt

9.1 Introduction

Chapter 6 introduced the attenuated Coulomb operator and showed that, for some molecular properties it performs quite well. There are properties, though, for which the approximation is poor (for example, ionization potentials). The obvious way to improve results is to move the CASE operator closer to the Coulomb operator, by decreasing ω. There are other, perhaps more efficient, ways to remove at least part of the neglected background. This chapter presents a second way of moving towards $1/r$, forming the CASE(m) approximation, which corrects for the neglected background by approximating it as a sum of Gaussians. These new terms are included in the short-range operator.

9.2 The CASE(m) Approximation

The Taylor expansion about $r = 0$ of the background is

$$\frac{\mathrm{erf}(\omega r)}{r} = \frac{2\omega}{\sqrt{\pi}}\left(1 - \frac{\omega^2 r^2}{3 \cdot 1!} + \frac{\omega^4 r^4}{5 \cdot 2!} - \frac{\omega^6 r^6}{7 \cdot 3!} + \ldots\right) \tag{9.1}$$

which is very similar to that of a sum of Gaussians

$$\sum_{j=1}^{m} a_{mj} \exp(-\omega^2 \alpha_{mj}^2 r^2) = \sum_{j=1}^{m} 1 - \frac{a_{mj}\omega^2 \alpha_{mj}^2 r^2}{1!} + \frac{a_{mj}\omega^4 \alpha_{mj}^4 r^4}{2!} - \frac{a_{mj}\omega^6 \alpha_{mj}^6 r^6}{3!} + \ldots \tag{9.2}$$

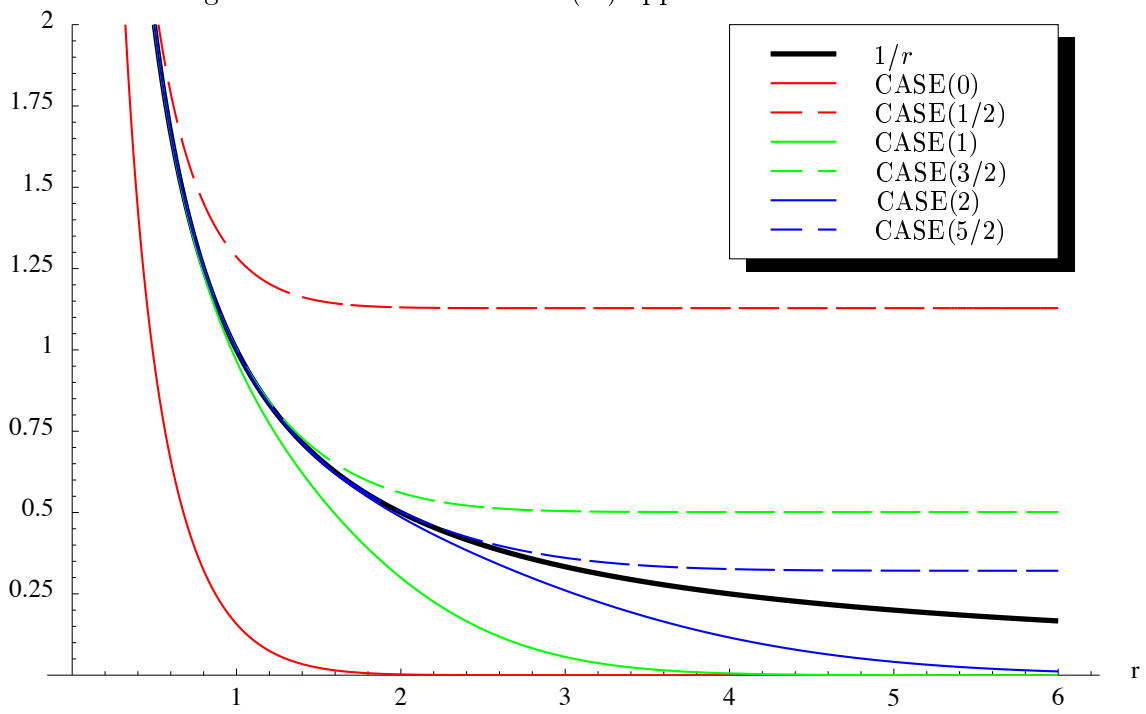
Equating the coefficients of the powers of r leads to a system of equations in a_{mj} and α_{mj}. The solution of these equations will yield a sum of Gaussians which reproduce the first $4m-2$ derivatives of the background at $r = 0$ for the operator:

$$S(r) = \frac{\operatorname{erfc}(\omega r)}{r} + \frac{2\omega}{\sqrt{\pi}} \sum_{j=1}^{m+1/2} a_{mj} \exp\left(-\omega^2 \alpha_{mj}^2 r^2\right). \tag{9.3}$$

The solutions up to and including three Gaussians are listed in Table 9.1. It should be noted that CASE(0) is the original $\operatorname{erfc}(\omega r)/r$, and that the remaining neglected term still has a Fourier term that decays as fast as a Gaussian.

The half integer values represent the inclusion of a constant, that is, a Gaussian with zero exponent. This will not alter the wavefunctions produced by the new operator, but it will reduce the magnitude of the background near the origin, thus improving accuracy in calculations where the total charge is not preserved. As can be seen in the following section, the addition of a constant to the operator introduces negligible extra work.

Figure 9.1: The first six CASE(m) approximations with $\omega = 1$



The first few CASE(m) approximations are illustrated in Figure 9.1. The simplest operator, CASE(0), is not equal to $1/r$ at any point. The addition of a constant produces CASE(1/2), which is correct at the origin, but overshoots thereafter. As m increases, the resulting CASE(m) remain faithful to the $1/r$ curve for longer before decaying to zero (when

Table 9.1: CASE(m) a_{mj} and α_{mj} for $m \leq 4$

m	Exponent, α_{mj}	Amplitude, a_{mj}
$\frac{1}{2}$	0	1.000000000000000
1	0.577350269189626	1.000000000000000
$1\frac{1}{2}$	0	0.444444444444444
	0.774596669241483	0.555555555555555
2	0.339981043584856	0.652145154862546
	0.861136311594053	0.347854845137454
$2\frac{1}{2}$	0	0.284444444444444
	0.538469310105683	0.478628670499366
	0.906179845938664	0.236926885056189
3	0.238619186083197	0.467913934572691
	0.661209386466265	0.360761573048139
	0.932469514203152	0.171324492379170
$3\frac{1}{2}$	0	0.208979591836735
	0.405845151377397	0.381830050505119
	0.741531185599394	0.279705391489277
	0.949107912342759	0.129484966168870
4	0.183434642495650	0.362683783378362
	0.525532409916329	0.313706645877887
	0.796666477413627	0.222381034453374
	0.960289856497536	0.101228536290376

m is integer) or a constant (half-integer m). Thus, the integer approximations are lower bounds to $1/r$, while the half-integer approximations are upper bounds to $1/r$. It should also be noted that the higher m values produce operators that decay to an insignificant value much more slowly, requiring the calculation of further integrals.

9.3 Addition Integrals

The introduction of these new terms to the operator necessitates the calculation of new integrals. Firstly, the constant is not treated as a Gaussian with zero exponent as above. It is more efficient to treat this as a correction to the various energy matrices.

The nuclear-nuclear correction is computationally trivial. It is computed as a double sum over the nuclei

$$E_{NN}^{\text{const}} = \sum_{AB} Z_A \frac{2\omega a_{mj}}{\sqrt{\pi}} Z_B \tag{9.4}$$

with Z_A and Z_B representing the nuclear charges.

For the Coulomb matrix, the correction is

$$J_{\mu\nu}^{\text{const}} = \sum_{\lambda\sigma} P_{\lambda\sigma} \iint \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1) \frac{2\omega a_{mj}}{\sqrt{\pi}} \phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{9.5}$$

$$= \frac{2\omega a_{mj}}{\sqrt{\pi}} \sum_{\lambda\sigma} P_{\lambda\sigma} S_{\mu\nu} S_{\lambda\sigma} \tag{9.6}$$

$$= \frac{2\omega a_{mj} N_{\text{elec}}}{\sqrt{\pi}} S_{\mu\nu}. \tag{9.7}$$

where N_{elec} is the number of electrons. Thus the correction is simply the addition of a scaled overlap matrix to the Coulomb matrix. This is formally $O(N^2)$ as there are N^2 elements of these matrices. However this is extremely fast $O(N^2)$ work and will not be noticed for even the largest systems currently studied. If the quadratic cost becomes a problem in the future, advantage can be made of the sparsity of $\mathbf{S}$ to determine the correction in linear work.

A similar addition to the Hamiltonian matrix is required, correcting for the change in the nuclear-electron attraction energy. The correction is

$$H_{\mu\nu}^{\text{const}} = \frac{2\omega a_{mj} Z}{\sqrt{\pi}} S_{\mu\nu} \tag{9.8}$$

where Z is the total nuclear charge. As would be expected, the nuclear-electron correction has the same scaling as the Coulomb correction.

The exchange matrix correction is

$$K_{\mu\nu}^{\text{const}} = \sum_{\lambda\sigma} P_{\lambda\sigma} \iint \phi_\mu(\mathbf{r}_1)\phi_\lambda(\mathbf{r}_1)\frac{2\omega a_{mj}}{\sqrt{\pi}}\phi_\nu(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{9.9}$$

$$= \frac{2\omega a_{mj}}{\sqrt{\pi}} \sum_{\lambda\sigma} P_{\lambda\sigma} S_{\mu\lambda} S_{\nu\sigma} \tag{9.10}$$

$$= \frac{2\omega a_{mj}}{\sqrt{\pi}} \mathbf{SPS} \tag{9.11}$$

Matrix multiplication is formally $O(N^3)$ — there are N^2 elements to calculate, each requiring N multiply-adds. Although, again this extremely fast work, and will not be noticed for any systems currently calculable. If it does become a problem in the future, sparsity of $\mathbf{S}$ can reduce this to $O(N^2)$. Then, if the system is an insulator, $\mathbf{P}$ will be sparse, allowing linear computation.

In order to add a constant to the Dirac DFT energy, the effect of a constant on the exchange energy of the uniform electron gas must be calculated. Beginning with equation (8.1) with the new operator and proceeding as before

$$K(\omega) = \frac{1}{4}\iint \left(\frac{\operatorname{erfc}(\omega r_{12})}{r_{12}} + \frac{2\omega a_{mj}}{\sqrt{\pi}}\right) |\rho_1(\mathbf{r}_1, \mathbf{r}_2)|^2 \, d\mathbf{r}_1 d\mathbf{r}_2 \tag{9.12}$$

$$= 9\pi \int \rho^2(\mathbf{r}) d\mathbf{r} \left[\int_0^\infty \operatorname{erfc}\left(\frac{\omega t}{k_F}\right) \frac{(\sin t - t\cos t)^2}{k_F^2 t^5} dt + \frac{2\omega a_{mj}}{\sqrt{\pi}} \int_0^\infty \frac{(\sin t - t\cos t)^2}{k_F^3 t^4} dt\right]. \tag{9.13}$$

The first term is the same as before, so continuing with only the second term gives

$$K(\omega)^{\text{const}} = 9\pi \int \rho^2(\mathbf{r}) \frac{2\omega a_{mj}}{\sqrt{\pi}} \frac{\pi}{6k_F^3} d\mathbf{r} \tag{9.14}$$

$$= \frac{\omega a_{mj}}{\sqrt{\pi}} \int \rho(\mathbf{r})\, d\mathbf{r} \tag{9.15}$$

The Gaussian terms are again trivial for the nuclear-nuclear term. For nuclear attraction, Coulomb and Hartree-Fock exchange terms, new two-electron integrals must be calculated. Proceeding as before, the ERI can be represented as

$$I_{CASE} = \frac{1}{2\pi^2 R^2} \int_0^\infty u \sin(u) e^{-u^2/4\mathrm{T}} \Im(u/R)\, du, \tag{9.16}$$

where the Fourier transform (of $\exp(-\alpha^2\omega^2 r^2)$) required is

$$\Im(k) = \frac{\pi^{3/2} e^{-k^2/4\alpha^2\omega^2}}{\alpha^3\omega^3} \tag{9.17}$$

which leads to

$$I_{CASE} = \frac{1}{2\sqrt{\pi}\alpha^3\omega^3 R^2} \int_0^\infty u \sin(u) \exp\left[-u^2\left(\frac{1}{4\mathrm{T}} + \frac{1}{4R^2\alpha^2\omega^2}\right)\right] du, \tag{9.18}$$

yielding

$$I_{CASE} = \left(\frac{v^2}{v^2+\alpha^2\omega^2}\right)^{3/2} \exp\left[-R^2/\left(\frac{1}{v^2}+\frac{1}{\alpha^2\omega^2}\right)\right]. \tag{9.19}$$

Thus the additional integrals are not as complex as those for CASE(0) allowing much faster computation. So, although there are more integrals to calculate, they are of a more efficient type and hence there is a trade-off.

As before, the effect of the additional terms in the operator on the Dirac functional must be determined. Beginning with the now familiar exchange energy functional

$$K(\omega)^{\text{addit}} = \frac{1}{4}\iint a\, e^{-\alpha^2\omega^2R^2} |\rho_1(\mathbf{r}_1,\mathbf{r}_2)|^2\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{9.20}$$

and substituting in the UEG density-matrix yields

$$K(\omega)^{\text{addit}} = 9\pi \int \rho^2(\mathbf{r}) d\mathbf{r} \int_0^\infty a\, e^{-\omega^2t^2/3k_F^2} \frac{(\sin t - t\cos t)^2}{t^4 k_F^3}\, dt \tag{9.21}$$

$$= \frac{a}{2\sqrt{\pi}} \int \rho(\mathbf{r}) F\left(\frac{\omega^2}{3k_F^2}\right) d\mathbf{r}. \tag{9.22}$$

9.4 Results

9.4.1 The Hydrogen Atom

As with the previous chapters, the accuracy of these new operators is first examined on a simple model system, the H atom. The total energy of the hydrogen atom is equal to its ionization energy, the worst property of those examined in Chapter 6. Thus the total energy is a good test for these new operators. Using the exact density the energy of the ground state is

$$E = \frac{-1}{2\pi}\int e^{-r}\nabla^2 e^{-r}\, d\mathbf{r} - \frac{1}{\pi}\int e^{-2r}\frac{\operatorname{erfc}(\omega r)}{r}\, d\mathbf{r} + \sum_{j=1}^{m+1/2} \frac{2\omega a_{mj}}{\sqrt{\pi}} \int e^{-\alpha_{mj}\omega^2 r^2}\frac{e^{-2r}}{\pi}\, d\mathbf{r} \tag{9.23}$$

$$= \frac{-1}{2} + \frac{2}{\omega\sqrt{\pi}} - \left(\frac{2}{\omega^2}-1\right)\exp\left(-\omega^{-2}\right)\operatorname{erfc}\left(\omega^{-1}\right) + \frac{2}{\omega^2}\sum_{j=1}^{m+1/2}\frac{a_{mj}}{\alpha_{mj}^3}\left[\frac{2}{\sqrt{\pi}\omega\alpha_{mj}} - \left(\frac{2}{\omega^2\alpha_{mj}^2}+1\right)\exp\left(\alpha_{mj}^{-2}\omega^{-2}\right)\operatorname{erfc}\left(\alpha_{mj}^{-1}\omega^{-1}\right)\right]. \tag{9.24}$$

The CASE(m) errors for the H atom are listed in Table 9.2 for a variety of m and ω values. As expected, CASE(0) produces poor total energies for all of the ω values listed. The most surprising feature, however, is the dramatic improvement with CASE(1/2), simply a

Table 9.2: CASE(m) - Exact energies for the ground state of the H atom

m	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0.1$
0	0.468989	0.396675	0.312866	0.217369	0.111734
1/2	-0.095201	-0.054677	-0.025648	-0.008307	-0.001104
1	0.013634	0.005886	0.001803	0.000295	0.000011
3/2	-0.002336	-0.000744	-0.000148	-0.000012	-0.000000
2	0.000429	0.000103	0.000014	0.000001	0.000000
5/2	-0.000086	-0.000016	-0.000001	-0.000000	-0.000000
3	0.000018	0.000003	0.000000	0.000000	0.000000
7/2	-0.000004	-0.000000	-0.000000	-0.000000	-0.000000

constant added to the operator. CASE(1/2) systematically overestimates the total energy (as would be expected from a comparison with the Coulomb operator), but the error is now as much as two orders of magnitude smaller than CASE(0).

As m increases further the errors continue to fall, providing accurate results, especially when ω is small. Because the hydrogen atom involves only a single Coulomb potential and CASE(integer m) bounds $1/r$ below and CASE(half-integer m) bounds $1/r$ above, the errors are consistently positive and negative respectively. Significantly, CASE($m + 1/2$) yields smaller errors than CASE(m) which is important as, computationally, they are virtually identical.

9.4.2 Madelung Constant of NaCl

All the results for the CASE approximation presented so far are for small molecules. It can be argued that, because these systems lack significant long-range effects, they are not very demanding tests of an operator which neglects long-range terms. As a more demanding test of the approximation, CASE(m) has been used to find the Madelung constant (the potential of an ion in the field of all others) of NaCl with a lattice spacing of 5.33 au. This is a harsher test as the system involved is of infinite size and the property sought includes very long-range contributions. The exact value [204] of 0.327873 can not be found by straightforward summation as the convergence is too slow. However the rapid decay of the CASE(m) operators produces rapidly converging sums. The deviations from the exact result for a variety of m and ω values are presented in Table 9.3.

Table 9.3: CASE(m) - Exact values for the Madelung constant of NaCl

m	$\omega = 0.5$	$\omega = 0.4$	$\omega = 0.3$	$\omega = 0.2$	$\omega = 0.1$
0	-0.327689	-0.325013	-0.303278	-0.224731	-0.112838
1/2	0.236501	0.126338	0.035236	0.000944	0.000000
1	-0.066350	0.044884	0.031996	0.000944	0.000000
3/2	-0.051156	-0.038349	-0.000148	0.000753	0.000000
2	0.043585	-0.000494	-0.005272	0.000252	0.000000
5/2	-0.006689	0.010898	0.000139	-0.000004	0.000000
3	-0.012523	-0.004867	0.001190	-0.000037	0.000000
7/2	0.011127	-0.000771	-0.000313	-0.000006	0.000000

Again the CASE(0) estimates are poor. However, as for the hydrogen atom, the errors are markedly smaller for the higher CASE(m) approximations, especially when ω is small. The Madelung constant results from large positive and negative Coulomb interactions, and therefore it should not be surprising that patterns are less obvious here. One pattern that does emerge is that the error is closely linked to the rate of decay of the operator. Those operators which decay to insignificance the slowest are the most accurate. This is an expected, though not entirely useful, property as the speed of the method is directly linked to the decay speed. The other pattern to note is that the half-integer approximations are more accurate than the integer values, yet they are equally demanding computationally.

9.5 Conclusions

This chapter has presented a second way to extend the CASE operator towards the Coulomb operator. By the addition of Gaussian terms to the short-range operator the neglected background can be corrected for. The Gaussian exponents and amplitudes are chosen to match as many derivatives as possible of the background at $r = 0$. The accuracy of the resulting operators is closely linked to the overall decay speed of the operator. So, while the Gaussian corrections are not extremely useful from an accuracy perspective, they may provide a useful way to share the short- and long-range work in KWIK theories, which use CASE as the short-range operator.

The addition of a constant, correcting for the magnitude of the background at $r = 0$, is surprisingly useful. Total energies are now overestimated, but are more accurate than before.

This does cause some new problems, however. With the half-integer approximations, two infinitely separated charged moieties have a non-zero interaction energy, thus the operators can behave in a non-physical manner. The effect of CASE(0) and CASE(1/2) will be examined in the following chapter.

Chapter 10

Effects of Coulomb Attenuation on Chemical Properties

"After having spent years trying to be accurate, we must spend as many more in discovering when and how to be inaccurate." - Ambrose Bierce

10.1 Introduction

The various CASE approximations have been presented in the previous chapters. The one thing that each of these chapters lacks, however, is an examination of the performance of the approximations on a variety of chemical properties. It is useful to have an idea of the maximum value of ω permissible for each property under investigation, remembering that the larger ω, the faster the calculation. This chapter investigates the performance of CASE on a selection of chemical properties, so that sensible choices for ω can be made. This will also be useful if, in the future, a correction for the background is calculated via the long-range operator, as there is the opportunity to implement background corrections only approximately, depending on the accuracy required for that property.

10.2 Wavefunctions

The effect of CASE on the wavefunction can be seen from the density matrix. Table 6.2 presented two typical molecular orbitals. From this it can be seen that any changes that do occur, seem to occur in a reasonably uniform manner. Thus, an averaging of the error across the wavefunction does not remove too much information. The RMS error from each

element of the density matrix, computed for a variety of molecules and ω values, is listed in Table 10.1. The STO-3G basis has been used throughout. To use a larger basis would allow a small degree of linear dependence, thus providing flexibility in the density matrix where two apparently different density matrices could be represented by the same wavefunction.

Table 10.1: RMS Errors in the density matrix for a variety of molecules and ω values

Molecule	$\omega = 0.05$	$\omega = 0.10$	$\omega = 0.15$	$\omega = 0.20$	$\omega = 0.30$	$\omega = 0.40$	$\omega = 0.50$
H_2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
CN^-	0.000031	0.000215	0.000576	0.000993	0.001501	0.002691	0.005751
AlF_3	0.000023	0.000138	0.000350	0.000782	0.002521	0.004343	0.006320
C_2Cl_4	0.000024	0.000243	0.001038	0.002433	0.005276	0.006392	0.006361
$SiCl_4$	0.000022	0.000201	0.001098	0.002898	0.006857	0.008314	0.007576
ATP	0.008373	0.010556	0.000950	0.003840	0.011166	0.012931	0.013982
L-glycerate	0.000060	0.000314	0.000722	0.001520	0.004601	0.008989	0.014194
SO_2	0.000038	0.000275	0.000866	0.001997	0.006242	0.012074	0.017615
CH_3COCl	0.000030	0.000218	0.000777	0.001953	0.006003	0.011341	0.017623
C_6H_6	0.000016	0.000192	0.000865	0.002227	0.006586	0.012261	0.018706
C_4H_{10}	0.000016	0.000177	0.000751	0.001914	0.006135	0.012501	0.020285
CH_3CONH_2	0.000028	0.000221	0.000790	0.001984	0.006411	0.012805	0.020515
C_4H_6	0.000032	0.000266	0.000958	0.002259	0.006663	0.013097	0.020816
$HCOOCH_3$	0.000023	0.000203	0.000779	0.001989	0.006445	0.013063	0.021186
C_5H_8	0.000017	0.000183	0.000827	0.002229	0.007176	0.013858	0.021208
CH_3SCH_3	0.000020	0.000182	0.000784	0.002156	0.007461	0.015258	0.024595
C_2H_6	0.000016	0.000142	0.000599	0.001705	0.006591	0.014841	0.025518
Si_2H_6	0.000021	0.000172	0.000821	0.002355	0.008052	0.016427	0.026246
OH_2	0.000014	0.000116	0.000427	0.001137	0.004808	0.013201	0.027606
CH_4	0.000012	0.000102	0.000422	0.001252	0.005677	0.014865	0.028510
NH_3	0.000021	0.000169	0.000594	0.001505	0.005894	0.015187	0.029924
SiH_5^+	0.000037	0.000315	0.001163	0.002960	0.010285	0.021994	0.035606
SiH_3	0.000036	0.000174	0.000784	0.002348	0.009934	0.023161	0.038816

CASE(1/2) values are not listed as this operator differs from CASE(0) by only a constant, thus CASE(1/2) wavefunctions are exactly the same as CASE(0).

As could be expected, the smallest ω value produces the best wavefunction, and very large ω values produce effectively unusable wavefunctions. An ω value of around 0.15 a_0^{-1} seems optimal, as this generally produces a density matrix whose elements are correct to three decimal places. The energy of a traditional calculation using such a matrix will be of approximately microhartree accuracy.

10.3 Chemical Energetics

The effect of the CASE approximation on the properties of the large G2 set [205] has been examined. The large G2 set consists of the original set mentioned earlier, supplemented by 93 atomization energies. The results for a variety of ω values are listed in Table 10.2.

Table 10.2: HF/6-31+G* RMS Errors (kJ/mol) for the properties of the large G2 set using CASE(0) and CASE(1/2)

Property	$\omega = 0.05$	$\omega = 0.10$	$\omega = 0.15$	$\omega = 0.20$	$\omega = 0.30$	$\omega = 0.40$	$\omega = 0.50$
				CASE(0)			
Atomiz. E.	0.58	9.69	44.82	107.63	251.30	360.02	485.67
Ioniz. P.	149.46	305.62	470.02	637.60	944.08	1155.95	1283.58
Elec. A.	4.04	28.07	77.65	145.02	277.84	361.73	390.97
Proton A.	1.96	14.09	40.18	76.31	138.12	119.65	143.00
Total	63.89	131.23	206.00	290.84	463.04	588.15	690.20
				CASE(1/2)			
Atomiz. E.	0.58	9.69	44.82	107.63	251.30	360.02	485.67
Ioniz. P.	9.68	73.11	224.67	468.82	1123.56	1782.59	2271.64
Elec. A.	4.04	28.07	77.65	145.02	277.84	361.73	390.97
Proton A.	1.96	14.09	40.18	76.31	138.12	119.65	143.00
Total	4.40	33.71	106.33	224.68	531.21	825.99	1057.34

As expected, the RMS errors generally increase with increasing ω. Extremely high ω values are nothing short of a disaster. Ionization potentials for CASE(0) are very poor. There is a large improvement in ionization potentials using small ω with the addition of a constant term to the operator; however, they are still the most difficult property for CASE to reproduce. As ω becomes large the addition of a constant is less worthwhile. As can be seen from Figure 10.1, when the operator is more rapidly decaying (that is, high ω), the

background is better approximated without any constant for moderate values of r. Thus, for large ω, CASE(0) is better than CASE(1/2). Caution should be taken in reading too much into the high ω results of Table 10.2, as the errors are of the order of the magnitude of the ionization potential. Notice that the other energy properties are unaffected by the change to the CASE(1/2) operator.

Figure 10.1: Backgrounds and the constants to approximate them for $\omega = 0.1$ and $\omega = 1$.

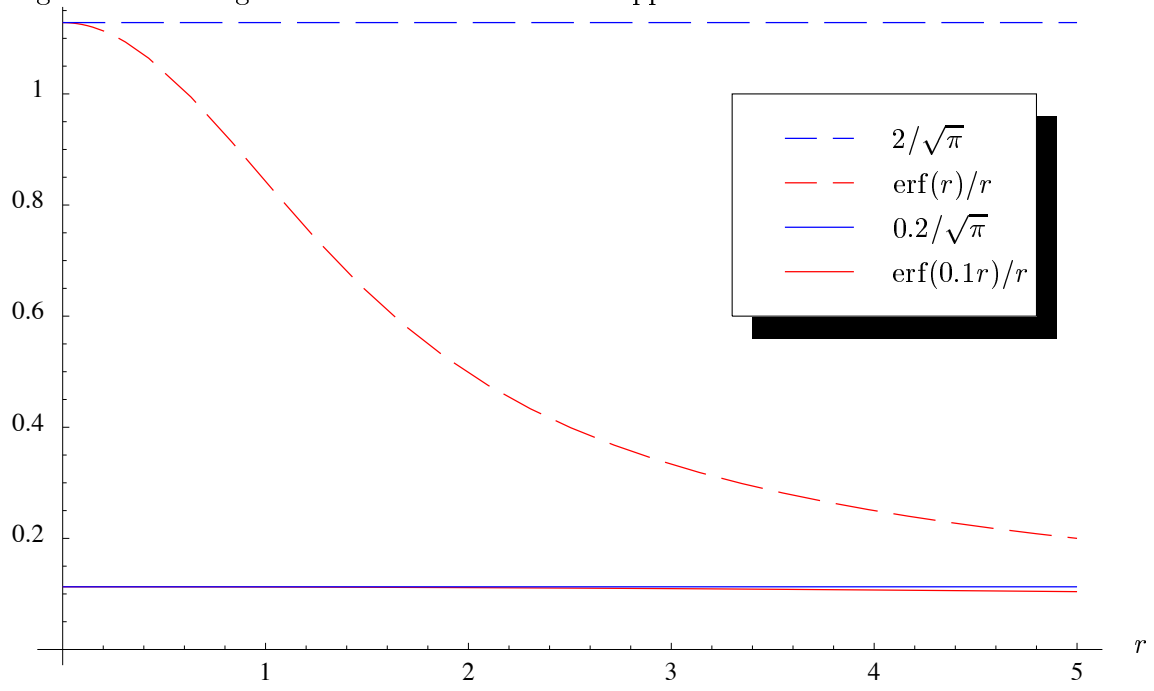


Table 10.3: HFS/6-31+G* RMS Errors (kJ/mol) for the properties of the large G2 set using CASE(0)

Property	$\omega = 0.05$	$\omega = 0.10$	$\omega = 0.15$	$\omega = 0.20$	$\omega = 0.30$	$\omega = 0.40$	$\omega = 0.50$
				CASE(0)			
Atomiz. E.	10.70	33.44	52.10	67.96	124.58	279.05	368.61
Ioniz. P.	71.75	146.32	229.08	317.40	480.09	581.83	622.15
Elec. A.	75.26	135.74	171.70	189.29	212.66	243.43	278.75
Proton A.	1.04	11.47	36.52	72.80	143.96	158.98	117.10
Total	40.78	82.25	121.69	160.51	241.59	348.47	413.10

The effect of CASE on HFS energetic properties is listed in table 10.3. The HFS results follow a very similar pattern to HF with a gradual increase in error. Thus there is no great

effect on the CASE approximation by the replacement of Fock exchange with the Dirac density functional.

10.4 Effect of Basis Set

Table 10.4: HF, $\omega = 0.1\ a_0^{-1}$ RMS Errors (kJ/mol) for the properties of the large G2 set using various basis sets.

Property	STO-3G	3-21G	6-31G	6-31+G(d)
		CASE(0)		
Atomiz. E.	14.44	12.70	12.48	9.69
Ioniz. P.	309.13	307.43	306.73	305.62
Elec. A.	25.66	29.70	29.45	28.07
Proton A.	17.99	16.35	15.46	14.09
Total	133.00	132.21	131.89	131.23
		CASE(1/2)		
Ioniz. P.	73.31	72.47	73.00	73.11
Total	34.75	34.31	34.42	33.70

The same properties of the previous section are examined, with ω fixed at $0.1\ a_0^{-1}$, for a variety of small basis sets. The results, listed in Table 10.4, show a remarkable constancy to changes in basis. Due to this stability with basis, the small STO-3G basis set will be used for the following sections.

10.5 Geometry Optimizations

Geometry optimizations are inherently computationally expensive, especially when many variables are involved. To reduce the number of variables, only the bond lengths have been optimized. This should also allow a more quantitative estimate of the effect of the CASE approximation. However, to convince the reader that the equilibrium of bond angles are affected in a very similar way to that of bond lengths, a simple investigation of the bond angle in ethane (Figure 10.2) has been performed.

The results, listed in Table 10.5, show a progressive deterioration in the bond angle as ω is increased. This deterioration is more rapid for large ω values. This is a pattern very

Figure 10.2: Geometry of Ethane, marking angle to be optimized

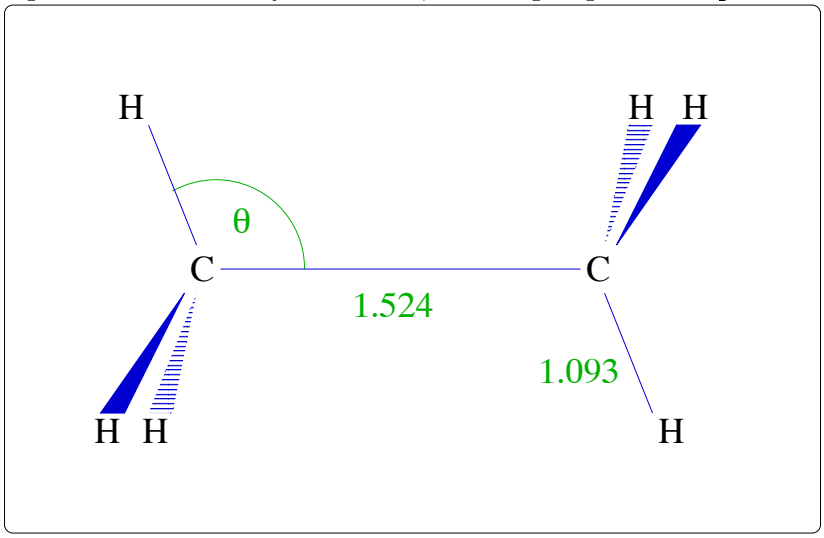


Table 10.5: HF/STO-3G Bond angle (°) of ethane

HF	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.15$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
110.8902	110.8909	110.8898	110.8643	110.7938	110.7727	111.5864	112.8320

similar to that seen for the variety of bond lengths listed in Table 10.6.

The numerous bonds of aspartam have been averaged over each type. So that less information is lost, the standard deviation at each value of ω for the bond is presented in brackets below the average error.

The effect of increasing ω here shows some counter-intuitive results. For a simple system like H_2, removing all Coulomb interactions increases the bond length (and decreases the potential well, as can be seen from the vibrational frequency in the next section). Taking a simplified view, there is an overall Coulomb attraction between two H atoms, so it is not surprising that attenuation reduces the attraction between the atoms. However, for more complex systems, Table 10.6 shows a more complex pattern, and with so many particles interacting it is hard to predict a pattern. The results, though, do clearly show that for several molecules the bond lengthening changes to a bond contraction at high ω.

10.6 Vibrational Frequencies

Vibrational frequencies, which require second derivatives of the potential surface, are a reflection of the quality of the optimized geometry of the molecule. The results described

Table 10.6: HF/STO-3G Bond lengths (Å) for a range of ω values

		CASE bond length - HF bond length						
Bond Type	HF	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.15$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
H_2								
H-H	0.7122	0.0000	0.0004	0.0012	0.0029	0.0072	0.0118	0.0139
BeH								
H-Be	1.3014	0.0001	0.0013	0.0041	0.0083	0.0170	0.0211	0.0273
CH_4								
H-C	1.0830	0.0001	0.0006	0.0019	0.0038	0.0076	0.0080	0.0065
H_2O								
H-O	0.9873	0.0001	0.0006	0.0019	0.0041	0.0099	0.0144	0.0176
CH_3SCH_3								
S-C	1.7989	0.0001	0.0009	-0.0001	-0.0089	-0.0579	-0.1178	-0.1392
H-C	1.0860	0.0001	0.0007	0.0021	0.0042	0.0114	0.0274	0.0606
CH_3COCl								
C-C	1.5370	0.0001	0.0003	-0.0018	-0.0105	-0.0386	-0.0450	0.0045
Cl-C	1.8295	0.0004	0.0022	0.0013	-0.0141	-0.0888	-0.1596	-0.1663
O-C	1.2007	0.0000	0.0004	0.0012	0.0023	0.0040	0.0019	-0.0165
H-C	1.0868	0.0001	0.0007	0.0017	0.0030	0.0069	0.0147	0.0280
Aspartam								
C-C (11)	1.4668	0.0001	0.0002	-0.0019	-0.0063	-0.0161	-0.0140	-0.0001
	(0.0790)	(0.0002)	(0.0015)	(0.0041)	(0.0066)	(0.0070)	(0.0136)	(0.0442)
O-C (6)	1.3046	0.0001	0.0003	0.0010	0.0019	-0.0027	-0.0225	-0.0743
	(0.0913)	(0.0003)	(0.0021)	(0.0057)	(0.0099)	(0.0210)	(0.0316)	(0.0415)
N-C (3)	1.4623	0.0000	-0.0018	-0.0083	-0.0178	-0.0362	-0.0412	-0.0520
	(0.0641)	(0.0001)	(0.0005)	(0.0014)	(0.0088)	(0.0284)	(0.0444)	(0.0586)
H-C (14)	1.0884	0.0001	0.0007	0.0019	0.0041	0.0121	0.0297	0.0779
	(0.0054)	(0.0000)	(0.0002)	(0.0005)	(0.0011)	(0.0027)	(0.0057)	(0.0268)
H-N (4)	1.0423	0.0000	-0.0001	-0.0006	-0.0013	-0.0031	-0.0057	-0.0003
	(0.0045)	(0.0001)	(0.0009)	(0.0019)	(0.0023)	(0.0029)	(0.0069)	(0.0264)

Table 10.6 (continued)

		CASE bond length - HF bond length						
Bond Type	HF	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.15$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
LiF								
F-Li	1.4070	0.0001	0.0015	0.0040	0.0062	-0.0018	-0.0332	-0.0732
CH_3CH_3								
C-C	1.5355	0.0001	0.0006	0.0008	-0.0020	-0.0263	-0.0670	-0.1011
H-C	1.0862	0.0001	0.0007	0.0020	0.0040	0.0087	0.0139	0.0226
PH_4^+								
H-P	1.3824	0.0000	0.0002	0.0001	-0.0018	-0.0141	-0.0308	-0.0322
AlF_3								
F-Al	1.5995	0.0001	0.0012	0.0033	0.0043	-0.0089	-0.0360	-0.0436
C_6H_6								
C-Du	1.3868	0.0001	0.0008	0.0011	-0.0010	-0.0148	-0.0311	-0.0450
H-C	1.0826	0.0001	0.0006	0.0018	0.0040	0.0131	0.0313	0.0685

here support this. Table 10.7 presents the CASE errors in the vibrational frequencies for a range of small molecules.

The results show the now familiar pattern with the gradual decline in accuracy on increasing ω, that becomes a more rapid decline as ω gets larger as well. On the whole, frequencies are one of the more resilient properties examined in this chapter. Considering that HF frequencies are typically only within about 10%, the CASE deterioration (especially for small ω) is quite acceptable. However, frequencies generated with very large ω values are extremely unreliable, and hence results from $\omega > 0.2\, a_0^{-1}$ should be treated with great scepticism.

10.7 Conclusions

The effect of CASE on a variety of molecular properties has been examined in this chapter. Not surprisingly, with all properties investigated, there is an increase in error with increasing ω. Wavefunctions are perhaps the best property for CASE to reproduce. With $\omega = 0.15\, a_0^{-1}$ a wavefunction which reproduces the energy to within a microhartree can be generated. This is significant as it means CASE can be used for the first few cycles of the SCF, allowing fast generation of the wavefunction, and then a final traditional (or KWIK-type) iteration

Table 10.7: HF/STO-3G Vibrational Frequencies (cm^{-1}) for a range of ω values

	HF Frequency - CASE Frequency						
HF Frequency	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.15$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
H_2							
5480.82	1.11	7.96	25.24	57.76	142.71	237.58	299.70
BeH							
2556.23	0.61	8.10	24.85	51.23	122.52	231.05	411.58
LiF							
1303.79	0.63	6.46	16.55	24.47	-5.85	-101.02	-149.48
CH_4							
1675.56	0.25	1.77	5.08	10.21	19.84	10.92	-31.45
1903.60	0.21	1.52	4.48	8.93	18.04	15.98	-9.51
3525.58	0.74	6.11	19.34	38.83	82.47	126.15	213.97
3786.12	0.78	6.49	20.54	41.21	88.66	149.53	297.61
H_2O							
2116.84	0.11	1.08	3.09	5.8	8.5	-5.79	-53.58
4134.35	1.29	8.16	25.47	53.54	126.27	198.13	322.25
4424.22	1.29	8.13	25.27	52.86	123.51	198.89	356.81
AlF_3							
263.57	0.27	1.75	4.38	6.68	1.55	-32.65	-60.53
293.44	0.59	3.38	7.69	11.57	8.56	-10.38	-3.46
774.13	0.24	2.14	5.44	5.92	-12.10	-32.07	27.14
1135.36	0.43	3.68	10.00	13.70	-1.96	4.30	154.22
NH_3							
1276.95	0.02	0.49	1.55	3.21	2.98	-24.38	-105.96
2055.35	0.21	1.59	4.75	9.59	19.98	17.68	-15.07
3846.81	1.17	8.21	25.19	51.82	116.83	180.16	289.41
4139.04	1.11	7.72	23.49	47.62	103.18	160.6	296.26

Table 10.7 (continued)

	HF Frequency - CASE Frequency						
HF Frequency	$\omega = 0.05$	$\omega = 0.1$	$\omega = 0.15$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$
C_6H_6							
477.91	0.12	0.91	3.28	7.20	7.80	-23.09	-63.89
699.87	0.15	1.12	3.85	10.39	39.8	45.6	-4.93
811.00	0.32	2.78	8.58	14.72	12.59	-5.71	-19.09
840.90	0.29	2.24	6.94	13.69	21.32	-0.45	-30.87
1037.65	0.47	3.98	12.49	23.27	34.61	33.18	-115.11
1153.78	-1.54	-0.80	3.40	12.46	48.62	68.85	39.85
1156.23	0.20	-10.-06	-12.12	-8.84	5.73	3.83	3.47
1171.93	0.40	2.39	-4.84	6.86	21.43	19.53	19.17
1190.44	0.56	4.52	13.67	25.37	39.94	38.04	-66.97
1214.51	0.60	4.73	14.03	-11.8	-32.72	-49.07	-42.90
1225.57	0.30	1.83	2.80	-0.74	-21.66	-38.02	-31.84
1371.19	0.25	1.65	3.53	3.69	-3.09	-24.19	-61.3
1376.87	0.17	1.10	1.80	-4.34	-137.21	-244.91	-220.47
1595.04	0.00	0.28	1.76	3.54	3.16	-6.38	-49.00
1772.06	0.40	2.29	3.05	-2.65	-34.65	-63.57	-65.99
1931.65	0.74	4.40	6.44	-3.67	-68.78	-115.26	-63.50
3704.05	1.01	7.79	25.30	60.19	209.21	515.35	1098.67
3722.47	1.04	7.93	25.50	60.23	208.50	515.22	1113.52
3736.14	1.02	7.66	24.11	56.31	193.96	482.78	1049.95
3746.99	1.03	7.60	23.54	54.01	184.66	457.72	982.1

to produce an accurate energy.

To investigate chemical energetics, a more conservative ω value is required. Ionization potentials are the hardest property studied for CASE to reproduce. This is because, as mentioned in Chapter 6, ionization potentials involve removing an electron from a charged system, and thus there is an uncancelled interaction. For small ω values the addition of a constant to the operator is very worthwhile for properties involving an uncancelled interaction, but the usefulness of this additional term decays with increasing ω until, eventually, it becomes disadvantageous.

The error introduced by CASE seems to be practically independent of the basis set used. Bond lengths, at least initially, usually increase with increasing small ω. The decline in accuracy is slower than the energetic properties, and ω values of 0.2 a_0^{-1}, or even a little higher, can be used to generate reasonable geometries.

Chapter 11

Reintroducing the Background[1]

"The great tragedy of science — the slaying of a beautiful hypothesis by an ugly fact." - Thomas H. Huxley

11.1 Introduction

While CASE is a useful method in its own right, it may become more important as the first term of an expansion allowing computation of the Coulomb energy in only $O(N)$ work. At the moment there is no completely satisfactory way of wholly including the background (in $O(N)$ work), but this is an area of active research [206]. This Chapter presents just a taste of what may be available in the future. The method presented is not without serious deficiencies, but it is included to convince the reader that CASE can be used as a starting point for more exact (and equally fast) methods.

11.2 An Expression for E_L

Applying the KWIK separator to the electron-electron Coulomb term partitions the energy, as before, into short- and long-range pieces

$$E_J = \frac{1}{2}\left\langle \rho \left| \frac{1}{r_{12}} \right| \rho \right\rangle \tag{11.1}$$

$$= \frac{1}{2}\left\langle \rho \left| \frac{\mathrm{erfc}(\omega r_{12})}{r_{12}} \right| \rho \right\rangle + \frac{1}{2}\left\langle \rho \left| \frac{\mathrm{erf}(\omega r_{12})}{r_{12}} \right| \rho \right\rangle \tag{11.2}$$

$$= E_S + E_L. \tag{11.3}$$

[1] The work described in this chapter has been carried out in collaboration with Dr Aaron Lee

The short-range is calculated as in Chapter 7. But, instead of neglecting the long-range energy, it is estimated from the systems electronic multipole moments.

First, the long-range operator is expressed as its one-dimensional spherical Bessel transform

$$\frac{\mathrm{erf}(\omega r_{12})}{r_{12}} = \frac{2}{\pi}\int_0^\infty e^{-k^2/4\omega^2}\, j_0(kr_{12})\, dk, \tag{11.4}$$

where

$$j_0(x) = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots \tag{11.5}$$

Interacting this operator with the density produces an expression for the long-range energy

$$E_L = \frac{2}{\pi}\int_0^\infty e^{-k^2/4\omega^2} G(k)\, dk, \tag{11.6}$$

where $G(k)$, termed the "fingerprint" of the density, is defined as

$$G(k) = \frac{1}{2}\left\langle \rho \left| j_0(kr_{12}) \right| \rho \right\rangle \tag{11.7}$$

$$= W_0 - W_2\frac{k^2}{3!} + W_4\frac{k^4}{5!} - W_6\frac{k^6}{7!} + \dots, \tag{11.8}$$

with $W_{2\lambda}$ defined as

$$W_{2\lambda} = \frac{1}{2}\left\langle \rho \left| r_{12}^{2\lambda} \right| \rho \right\rangle. \tag{11.9}$$

The advantage of all this rearrangement is that Gill [207] has shown that the $W_{2\lambda}$ can be calculated in only $O(N)$ work from the multipole moments of the system. For example (using the Einstein summation convention),

$$W_0 = \frac{1}{2}\mathbf{M}_0^2 \tag{11.10}$$

$$W_2 = \mathbf{M}_0\mathbf{M}_{ii} - \mathbf{M}_i\mathbf{M}_i \tag{11.11}$$

$$W_4 = \mathbf{M}_0\mathbf{M}_{iijj} - 4\mathbf{M}_i\mathbf{M}_{ijj} + \mathbf{M}_{ii}\mathbf{M}_{jj} + 2\mathbf{M}_{ij}\mathbf{M}_{ij} \tag{11.12}$$

$$W_6 = \mathbf{M}_0\mathbf{M}_{iijjkk} - 6\mathbf{M}_i\mathbf{M}_{ijjkk} + 3\mathbf{M}_{ii}\mathbf{M}_{jjkk} + 12\mathbf{M}_{ij}\mathbf{M}_{ijkk} - 6\mathbf{M}_{iij}\mathbf{M}_{jkk} - 4\mathbf{M}_{ijk}\mathbf{M}_{ijk} \tag{11.13}$$

where $\mathbf{M}$ are the Cartesian multipole moment tensors of the electron density, for example

$$\mathbf{M}_0 = \int \rho(\mathbf{r})\, d\mathbf{r} \tag{11.14}$$

$$\mathbf{M}_i = \int r_i\rho(\mathbf{r})\, d\mathbf{r} \tag{11.15}$$

$$\mathbf{M}_{ij} = \int r_i r_j\rho(\mathbf{r})\, d\mathbf{r} \tag{11.16}$$

with r_i representing the ith Cartesian component of $\mathbf{r}$. These moments are trivially computed in $O(N)$ work.

11.3 A Gaussian Basis for $G(k)$

Thus, in order to calculate the long-range energy in $O(N)$ work all that needs to be done is generate the electronic multipole moments, combine to form the $W_{2\lambda}$ and perform the integration in equation (11.6). There is however, a problem: the $W_{2\lambda}$ increase factorially, making the series in equation (11.8) only asymptotically convergent.

Equation (11.8) shows that the $W_{2\lambda}$ provide the derivatives of $G(k)$ at $k = 0$. These can be used to generate an approximation to $G(k)$ which allows the integration in equation (11.6) to be carried out. Just what basis to use to represent $G(k)$ is an open question. A suitable basis must allow analytic integration of $G(k)$ and provide an accurate representation of $G(k)$, using as few derivatives as possible. By using a sum of m Gaussians,

$$G(k) \approx G_m(k) = \sum_{i=1}^{m} A_{im} e^{-\alpha_{im} k^2} \tag{11.17}$$

the approximate function's Taylor expansion matches that of $G(k)$ up to order k^{4m-2}. Obviously, the more Gaussians used, the more moments have to be calculated, thus increasing the cost of the algorithm.

The coefficients and exponents of the Gaussians are found by equating the first $2m$ terms of the Taylor expansions for $G(k)$ and $G_m(k)$, leading to a set of linear equations which are solved for polynomial expressions of the exponents. The exponents are then found by finding the roots of a polynomial of order m.

Inserting this sum of Gaussians into equation (11.6) and performing the integration produces the rather simple expression for the long-range energy

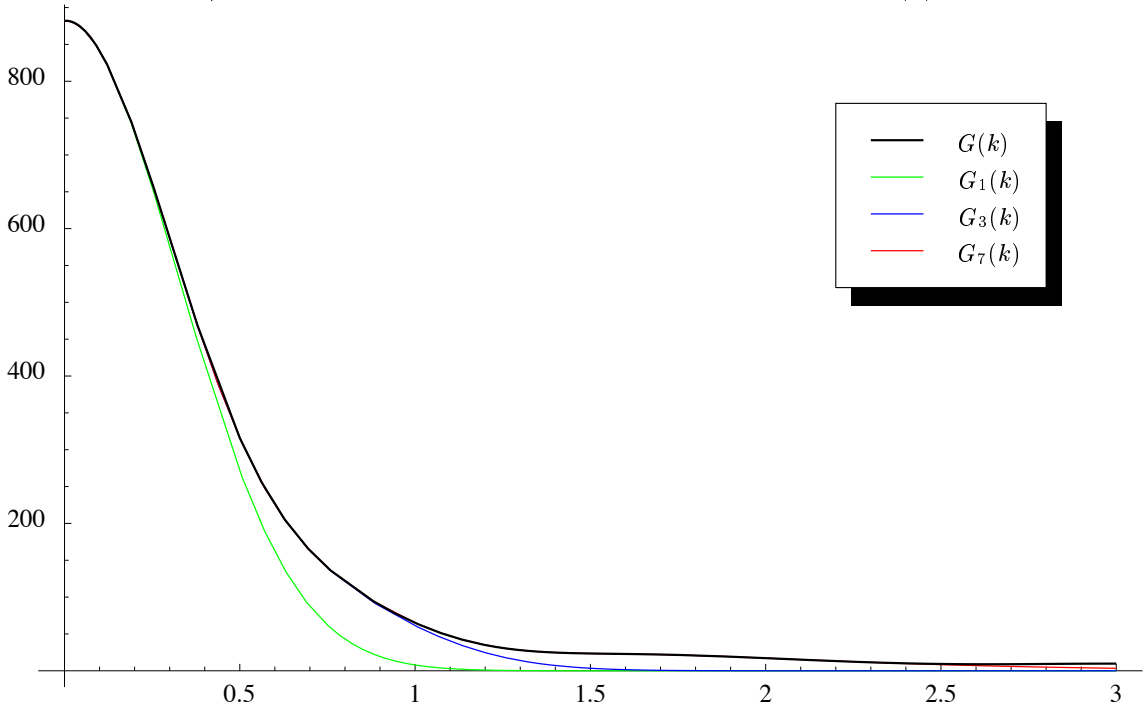
$$E_L = \frac{2\omega}{\sqrt{\pi}} \sum_{i=1}^{m} \frac{A_{im}}{\sqrt{4\omega^2 \alpha_{im} + 1}}. \tag{11.18}$$

The higher the value of m in $G_m(k)$, the more accurately $G_m(k)$ represents $G(k)$, providing a more accurate estimate for E_L.

11.4 The Ugly Fact

Dr Aaron Lee has modified the Q-Chem program to calculate the two-electron integrals over the $j_0(kr_{12})$ operator, allowing the $G(k)$ curves to be plotted from an arbitrary density. This curve for a pentane density is plotted in Figure 11.1. Superimposed on this are the first few $G_m(k)$ approximations to it. Each successive $G_m(k)$ follows the exact $G(k)$ for longer, before decaying to zero. G_2, G_4, G_5 and G_6 have all been omitted because they possess an

Figure 11.1: HF/6-31+G* fingerprint of pentane and the first few $G_m(k)$ approximations



exponential with a positive exponent. These curves follow $G(k)$ before rising to infinity. This will, unless ω is small enough, produce a nonsensical result (remembering that this curve is integrated)! This is a major draw-back of using a sum of Gaussians as a basis function to represent $G(k)$. Several basis functions have been tried, and no satisfactory answer has yet been found. A drawback is that the basis functions chosen have to be very efficient in reproducing $G(k)$, as numerical errors become a problem for large m values.

11.5 $G_m(k)$ Chemistry

Table 11.1: CASE Coulomb energy error for pentane (HF/6-31+G*)

	$E_J - E_L^m$ (hartrees)						
m	$\omega = 0.1$	$\omega = 0.2$	$\omega = 0.3$	$\omega = 0.4$	$\omega = 0.5$	$\omega = 1.0$	$\omega = 2.0$
1	0.1646	2.4500	7.5051	13.2132	18.3929	34.8348	48.4376
3	0.0000	0.0035	0.0846	0.4548	1.2316	7.6881	17.5541
7	0.0000	0.0000	0.0000	0.0000	0.0006	0.3954	3.9760
8	0.0000	0.0000	0.0000	-0.0001	0.0001	0.3634	3.8471

As the purpose of this chapter is only to convince the reader of the viability of an $O(N)$ correction for the background (and because no better basis has yet been found), Gaussians will be used as a basis, and calculations producing positive exponents will simply be admitted. Table 11.1 presents the energies obtained from each of the $G_m(k)$ approximations with varying ω. As ω increases, the amount of long-range work increases, and so the $G_m(k)$ approximation becomes more drastic. It should be noted, though, that a large proportion of the CASE Coulomb energy error can be regained with only a few Gaussians.

Table 11.2: Chemical energetics RMS errors (kJ/mol) for $G_m(k)$ approximations

Property	G_1	G_2	G_3	G_4	G_5
32 Atomiz. E.	3002.5	410.3	80.5	64.2	14.0
30 Ioniz. E.	1900.8	380.5	71.6	18.1	3.5
19 Electron A.	3553.4	1306.5	329.3	88.6	11.0
7 Proton A.	624.3	324.6	106.4	13.2	2.8

Table 11.2 shows the (HF/6-31+G*) atomization energy, ionization potential, electron affinity and proton affinity RMS errors using $G_1(k)$ to $G_5(k)$ for a variety of small molecules using $\omega = 1/2$. These are the properties of the G2 set for which there were no problematic Gaussian exponents. Note that this is a large value of ω, with small m values. Even with only five Gaussians most of the chemistry is reproduced.

11.6 Conclusions

The algorithm presented in this chapter shows a way to generate a correction for the neglected part of a CASE calculation in only $O(N)$ work, by representing the long-range energy in terms of electronic multipole moments of the system. The most straightforward derivation generates a representation for the energy via a series that is only asymptotically convergent. To overcome this, the series can be approximated by basis functions using knowledge of the derivatives at the origin. Choosing just what type of basis function, however, is still an unsolved problem. A basis function of Gaussians has been tried here and has shown promising results. Most of the neglected energy can be recovered with only a few of these basis functions (and hence only small order moments of the system). Gaussians have the serious drawback of sometimes producing a nonsensical energy through the need for positive exponents to accurately represent $G(k)$.

The most encouraging point from this chapter, though, is that even if the CASE approximation proves to be unworkable in the future, the code written and methods developed in this thesis will still be useful as they will provide the short-range energy for an $O(N)$ Coulomb algorithm.

Chapter 12

Concluding Remarks

Novel methods for the application of Quantum Chemistry to large molecules has been examined in this thesis. The first piece of research presented is a new way of examining the quality of density functionals by partitioning the electron density. When applied to the LSDA functional, one of the few ab initio functionals in use today, some major drawbacks were highlighted. To make advances in DFT a more pragmatic approach is required, and functionals of a more empirical nature should be used. This idea is taken to its logical conclusion with the introduction of an entirely empirical density functional, EDF1. EDF1 has been specifically designed for a small basis set (6-31+G*) allowing application to large molecules. EDF1 is more accurate than the currently most popular functional, B3LYP, for this basis set. One of the interesting side effects of this research was that the addition of Fock exchange was of no benefit. This is advantageous computationally as the Fock term is very expensive.

The most time-consuming aspect of quantum chemical calculations is the generation of the two-electron integrals. The time scaling of this has been tackled in Chapter 5. Specifically, the scaling with the contraction of the Gaussian basis set has been removed from the $O(N^2)$ part of the code, at the cost of a little extra (and mainly insignificant) $O(N)$ work.

The CASE approximation has been introduced in Chapter 6. CASE assumes that neutral distributions of charge a large distance away do not interact with a charge distribution. This is achieved by attenuating the Coulomb operator in a smooth manner, forcing the operator to decay to insignificance much faster. Thus the number of significant interactions grows as only $O(N)$, rather than the traditional $O(N^2)$. With the introduction of a boxing scheme the required two-electron integrals can be computed in only $O(N)$ work, providing massive computational savings.

CASE can produce reasonable results, if the attenuation is not too great. It may be especially useful in determining wavefunctions, which can then be used with the full Coulomb operator to give accurate properties. CASE can also be viewed as a first order approximation, upon which higher orders can be made. The correction for the background, although not yet fully developed, is currently a promising route to a very fast $O(N)$ Coulomb method. Research in this area is being continued by Dr Aaron Lee and Mr Nikhil Nair.

Bibliography

[1] L. de Broglie, Ann. Physik. 3 (1925) 22.

[2] C. J. Davisson, Congres Intern. Elec. 2 (1932) 593.

[3] E. Schrödinger, Ann. Physik. 79 (1926) 361.

[4] W. Heisenberg, Z. Physik. 43 (1927) 172.

[5] M. Born, Z. Physik. 38 (1926) 803.

[6] M. Born, J. R. Oppenheimer, Ann. Physik. 84 (1927) 457.

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

[8] E. A. Hylleraas, Z. Physik. 54 (1929) 347.

[9] E. A. Hylleraas, Z. Physik. 65 (1930) 209.

[10] D. R. Hartree, Proc. Cam. Phil. Soc. 24 (1928) 89.

[11] D. R. Hartree, Proc. Cam. Phil. Soc. 24 (1928) 111.

[12] D. R. Hartree, Proc. Cam. Phil. Soc. 24 (1928) 426.

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

[14] W. Pauli, Z. Physik. 31 (1925) 765.

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

[16] J. C. Slater, Phys. Rev. 35 (1930) 509.

[17] A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York, first (revised) edition, (1989).

[18] M. Wolfsberg, L. Helmholtz, J. Chem. Phys. 20 (1952) 837.

[19] B. G. Johnson, unpublished.

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

[21] G. G. Hall, Proc. Roy. Soc. (London) A205 (1951) 541.

[22] C. C. J. Roothaan, Rev. Mod. Phys 32 (1960) 179.

[23] J. S. Binkley, J. A. Pople, P. A. Dobosh, Mol. Phys. 28 (1974) 1423.

[24] J. A. Pople, R. K. Nesbet, J. Chem. Phys. 22 (1954) 571.

[25] P. M. W. Gill, L. Radom, Chem. Phys. Lett. 132 (1986) 16.

[26] R. H. Nobes, J. A. Pople, L. Radom, N. C. Handy, P. J. Knowles, Chem. Phys. Lett. 138 (1987) 481.

[27] P. J. Knowles, N. C. Handy, J. Phys. Chem. 92 (1988) 3097.

[28] P. M. W. Gill, J. A. Pople, L. Radom, J. Chem. Phys. 89 (1988) 7307.

[29] D. L. Strout, G. E. Scuseria, J. Chem. Phys. 102 (1995) 8448.

[30] P. O. Löwdin, Adv. Chem. Phys. 2 (1959) 207.

[31] W. J. Hehre, L. Radom, P. v. R. Schleyer, J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, (1986).

[32] J. A. Pople, M. Head-Gordon, K. Raghavachari, J. Chem. Phys. 87 (1987) 5968.

[33] S. F. Boys, Proc. Roy. Soc. (London) A201 (1950) 125.

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

[35] L. Brillouin, Actualities Sci. Ind. 71 (1934) 159.

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

[37] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, Chem. Phys. Lett. 157 (1989) 479.

[38] K. Raghavachari, J. A. Pople, E. S. Replogle, M. Head-Gordon, J. Phys. Chem. 94 (1990) 5579.

[39] J. Čižek, J. Chem. Phys. 45 (1966) 4256.

[40] J. Čižek, Adv. Chem. Phys. 14 (1969) 35.

[41] G. D. Purvis, R. J. Bartlett, J. Chem. Phys. 76 (1982) 1910.

[42] K. Raghavachari, J. B. Anderson, J. Phys. Chem. 100 (1996) 12960.

[43] M. Head-Gordon, J. Phys. Chem. 100 (1996) 13213.

[44] N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, G. W. Trucks, Chem. Phys. Lett. 164 (1989) 185.

[45] K. A. Brueckner, Phys. Rev. 96 (1954) 508.

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

[47] J. A. Pople, J. S. Binkley, R. Seeger, Int. J. Quant. Chem. 10 (1976) 1.

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

[49] O. Christiansen, J. Olsen, P. Jorgensen, H. Koch, P. A. Malmqvist, Chem. Phys. Lett. 261 (1996) 369.

[50] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, L. A. Curtiss, J. Chem. Phys. 90 (1989) 5662.

[51] L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, J. A. Pople, J. Chem. Phys. 93 (1990) 2537.

[52] L. A. Curtiss, K. Raghavachari, G. W. Trucks, J. A. Pople, J. Chem. Phys. 94 (1991) 7221.

[53] L. A. Curtiss, K. Raghavachari, J. A. Pople, J. Chem. Phys. 103 (1995) 4192.

[54] A. Nicolaides, L. Radom, J. Phys. Chem. 98 (1994) 3092.

[55] J. C. Slater, Phys. Rev. 36 (1930) 57.

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

[57] E. R. Davidson, D. Feller, Chem. Rev. 86 (1988) 681.

[58] W. J. Hehre, R. F. Stewart, J. A. Pople, J. Chem. Phys. 51 (1969) 2657.

[59] E. Clementi, IBM J. Res and Dev. 9 (1965) 2.

[60] T. H. Dunning, J. Chem. Phys. 53 (1970) 2823.

[61] R. Ditchfield, W. J. Hehre, J. A. Pople, J. Chem. Phys. 54 (1971) 724.

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

[63] P. C. Hariharan, J. A. Pople, Theor. Chim. Acta. 28 (1973) 213.

[64] T. Kato, Commun. Pure Appl. Math. 10 (1957) 151.

[65] R. T. Pack, W. B. Brown, J. Chem. Phys. 45 (1966) 556.

[66] W. A. Bingel, Z. Naturforsch. Teil A 18 (1963) 1249.

[67] V. A. Rassolov, D. M. Chipman, J. Chem. Phys. 104 (1996) 9908.

[68] W. Klopper, W. Kutzelnigg, J. Chem. Phys. 96 (1991) 2020.

[69] A. Rich, N. Davidson, editors, *Structural Chemistry and Biology*, 753–760, W. H. Freeman, San Fransisco, (1968).

[70] P. Hohenberg, W. Kohn, Phys. Rev. B136 (1964) 864.

[71] R. M. Dreizler, E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem*, Springer-Verlag, (1990).

[72] M. Levy, Phys. Rev. A. 26 (1982) 1200.

[73] E. H. Lieb, Int. J. Quant. Chem. 24 (1983) 243.

[74] M. Levy, Proc. Natl. acad. Sci. USA 76 (1979) 6062.

[75] R. McWeeny, Rev. Mod. Phys 32 (1960) 335.

[76] J. C. Slater, Phys. Rev. 81 (1951) 385.

[77] C.-O. Almbladh, A. C. Pedroza, Phys. Rev. A. 29 (1984) 2322.

[78] L. H. Thomas, Proc. Cam. Phil. Soc. 23 (1927) 542.

[79] E. Fermi, Rend. Accad. Lincei 6 (1927) 602.

[80] E. Fermi, Z. Physik. 48 (1928) 73.

[81] E. Fermi, Rend. Accad. Lincei 7 (1928) 342.

[82] P. A. M. Dirac, Proc. Cam. Phil. Soc. 26 (1930) 376.

[83] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, (1989).

[84] A. D. Becke, J. Chem. Phys. 85 (1986) 7184.

[85] S. J. Vosko, L. Wilk, M. Nusair, Can. J. Phys. 58 (1980) 1200.

[86] D. M. Ceperly, B. J. Alder, Phys. Rev. Lett. 45 (1980) 566.

[87] J. P. Perdew, Phys. Rev. B. 33 (1986) 8822.

[88] C. F. von Weizsacker, Z. Physik. 96 (1935) 431.

[89] C. H. Hodges, Can. J. Phys. 51 (1973) 1428.

[90] D. R. Murphy, Phys. Rev. A. 24 (1981) 1682.

[91] P. J. Marcus, J. F. Janak, A. R. Williams, editors, *Computational Methods in Bond Theory*, 458–468, Plenum, New York, (1971).

[92] L. Kleinman, Phys. Rev. B. 30 (1984) 2223.

[93] W. Kohn, L. J. Sham, Phys. Rev. A140 (1965) 1133.

[94] T. Koopmans, Physica 1 (1934) 104.

[95] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, 149, Oxford University Press, New York, (1989).

[96] J. A. Pople, P. M. W. Gill, N. C. Handy, Int. J. Quant. Chem. 56 (1995) 303.

[97] R. M. Driezler, J. da Providencia, editors, *Density Functional Methods in Physics*, 11–30, Plenum, New York, (1985).

[98] P. M. W. Gill, B. G. Johnson, J. A. Pople, M. J. Frisch, Chem. Phys. Lett. 197 (1992) 499.

[99] N. C. Handy, C. W. Murray, R. D. Amos, J. Phys. Chem. 97 (1993) 4392.

[100] B. G. Johnson, P. M. W. Gill, J. A. Pople, J. Chem. Phys. 98 (1993) 5612.

[101] M. M. Morrell, R. G. Parr, M. Levy, J. Chem. Phys. 62 (1975) 549.

[102] A. D. Becke, Phys. Rev. A. 38 (1988) 3098.

[103] C. Lee, Z. Zhou, Phys. Rev. A. 44 (1991) 1536.

[104] E. Engel, J. A. Chevary, L. D. Macdonald, S. H. Vosko, Z. Phys. D 23 (1992) 7.

[105] J. B. Krieger, Y. Li, G. J. Iafrate, Phys. Rev. A. 45 (1992) 101.

[106] P. M. W. Gill, Mol. Phys. 89 (1996) 433.

[107] P. Ziesche, H. Eschrig, editors, *Electronic Structure of Solids '91*, Akademie Verlag, Berlin, (1991).

[108] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, C. Fiolhais, Phys. Rev. B. 46 (1992) 6671.

[109] A. D. Becke, J. Chem. Phys. 97 (1992) 9173.

[110] R. Colle, O. Salvetti, Theor. Chim. Acta. 37 (1975) 329.

[111] C. Lee, W. Yang, R. G. Parr, Phys. Rev. B. 37 (1988) 785.

[112] B. Miehlich, A. Savin, H. Stoll, H. Press, Chem. Phys. Lett. 157 (1989) 200.

[113] B. G. Johnson, G. A. Gonzalez, P. M. W. Gill, J. A. Pople, Chem. Phys. Lett. 221 (1994) 100.

[114] E. P. Wigner, Trans. Faraday Soc. 34 (1938) 678.

[115] L. C. Wilson, Chem. Phys. 181 (1994) 337.

[116] B. Pullman, R. Parr, editors, *The new world of quantum chemistry*, 3, Reidel, Dordrecht, (1976).

[117] G. Brual Jr., S. M. Rothstein, J. Chem. Phys. 69 (1978) 1177.

[118] P. A. Stewart, P. M. W. Gill, J. Chem. Soc. Faraday Trans. 91 (1995) 4337.

[119] A. D. Becke, J. Chem. Phys. 98 (1993) 5648.

[120] A. D. Becke, J. Chem. Phys. 98 (1993) 1372.

[121] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, J. Phys. Chem. 98 (1994) 11623.

[122] V. Barone, Chem. Phys. Lett. 226 (1994) 392.

[123] J. Andzelm, E. Wimmer, J. Chem. Phys. 96 (1992) 1280.

[124] S. F. Boys, P. Rajagopal, Adv. Quant. Chem. 2 (1965) 1.

[125] B. G. Johnson, P. M. W. Gill, M. Head-Gordon, C. A. White, J. Baker, D. R. Maurice, T. R. Adams, J. Kong, M. Challacombe, E. Schwegler, M. Oumi, C. Ochsenfeld, N. Ishikawa, J. Florian, R. D. Adamson, J. P. Dombroski, R. L. Graham, A. Warshel, Q-Chem, Version 1.1, Q-Chem, Inc., Pittsburgh, PA (1997).

[126] A. D. Becke, J. Chem. Phys. 88 (1988) 2547.

[127] J. C. Slater, J. Chem. Phys. 41 (1964) 3199.

[128] J. C. Slater, *Quantum Theory of Molecules and Solids*, volume 2, McGraw-Hill, New York, (1965).

[129] W. Yang, J. Chem. Phys. 94 (1991) 1208.

[130] G. te Velde, E. J. Baerends, J. Comput. Phys. 99 (1992) 84.

[131] O. Treutler, R. Alrichs, J. Chem. Phys. 102 (1995) 346.

[132] M. E. Mura, P. J. Knowles, J. Chem. Phys. 104 (1996) 9848.

[133] C. W. Murray, N. C. Handy, G. J. Laming, Mol. Phys. 78 (1993) 997.

[134] S. F. Boys, N. C. Handy, Proc. R. Soc. A. 311 (1969) 309.

[135] N. C. Handy, S. F. Boys, Theor. Chim. Acta. 31 (1973) 195.

[136] A. D. McLaren, Math. Comput. 17 (1963) 361.

[137] A. H. Stroud, *Approximate Calculation of Multiple Integrals*, Prentice-Hall, Englewood Cliffs, (1971).

[138] S. L. Sobolev, Sibirsk. Mat. Zh. 3 (1962) 769.

[139] S. I. Konyaev, Mat. Zametki. 25 (1979) 629.

[140] V. I. Lebedev, Zh. Vychisl. Mat. mat. Fiz. 15 (1975) 48.

[141] V. I. Lebedev, Zh. Vychisl. Mat. mat. Fiz. 16 (1976) 293.

[142] V. I. Lebedev, Sibirsk. Mat. Zh. 18 (1977) 132.

[143] V. I. Lebedev, A. L. Skorokhodov, Russian Acad. Sci. Dokl. Math. 45 (1992) 587.

[144] B. G. Johnson, M. J. Frisch, Chem. Phys. Lett. 216 (1993) 133.

[145] R. S. Jones, J. W. Mintmire, B. I. Dunlap, Int. J. Quant. Chem. 22 (1988) 77.

[146] B. G. Johnson, P. M. W. Gill, J. A. Pople, Chem. Phys. Lett. 220 (1994) 377.

[147] P. M. W. Gill, B. G. Johnson, J. A. Pople, Chem. Phys. Lett. 209 (1993) 506.

[148] B. G. Johnson, *Development, implementation and performance of efficient methodologies for density functional calculations*, Ph.D. thesis, Carnegie Mellon University, (December 1993).

[149] R. E. Stratman, G. E. Scuseria, M. J. Frisch, Chem. Phys. Lett. 257 (1996) 213.

[150] J. M. Pérez-Jordá, W. Yang, Chem. Phys. Lett. 241 (1995) 469.

[151] J. A. Pople, P. M. W. Gill, B. G. Johnson, Chem. Phys. Lett. 199 (1992) 557.

[152] H. Stoll, C. M. E. Pavlidou, H. Preuss, Theor. Chim. Acta. 49 (1978) 143.

[153] H. Stoll, E. Golka, H. Preuss, Theor. Chim. Acta. 55 (1980) 29.

[154] J. P. Perdew, E. R. McMullen, A. Zunger, Phys. Rev. A. 23 (1981) 2785.

[155] A. D. Becke, J. Chem. Phys. 96 (1992) 2155.

[156] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, C. Fischer, Phys. Rev. A. 44 (1991) 7071.

[157] C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon, Chem. Phys. Lett. 253 (1996) 268.

[158] M. Challacombe, E. Schwegler, J. Chem. Phys. 106 (1997) 5526.

[159] P. M. W. Gill, Adv. Quant. Chem. 25 (1994) 141.

[160] P. M. W. Gill, J. A. Pople, Int. J. Quant. Chem. 40 (1991) 753.

[161] J. A. Pople, W. J. Hehre, J. Comput. Phys. 27 (1978) 161.

[162] M. Dupuis, J. Rys, H. F. King, J. Chem. Phys. 65 (1976) 111.

[163] L. E. McMurchie, E. R. Davidson, J. Comput. Phys. 26 (1978) 218.

[164] S. Obara, A. Saika, J. Chem. Phys. 84 (1986) 3963.

[165] M. Head-Gordon, J. A. Pople, J. Chem. Phys. 89 (1988) 5777.

[166] S. Ten-no, Chem. Phys. Lett. 211 (1993) 259.

[167] P. M. W. Gill, B. G. Johnson, J. A. Pople, Int. J. Quant. Chem. 40 (1991) 745.

[168] J. Rys, M. Dupuis, H. F. King, J. Comp. Chem. 4 (1983) 154.

[169] T. P. Hamilton, H. F. Schaefer III, Chem. Phys. 150 (1991) 163.

[170] R. Lindh, U. Ryu, B. Liu, J. Chem. Phys. 95 (1991) 5889.

[171] T. Helgaker, P. R. Taylor, Theor. Chim. Acta. 83 (1992) 177.

[172] K. Ishida, Int. J. Quant. Chem. 59 (1996) 209.

[173] B. G. Johnson, P. M. W. Gill, J. A. Pople, Int. J. Quant. Chem. 40 (1991) 809.

[174] U. Ryu, Y. S. Lee, R. Lindh, Chem. Phys. Lett. 185 (1991) 562.

[175] B. G. Johnson, P. M. W. Gill, J. A. Pople, Chem. Phys. Lett. 206 (1993) 229.

[176] T. R. Adams, R. D. Adamson, P. M. W. Gill, J. Chem. Phys. 107 (1997) 124.

[177] E. Clementi, in *Robert A. Welch Foundation Conferences on Chemical Research*, volume XVI, Houston, (1972).

[178] L. Greengard, V. Rokhlin, J. Comput. Phys. 73 (1987) 325.

[179] L. Greengard, *The Rapid Evaluation of Potential Fields in Particle Systems*, MIT, Cambridge, (1987).

[180] C. A. White, M. Head-Gordon, J. Chem. Phys. 101 (1994) 6593.

[181] L. Greengard, Science 265 (1994) 909.

[182] C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon, Chem. Phys. Lett. 230 (1994) 8.

[183] C. A. White, M. Head-Gordon, Chem. Phys. Lett. 257 (1996) 647.

[184] C. A. White, M. Head-Gordon, J. Chem. Phys. 105 (1996) 5061.

[185] E. Schwegler, M. Challacombe, J. Chem. Phys. 105 (1996) 2726.

[186] W. Kohn, Phys. Rev. 115 (1959) 809.

[187] W. Kohn, Int. J. Quant. Chem. 56 (95) 229.

[188] J. P. Dombroski, S. W. Taylor, P. M. W. Gill, J. Phys. Chem. 100 (1996) 6272.

[189] J. P. Dombroski, *The KWIK Algorithm for Coulomb Interactions and its Applications*, Ph.D. thesis, Massey University, (1997).

[190] A. M. Lee, S. W. Taylor, J. P. Dombroski, P. M. W. Gill, Phys. Rev. A. 55 (1997) 3233.

[191] A. M. Lee, Chem. Phys. Lett. 292 (1998) 172.

[192] W. Kohn, Chem. Phys. Lett. 208 (1993) 167.

[193] S. Wolfram, Mathematica, Wolfram Research Inc., (1998).

[194] S. Sæbø, P. Pulay, Chem. Phys. Lett. 113 (1985) 13.

[195] J. Almlöf, Chem. Phys. Lett. 181 (1991) 319.

[196] M. Hässer, J. Almlöf, J. Chem. Phys. 96 (1992) 489.

[197] D. C. Rapaport, Comp. Phys. Rep. 9 (1988) 1.

[198] C. G. Lambert, T. A. Darden, J. A. Board Jr., J. Comput. Phys. 126 (1996) 274.

[199] R. W. Hockney, J. W. Eastwood, *Computer simulation using particles*, IOP Publishing, Bristol, student edition, (1988).

[200] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge University Press, Cambridge, second edition, (1992).

[201] M. P. Allen, D. J. Tildesley, *Computer simulation of liquids*, Clarendon Press, Oxford, (1987).

[202] M. Berkowitz, Chem. Phys. Lett. 129 (1986) 486.

[203] A. D. Becke, J. Chem. Phys. 104 (1996) 1040.

[204] Y. Sakamoto, J. Chem. Phys. 28 (1958) 164.

[205] L. A. Curtis, K. Raghavachari, P. C. Redfern, J. A. Pople, J. Chem. Phys. 106 (1997) 1063.

[206] A. M. Lee, R. D. Adamson, P. M. W. Gill, to be published.

[207] P. M. W. Gill, Chem. Phys. Lett. 270 (1997) 193.