One Dimensional Chemistry and the Generalised Local Density Approximation

Caleb James Ball

February 2017

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Declaration

The contents of this thesis have not been submitted for the award of any other degree or diploma. Except where otherwise noted, the material is my own original work.

_____

Caleb James Ball

20 February, 2017

Acknowledgements

Writing a Ph.D. thesis leaves one with many people to thank. Here are a few most important to me:

Thanks Peter. After supervising my honours thesis you were brave enough to take me on for a Ph.D. project. You've taught me a great deal about many different things, but most of all I thank you for the immense wisdom you've shared over the past years. I hope that when I look for a bug in some code in the future I will remember all the times you told me to go back to the simplest case.

Thanks Titou. You've spent so much time with me that I consider your supervision on equal footing with Peter's. Many thanks for taking that burden on. I know I caused you a lot of angst at various times, and I'm grateful that you kept putting up with me. Hopefully we can share another beer sometime in the future, because at the end of the day I'm going to miss your company.

Thanks Mum and Dad. I wouldn't be here without you, in more ways than one! You've given me more support than I could have asked for and it's opened up so many opportunities for me. I'm sorry I chose not to go to music school, hopefully this makes up for it!

Thanks Fiorella. Thanks for sticking with me through thick and thin. You were always there to cheer me up when I felt like things were looking pretty bleak, which probably helped a lot more than I thought. I promise I'll make our next adventure a little smoother.

And thanks to everyone else. Thanks to the other members of the Gill group, particularly Andrew, Giuseppe, Marat and Simon. Thanks to all my friends who've watched as it all happened. Life would have been a lot duller had you not all been there to brighten things up!

And thanks to the institutions who provided me funding: the Australian Government and the ANU Research School of Chemistry for their scholarships as well as the National Computing Infrastructure for their grants of supercomputer time.

Abstract

In this thesis we explore two distinct topics: a unique model of one dimensional chemistry and the development of the generalised local density approximation.

The specific one dimensional system we study is one in which three dimensional particles, both electrons and nuclei, have been strictly confined to move on a line. This means we retain the Coulomb interaction of the three dimensional particles. This is problematic, the singularity of this interaction is exceptionally strong in one dimension and requires special techniques be used to circumvent it.

In our study we employ the Dirichlet boundary conditions, which require the wavefunction to vanish whenever two particles touch. This brings some severe consequences including, most controversially, that the nuclei are impenetrable to the electrons. However it does permit finite binding energies of electrons to nuclei and has a number of unique features which make it an interesting system to study for insight into electron behaviour.

Here we explore the mechanics of chemistry within this model. We construct an unusual periodic table for one dimensional elements and explore the mechanisms by which they bind into molecules. Ultimately, we are able to develop a set of simple rules with which one can easily predict the outcome of a reaction in this unique model.

The generalised local density approximation is a new method for constructing density functional approximations. The local density approximation, the most simplistic of all density functionals, is built off the properties of the infinite uniform electron gas and as result replicates them exactly. It is also possible to construct finite uniform electron gases and, contrary to expectations, the regular local density approximation is unable to model these finite gases correctly.

The generalised local density approximation incorporates these finite gases into its construction. We account for their differences from the infinite gases by introducing a new parameter which measures the proximity of electrons at a point in space. We can then construct a correlation functional for one dimensional systems using a method which closely resembles that used for previous local density approximations.

Because of the definition of the new parameter, one would formally consider this functional to be a meta-GGA functional. However it retains the exceptionally simple form of a local density approximation.

This new functional is observed to have greatly improved accuracy compared to the standard local density approximation when tested against a variety of systems, including reactions of the one dimensional molecules identified earlier. Although we only construct a one dimensional functional here, the method is easily generalised to other dimensions.

Contents

List of Publications

The following papers have been published during the preparation of this thesis, and their content appears within:

1. P. F. Loos, C. J. Ball and P. M. W. Gill, "Uniform electron gases. II. The generalized local density approximation in one dimension", *J. Chem. Phys.* **140**, 18A524 (2014).

2. P. F. Loos, C. J. Ball and P. M. W. Gill, "Chemistry in one dimension", *Phys. Chem. Chem. Phys.* **17**, 3196 (2015).

3. C. J. Ball and P. M. W. Gill, "Chem1D: a software package for electronic structure calculations on one-dimensional systems", *Mol. Phys.* **113**, 1843 (2015).

4. C. J. Ball, P. F. Loos and P. M. W. Gill, "Molecular electronic structure in one-dimensional Coulomb systems", *Phys. Chem. Chem. Phys.* **19**, 3987 (2015).

The following paper has also been published during the preparation of this thesis, but does not appear within:

5. F. J. M. Rogers, C. J. Ball and P. F. Loos, "Symmetry-broken local-density approximation for one-dimensional systems", *Phys. Rev. B* **97**, 235114 (2016).

Chapter 1

Introduction

1.1 Quantum chemistry

With the advent of quantum mechanics in the early twentieth century chemists gained the necessary physical and mathematical tools to predict the outcome of many chemical processes from first principles. The computational demands of this problem are so immense that it has taken considerable time for this to become an attractive exercise. The advances in both technology and theoretical understanding that have occurred since have lead to this becoming a significant feature of the modern study of chemistry.

Unfortunately, many chemical systems remain well outside the reach of modern theoretical methods. And so extensive research still continues into the improvement of these methods. In this thesis we will explore two main topics which relate to such developments.

Our first topic concerns a newly defined model chemistry. Within this model are some unique features that may allow for new insights into the behaviour of particles in chemical systems. The second topic involves the initial development of a method which builds upon one of the oldest and simplest. This construction takes a new direction from the work preceding it which may also offer new insights and act as a new base for further development.

Before we discuss these topics we will present some of the necessary background to their discussion.

1.1.1 The Schrödinger wave equation

Within the field of quantum chemistry Schrödinger's wave equation [1, 2] is the dominant formulation of quantum mechanics. For our purposes we need only consider

the time-independent equation, which can be given as

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

where we have adopted atomic units, which will be used throughout this thesis unless otherwise stated. This simple equation hides a wealth of complexity. In order to observe this let us examine what the components of the equation are.

The Hamiltonian operator, $\hat{H}$, describes the energetic components of the system. Most importantly it contains terms which describe the kinetic and potential energies of a particle in the system being considered. In a complete treatment of a quantum system it is not limited to this, and may, for example, contain terms which describe the interaction of the particles with an external field. Such quantities will play only a small role throughout the remainder of this thesis. Removing them leaves us with the following Hamiltonian

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_{ee} + \hat{V}_{eN} + \hat{V}_{NN} \tag{1.2}$$

$\hat{T}_e$ and $\hat{T}_N$ are operators which describe the kinetic energies of the electrons and nuclei respectively, while $\hat{V}_{ee}$, $\hat{V}_{eN}$ and $\hat{V}_{NN}$ describe the potential energy from the interaction between electrons, electrons and nuclei and between nuclei.

The other key component of the Schrödinger wave equation is Ψ, the wavefunction. This is a high dimensional function which completely describes the quantum state of all the particles present within the system. In principle, with the knowledge of this object it is possible to determine any physical property of the system.

Finally, the E in the wave equation is simply the total energy of the system, a scalar quantity. Hence the equation describes what is known as an *eigenproblem*. We begin with knowledge of $\hat{H}$, which can be constructed for a given system of particles. From this we search for a function Ψ which when acted upon by $\hat{H}$ will be unchanged apart from a constant factor.

Unfortunately, despite such problems being widely studied by both mathematicians and physicists, the Schrödinger wave equation is unable to be solved analytically (except in a handful of special cases). This should not be surprising; the potential interactions couple the movement of particles together, resulting in this being equivalent to the many-body problem of classical physics. A solution to that problem has thus far eluded physicists for centuries, so we are reduced to finding approximations to the solution of the equation. Before we explore some of these approximations, we consider a few important points from quantum mechanics.

1.1.2 The Pauli exclusion principle

As stated above, the wavefunction is a complicated object which describes the quantum state of the particles in a system. An important principle of quantum mechanics is that two electrons (or more generally, two fermions) cannot occupy the same state. This is called the Pauli exclusion principle, after the physicist who first described it [3].

This has an important manifestation in the mathematics of the Schrödinger wave equation. It requires that the wavefunction be antisymmetric with respect to the exchange of any two electrons. That is to say

$$\Psi(\ldots, \mathbf{x}_i, \mathbf{x}_j, \ldots) = -\Psi(\ldots, \mathbf{x}_j, \mathbf{x}_i, \ldots) \tag{1.3}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the full coordinates of two electrons, including both their positional vectors ($\mathbf{r}_i$ and $\mathbf{r}_j$) and their spin coordinates (s_i and s_j).

Spin is a quantum mechanical property of an electron, and represents a type of momentum intrinsic to a particle. In the case of electrons it can take one of two values, either $+1/2$ or $-1/2$. This spin coordinate is part of the quantum state of an electron (the remainder being its spatial description), and a chemically important consequence of this principle is that any spatial quantum state can be occupied by up to two electrons, and no more, at any one time.

1.1.3 Observables and self-adjoint operators

From the wavefunction one can find *any* observable quantity of the system. One of the tenets of quantum mechanics is that an observable quantity can be described by what is called a self-adjoint operator. An operator $\hat{\mathcal{O}}$ is self adjoint if $\langle f, \hat{\mathcal{O}}g \rangle = \langle \hat{\mathcal{O}}f, g \rangle$. In the Hilbert space used in quantum chemistry this statement becomes

$$\int f^*(\mathbf{r})\hat{\mathcal{O}}g(\mathbf{r})d\mathbf{r} = \int g^*(\mathbf{r})\hat{\mathcal{O}}f(\mathbf{r})d\mathbf{r} \tag{1.4}$$

$$\langle f|\hat{\mathcal{O}}|g \rangle = \langle g|\hat{\mathcal{O}}|f \rangle \tag{1.5}$$

where $\langle f|\hat{\mathcal{O}}|g \rangle$ is the Dirac notation that is common in the study of quantum physics.

A simple example of this is determining the energy of a wavefunction. In this case the Hamiltonian itself is the self-adjoint operator, and the energy is given by

$$E = \langle \Psi|\hat{H}|\Psi \rangle = \int \Psi^*(\mathbf{x})\hat{H}\Psi(\mathbf{x})d\mathbf{x} \tag{1.6}$$

where we have assumed that the wavefunction is normalised.

Later, it will be important to note that the relationship between observable quantities and self-adjoint operators is an equivalence. Hence we can state that if an operator is *not* self-adjoint then it cannot describe an observable quantity.

1.1.4   The variational principle

Eigenproblems such as the Schrödinger wave equation permit a spectrum of solutions. For simplicity's sake, we'll assume that the spectrum is discrete rather than continuous. One of the solutions, which we label Ψ_0, must then have the lowest energy. Enumerating the other solutions as Ψ_i, the corresponding eigenvalues E_i satisfy $E_0 < E_i$ for all $i \neq 0$.

This solution is called the ground state, and is the state that the system will prefer to occupy. The variational principle states that any approximate, or trial, wavefunction must then be higher in energy, that is

$$E_0 \leq \langle \Psi^{\mathrm{trial}} | \hat{H} | \Psi^{\mathrm{trial}} \rangle \tag{1.7}$$

where Ψ^{trial} is the approximate wavefunction, which we assume here to be normalised. This results in the true energy being a lower bound to the energy of any approximation to the true wavefunction.

This is an incredibly useful result, as it allows us to judge the relative quality of different approximate wavefunctions. Not all quantum chemistry methods use the variational method to obtain their estimate however, but we will discuss one which does shortly.

1.1.5   The Born-Oppenheimer approximation

The complete Schrödinger wave equation treats all particles within a system in an identical manner. In any given molecule we consider two kinds of particles: electrons and nuclei. There is an enormous disparity in the mass of these two particles however; nuclei are much more massive than electrons. As a consequence of this electrons have far greater kinetic energy than the nuclei.

This observation leads to a first, and very common, approximation that was originally described by Born and Oppenheimer [4]. From the perspective of the electrons, the nuclei move so slow that they are essentially stationary. More rigorously, we can separate out part of the problem by considering a simplified Hamiltonian

operator where the nuclei are fixed in place

$$\hat{H}_{\text{elec}} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{eN} \tag{1.8}$$

$$\hat{H}_{\text{elec}} \Psi_{\text{elec}} = E_{\text{elec}} \Psi_{\text{elec}} \tag{1.9}$$

The result of this equation is an *electronic* wavefunction which describes only the state of the electrons. This is then *parametrically* dependent upon the nuclear coordinates; for every possible arrangement of the nuclei there is a unique electronic wavefunction. The total energy can then be obtained by simply adding the energy originating from the Coulombic repulsion between the nuclei to the electronic energy.

Having solved the electronic problem it is then possible to construct the total wavefunction by factoring in the motion of the nuclei. If we continue the same line of reasoning then we can approximate the effect of the electrons on the nuclei by considering only their average field.

Although not always the case, this assumption typically yields extremely accurate approximations. For the remainder of this thesis we will take the Born-Oppenheimer approximation as given, and will concern ourselves only with the determination of the electronic wavefunction. As such we will no longer explicitly refer to the electronic wavefunction, shortening this to just "the wavefunction". Any reference to the energy of a molecule will refer to the total energy, i.e. the sum of the electronic energy and the nuclear repulsion, unless otherwise stated.

1.2 The Hartree-Fock method

1.2.1 The Hartree-Fock wavefunction

Although the Born-Oppenheimer approximation significantly simplifies the wave equation the electronic wavefunction is still a much too complicated object to, at least in general, find analytically. This difficulty is caused by the coupling of electron motion via the Coulomb operator. If we were to decouple the electrons, however, the problem would be reduced to one that could be much more easily solved.

To achieve this decoupling we begin by introducing an ansatz wavefunction. This was a key suggestion by Hartree [5], who proposed approximating the wavefunction with the following form

$$\Psi_{\text{Hartree}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \prod_{i=1}^{N} \psi_i(\mathbf{r}_i, s_i) \tag{1.10}$$

where the functions ψ_i are referred to as molecular spin orbitals. These are wavefunction-like objects that describe the quantum state of a single electron, rather than the whole ensemble. Note that each spin orbital is a product of a spatial orbital $\psi_i(\mathbf{r}_i)$, which describes the movement of an electron through space, and a spin function that encodes the spin state associated with the orbital.

Such a wavefunction does not satisfy the antisymmetry requirement that Pauli's exclusion principle demands. This was independently observed by both Fock [6] and Slater [7] shortly after Hartree's initial proposal. Slater suggested correcting the problem by expanding the simple product into a determinant which correctly captures the required antisymmetry.

$$\Psi_{\text{HF}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \begin{vmatrix} \psi_1(\mathbf{r}_1, s_1) & \psi_1(\mathbf{r}_2, s_2) & \cdots \\ \psi_2(\mathbf{r}_1, s_1) & \psi_2(\mathbf{r}_2, s_2) & \cdots \\ \vdots & \vdots & \ddots \end{vmatrix} \tag{1.11}$$

A determinant of this kind is referred to as a Slater determinant. By invoking the variational principle and minimising the total energy of this determinant we obtain a useful approximate wavefunction. This wavefunction is referred to as the Hartree-Fock wavefunction and the method through which it is constructed the Hartree-Fock method (often simply shortened to Hartree-Fock and abbreviated as HF).

In this section we will give an overview of some of the details of this method, and is largely based upon the truly comprehensive description in the classic text by Szabo & Ostlund[8]. However in the following chapters spin will have a vanishing role in the mechanics of the calculations, and so from here on we will omit the details of its involvement in computation, preferring to work with the spatial orbitals $\psi_i(\mathbf{r}_i)$ where possible.

1.2.2   The Fock potential

The variational theorem tells us that to find the correct Hartree-Fock wavefunction we must obtain the Slater determinant with the lowest possible total energy. To begin with, it can be shown that the energy of an arbitrary determinant constructed

from the set of orbitals $\{\psi_a, \psi_b, \psi_c \ldots\}$ is given by the expression

$$E = \langle \Psi_{\mathrm{HF}} | \hat{H} | \Psi_{\mathrm{HF}} \rangle \tag{1.12}$$

$$= \sum_a \langle a | \hat{h}(1) | a \rangle + \frac{1}{2} \sum_{ab} (aa|bb) - (ab|ab) \tag{1.13}$$

$$= \sum_a \langle a | \hat{h}(1) | a \rangle + \frac{1}{2} \sum_{ab} (aa||bb) \tag{1.14}$$

$$\hat{h}(1) = -\frac{1}{2} \nabla_1^2 + \sum_A \frac{Z_A}{r_{1A}} \tag{1.15}$$

where Z_A is the charge of nucleus A and r_{1A} is the distance between electron 1 and nucleus A. We have used chemist's notation for the two electron integrals $(aa|bb)$, the use of which will continue throughout this thesis.

If we constrain the orbitals such that they are orthonormal to each other,

$$\langle \psi_i | \psi_j \rangle = \int \psi_i(\mathbf{r})^* \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{i,j} \tag{1.16}$$

where $\delta_{i,j}$ is the Kronecker delta function, then through a derivation of significant length we find that the correct HF orbitals are those which solve the equation

$$\left[\hat{h}(1) + \sum_b \mathcal{J}_b(1) - \sum_b \mathcal{K}_b(1) \right] \psi_a(1) = \varepsilon_a \psi_a(1) \tag{1.17}$$

where we have introduced the three new operators $\hat{h}(1)$, $\mathcal{J}_b(1)$ and $\mathcal{K}_b(1)$. The one particle operator $\hat{h}(1)$ collects the kinetic energies and the potential energies generated by the interaction between the nuclei and electrons, which are unchanged from the regular Schrödinger equation.

The other two operators $\mathcal{J}_b(1)$ and $\mathcal{K}_b(1)$, which are both two particle operators, do not appear in the Schrödinger equation. These operators are mean-field approximations to the correct electron-electron potential operator. That is, they give the potential interaction of an electron with the averaged field of the other electrons. $\mathcal{J}_b(1)$ gives the classical Coulomb interaction between two charged particles, while $\mathcal{K}_b(1)$ returns the so called exchange interaction of quantum mechanical systems, which arises as a consequence of the antisymmetry of the wavefunction. These expressions can be defined as

$$\mathcal{J}_b(1)\psi_a(1) = \int \frac{|\psi_b(2)|^2}{r_{12}} d\mathbf{r}_2 \psi_a(1) \tag{1.18}$$

$$\mathcal{K}_b(1)\psi_a(1) = \int \frac{\psi_b(2)^* \psi_a(2)}{r_{12}} d\mathbf{r}_2 \psi_a(1) \tag{1.19}$$

The key gain of this process is that the overall operator is identical for every electron in the system. This allows us to define a one electron operator called the Fock operator,

$$\hat{f}(1) = \hat{h}(1) + \sum_b \left[\mathcal{J}_b(1) - \mathcal{K}_b(1) \right], \tag{1.20}$$

and write the HF equations, those which give the orbitals that construct the lowest energy single Slater determinant, as

$$\hat{f}\psi_a = \varepsilon_a \psi_a. \tag{1.21}$$

Note that the Fock operator is dependent upon the orbitals ψ_a, resulting in this equation becoming a non-linear eigenfunction problem. While solving this problem and obtaining the correct HF solution is possible in principle, it is ill-suited to large scale attack by computer. Given the size of most molecular systems this quickly becomes necessary.

1.2.3   The Roothaan-Hall equations

The Fock potential allows us to approximate the complicated Schrödinger wave equation by a system of much simpler equations. This does not, however, permit a useful, generally applicable algorithm to obtain the solution. To achieve this we will introduce the Roothaan-Hall [9, 10] equations which, by introducing a basis set for expanding the orbitals, reduces the problem to an iteratively solvable matrix equation.

More specifically, we will introduce a slight variation of the Roothaan-Hall equations. This method was first described for closed-shell systems, i.e. where all the spatial orbitals are doubly occupied by a spin-up and a spin-down electron. Here we will describe the equations for a fully ferromagnetic system, i.e. where all the spatial orbitals are occupied by a single spin-up electron.

Strictly speaking, this is a different special case of the Pople-Nesbet equations, which generalise the Roothaan-Hall equations to any open-shell system [11]. Later in this thesis we will be considering ferromagnetic systems exclusively, and so a full description of the Pople-Nesbet equations will not be required.

The first step in this process is introducing a basis set for describing the spatial component of the HF orbitals. Given a set of functions $\{\phi_1, \phi_2, \ldots, \phi_M\}$ we can

expand the orbitals as a linear combination

$$\psi_a = \sum_{\mu=1}^{M} c_{a,\mu} \phi_\mu \tag{1.22}$$

This is only exact if the chosen basis set is complete over the space in which the function ψ_a resides. In this case, such a requirement usually demands that the basis set be of infinite size. This is not possible from a computational perspective, and so finite basis sets are used and the expansion in Eq. (1.22) is only an approximation. The closer the basis set is to completeness, the better this approximation becomes.

There are many possible choices of basis set, and later in this thesis such choices will be discussed. For the remainder of this section the precise nature of the basis set is unimportant and so we will not concern ourselves with these details.

If we substitute the Eq. (1.22) into the HF equation obtained above, Eq. (1.21), it can be reduced to a matrix equation by introducing the four matrices

$$\mathbf{C}_{i,j} = c_{i,j} \tag{1.23}$$

$$\varepsilon_{i,j} = \delta_{i,j}\varepsilon_i \tag{1.24}$$

$$\mathbf{S}_{\mu,\nu} = \int \phi_\mu^* \phi_\nu d\mathbf{r} \tag{1.25}$$

$$\mathbf{F}_{\mu,\nu} = \int \phi_\mu^* \hat{f} \phi_\nu d\mathbf{r} \tag{1.26}$$

like so

$$\mathbf{FC} = \varepsilon \mathbf{SC}. \tag{1.27}$$

The matrices $\mathbf{C}$ and ε now contain the coefficients in the expansion of Eq. (1.22) and the energies of the molecular orbitals. $\mathbf{S}$ contains the overlap between the basis functions used and is, unsurprisingly, called the overlap matrix. We can also define a charge density matrix $\mathbf{P}$

$$\mathbf{P}_{\mu,\nu} = \sum_a \mathbf{C}_{\mu,a} \mathbf{C}_{\nu,a}^*. \tag{1.28}$$

which simplifies the definition of the matrix $\mathbf{F}$, typically called the Fock matrix, as follows

$$\begin{aligned} \mathbf{F}_{\mu,\nu} &= \int \phi_\mu^* \hat{h}(1)\phi_\nu d\mathbf{r} + \sum_{\lambda\sigma} \mathbf{P}_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - (\mu\lambda|\nu\sigma) \right] \\ &= \int \phi_\mu^* \hat{h}(1)\phi_\nu d\mathbf{r} + \sum_{\lambda\sigma} \mathbf{P}_{\lambda\sigma} (\mu\nu||\lambda\sigma) \end{aligned} \tag{1.29}$$

It is important to observe that for an orthonormal basis the overlap matrix would simply be the identity matrix, since by the definition of orthonormal functions (in the space L^2) we find

$$\int \phi_\mu^* \phi_\nu d\mathbf{r} = \delta_{i,j}. \tag{1.30}$$

While it is common to employ a normalised basis set it is rare for it to also be orthonormal, at least for computations on molecular systems. This makes it necessary to introduce an orthogonalisation procedure which involves diagonalising the overlap matrix $\mathbf{S}$. This can be problematic if there is near linear dependence in the basis set, as it can lead to numerical instabilities.

The result of the orthogonalisation is a transformation matrix, usually called $\mathbf{X}$, which moves between the chosen basis set to an orthogonalised set. This allows us to recast Eq. (1.27) as a simple matrix eigenproblem

$$\mathbf{F}'\mathbf{C}' = \varepsilon\mathbf{C}' \tag{1.31}$$

where the primes indicate the matrices have been transformed into the orthogonal basis. If we can construct the Fock matrix $\mathbf{F}$, then diagonalising $\mathbf{F}'$ allows us to generate both ε and $\mathbf{C}'$. We need only apply the transformation $\mathbf{X}$ to return the coefficient matrix back to the original basis to obtain the solution $\mathbf{C}$.

Unfortunately, the dependence of the Fock matrix $\mathbf{F}$ on the coefficient matrix $\mathbf{C}$ (via the density matrix $\mathbf{P}$ in Eq. (1.29)) causes this to remain a non-linear eigenproblem. While this makes a direct solution difficult, an iterative scheme can be used to solve the problem.

The equivalence in Eq. (1.27) only holds for the correct coefficient matrix $\mathbf{C}$. That is, if we generate the Fock matrix $\mathbf{F}$ from $\mathbf{C}$ this relationship is only true if $\mathbf{C}$ is the correct solution. In an iterative scheme we begin with a trial set of coefficients $\mathbf{C}_0$ that we use to generate a Fock matrix. We can then follow the procedure to obtain a new coefficient matrix $\mathbf{C}_1$.

Should $\mathbf{C}_1 \approx \mathbf{C}_0$, using some chosen metric and to within some threshold, then we know that $\mathbf{C}_0$ solves the equations. In the more likely case that it is not, then we can construct a new Fock matrix from $\mathbf{C}_1$ and repeat the process. Ideally, if we continue this process then we will eventually converge to a solution. This method is called the self consistent field (SCF) method, since we search for a solution which is consistent with the field that it generates.

There is no guarantee of convergence during this process, and both the success of the process and the speed with which the solution is found depend heavily upon the quality of the guess. As such, choosing the initial starting point for the SCF procedure

is not necessarily a trivial process. Additionally, there are multiple techniques for improving the convergence behaviour which can be added to the algorithm (perhaps most notably the DIIS method of Pulay [12, 13]). Fortunately, in this thesis we will not find ourselves grappling with such difficulties, and so we will omit further discussion of this topic.

1.2.4 Aftermath

With the Roothaan-Hall equations constructed, we now have a general method for solving an approximation to the Schrödinger wave equation. The question remains: is this approximation any good? As it turns out, the Hartree-Fock wavefunction yields an excellent approximation to the total energy of a molecular system.

It is unfortunate that quantum chemists are rarely interested in the raw total energies of molecules. It is differences that are typically required when asking chemical questions. At the simplest level, a reaction being favourable depends upon whether the products of a reaction have a lower energy than the reactants. Similarly, when considering the kinetics of a reaction it is the height of the transition barrier that is essential, which is the difference between the energy of the transition state and the reactants.

During most reactions the vast majority of a molecule's total energy is unchanged. The nuclei still attract the electrons by a similar amount, and the electrons, particularly those close to the nuclei, maintain a similar level of kinetic energy. The consequence is that the small fraction of the total energy missed by the Hartree-Fock approximation is essential to describing the energetics involved in chemistry.

1.3 Wavefunction-based correlated methods

The difference between the Hartree-Fock energy and the exact energy is called the correlation energy,

$$E_c = E - E_{\text{HF}}. \tag{1.32}$$

Because the Hartree-Fock energy is a variationally minimised approximation to the exact energy it must be greater than the exact energy. As a result the correlation energy is a negative number, and represents stabilising forces which are not accounted for in the Hartree-Fock method and wavefunction.

This difference is a consequence of the Hartree-Fock potential being a mean-field approximation: each electron feels the averaged electric field generated by the other electrons. If we were to allow the electrons to interact with the instantaneous

positions of each other they would be capable of relaxing into more energetically favourable orbits by correlating their movement with one another. This is the source of both the correlation energy's physical nature and the name given to it.

In this section we will describe a collection of methods for estimating the correlation energy of a molecule that will be used later in this thesis. For the moment we will restrict ourselves to wavefunction-based methods, i.e. those which primarily work with the wavefunction to obtain their estimate. There is another class of correlated methods which are based on the electron density that we will discuss afterwards.

1.3.1   Møller-Plesset perturbation theory

A powerful and well established method for finding an approximate solution to a difficult problem is perturbation theory. The idea is to find an exact solution to a simplified version of the problem that needs to be solved. This is also an approximation to the solution of the original problem, and the closer the simplified problem is to the original, the better the approximation.

The difference between the original and the simplified problem can then be seen as a perturbation to the simplified problem. Perturbation theory observes how this difference affects the approximate solution and from this attempts to correct the approximation. The standard perturbation technique used in quantum chemistry is Møller-Plesset theory [14], which is an application of Rayleigh-Schrödinger (RS) perturbation theory [2, 15] to the correlation problem.

RS perturbation theory separates the Schrödinger equation in the following manner

$$\hat{H}\Psi = \hat{H}_0\Psi + \lambda\hat{V}\Psi = E\Psi \tag{1.33}$$

where the operator $\hat{H}_0$ is chosen such that the solution to the equation

$$\hat{H}_0\Psi^{(0)} = E^{(0)}\Psi^{(0)} \tag{1.34}$$

is known.

To obtain a solution to this problem both the wavefunction and energy are expanded as a power series in λ

$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots \tag{1.35}$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \tag{1.36}$$

This expansion can be substituted into Eq. (1.33) and the coefficients of the powers of λ equated. This gives a sequence of equations that relate the known solution $\Psi^{(0)}$

and $E^{(0)}$ to the unknown terms of Eqs. (1.35) and (1.36). Ultimately, this allows us to express the terms of the energy expansion using the spectrum of $\hat{H}_0$.

It was Møller and Plesset who first described the application of RS perturbation theory to molecular systems [14]. Their formulation uses the Hartree-Fock wavefunction as the starting point for the perturbation, i.e. $\hat{H}_0$ is taken to be the Fock operator. This leaves the perturbation operator $\hat{V}$ as the difference between the exact Coulomb potential and the mean-field approximation of the Fock potential

$$\hat{V} = \sum_{ij} r_{ij}^{-1} - \sum_{b} \left[\mathcal{J}_b(1) - \mathcal{K}_b(1) \right] \tag{1.37}$$

where both $\mathcal{J}_b(1)$ and $\mathcal{K}_b(1)$ have been defined in Eqs. 1.18 and 1.19 respectively.

Having made this separation the terms of Eq. (1.36) can be cast in terms of integrals over the HF molecular orbitals. These quickly become lengthy (but not necessarily complicated), however the first few are given by the relatively simple expressions

$$E_0 = \sum_{a} \epsilon_a \tag{1.38}$$

$$E_1 = -\frac{1}{2} \sum_{ab}^{\text{occ}} (aa||bb) \tag{1.39}$$

$$E_2 = \frac{1}{4} \sum_{ab}^{\text{occ}} \sum_{rs}^{\text{virt}} \frac{|(ar||bs)|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} \tag{1.40}$$

where ψ_a and ψ_b represent occupied HF orbitals and ψ_r and ψ_s are unoccupied (virtual) HF orbitals. Partial summations of this series give approximations to the true energy of the system. The energy of the HF wavefunction is given by the sum of the zero-th and first order term ($E_{\text{HF}} = E_0 + E_1$), so the first correlated energy estimate is given by the second-order sum. This approximation is referred to as MP2, and including higher terms gives the MP3 energy, MP4 energy and so on.

In principle the infinite summation $\sum_n E_n$ gives the exact energy. Unfortunately, for many molecular systems, this summation begins diverging shortly into the sequence of partial sums[16]. Furthermore, higher order corrections quickly escalate in computational difficulty. For these reasons only the low order corrections are commonly used (i.e. MP2, MP3 and occasionally MP4).

1.3.2 Configuration interaction

While Møller-Plesset theory is a useful method for obtaining low cost corrections for the correlation energy, its use is limited when one is looking for extremely high

accuracy. Configuration interaction (CI) is another method, even simpler in concept, for obtaining a correlation energy [8]. The key draw for the CI method is that, when used to its full extent, it is capable of finding the exact solution to the Schrödinger equation.

Where the HF wavefunction is a Slater determinant constructed from a set of one-electron spin orbitals, the CI wavefunction builds upon HF by using a linear combination of Slater determinants. The set of determinants used for the CI wavefunction is composed of all the determinants Ψ_i that can be constructed by occupying different configurations of the HF orbitals.

The final CI wavefunction is found by choosing the coefficients in the linear expansion which, according to the variational theory, minimise the total energy. This is achieved by constructing a Hamiltonian matrix which describes the interactions between the determinants, and has the elements

$$\mathbf{H}_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle. \tag{1.41}$$

The lowest eigenvalue of this matrix is the exact ground state energy of the system, within the limits of the basis set employed during the initial HF calculation. The corresponding eigenvector contains the coefficients of the determinants in the expansion of the ground state wavefunction.

The method described above is referred to as full CI since it makes use of the full set of determinants. Although this method gives the correct answer to the Schrödinger equation (when used within a complete basis set) it is rarely used due to its computational expense. Constructing the determinants is a combinatorial problem, and leads to a factorial scaling with system size. For all but the smallest and simplest of systems this problem is intractable.

It is, however, possible to truncate the set of determinants used in the wavefunction expansion. For example, in the CISD method only those determinants which differ from the ground state determinant by one (single excitations) or two (double excitations) orbitals are included in the basis set. This makes a significant difference to the computational scaling of the method by dramatically reducing the number of possible determinants. However, it does introduce its own difficulties. In this thesis we will employ the CI method in order to generate high accuracy reference values for systems where the Full CI problem is tractable, and so we will refrain from describing the truncated CI methods further.

1.3.3 Quantum Monte Carlo

Quantum Monte Carlo (QMC) is a stochastic numerical method. It differs from the previously described correlated methods by not being a strictly post HF correction. In QMC, the energy of an ansatz wavefunction is directly evaluated via numerical methods and the wavefunction manipulated to achieve the best possible energy. Like Full CI, QMC is capable of obtaining the exact energy and wavefunction in principle. Depending on the required precision, however, it may become intractable to do so, and the method has other limitations which may prevent it from achieving the correct solution.

A QMC calculation typically employs two separate steps. A Variational Monte Carlo (VMC) step generates a high accuracy first approximation to the solution. This result is then fed to a Diffusion Monte Carlo (DMC) step, which relaxes the wavefunction further and is capable of extremely high accuracy at the cost of slowly converging results. This theory only makes a brief appearance later in this thesis, and so we will refrain from exploring the details of the construction.

In VMC a trial wavefunction $\Psi_T(q; \mathbf{r})$ is chosen with a set of variable parameters q. The variational principle is invoked and these parameters are optimised to minimise the energy of the wavefunction. The key feature of VMC is that the energy is found by directly evaluating the integral

$$E = \frac{\langle \Psi_T | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \frac{\int |\Psi_T(q; \mathbf{r})|^2 E_L \, d\mathbf{r}}{\int |\Psi_T(q; \mathbf{r})|^2 \, d\mathbf{r}} \qquad (1.42)$$

where

$$E_L = \frac{\hat{H} \Psi_T(q; \mathbf{r})}{\Psi_T(q; \mathbf{r})} \qquad (1.43)$$

is known as the local energy. This is done by using an importance sampling of the distribution $|\Psi_T(q; \mathbf{r})|^2$.

Unlike what has been discussed previously, DMC makes use of the time-dependent Schrödinger wave equation. Again, a trial wavefunction is used, and it is allowed to propagate through imaginary time. One can show that through this process the wavefunction will eventually reach the exact ground-state solution. The propagation step requires a high-dimensional integral involving the wavefunction, which is evaluated using a stochastic Monte-Carlo method.

1.4 Density Functional Theory

For most of the twentieth century wavefunction based methods were the prevailing approach for solving the Schrödinger equation for chemical systems. It is not the only way to approach the problem however. Modern computational chemistry is dominated by Density Functional Theory (DFT), a very different approach to working with the Schrödinger equation.

Rather than focussing on describing the wavefunction, DFT methods work with the electronic density instead. This was shown to be possible by Hohenberg and Kohn in their seminal 1964 paper [17], in which they prove that the ground-state properties of a system are uniquely determined by the electronic density. Modern applications of this theory also rely upon the Kohn-Sham equations [18], which provide a similar framework for DFT as HF theory does for wavefunction methods.

1.4.1 The Hohenberg-Kohn theorems

The validity of DFT rests upon two proofs given by Hohenberg and Kohn [17]. In order to explore these theorems let us define a Hamiltonian

$$\hat{H} = \hat{T} + \hat{U} + \hat{V} \tag{1.44}$$

$$\hat{T} = -\frac{1}{2}\sum_i \nabla_i^2; \qquad \hat{U} = \sum_{i,j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}; \qquad \hat{V} = \sum_i \nu(\mathbf{r}_i) \tag{1.45}$$

where $\nu(r)$ is an external potential in which the electrons are moving. In a molecular system this potential would be the one generated by the nuclei. We will also assume that this Hamiltonian permits a non-degenerate ground-state wavefunction Ψ, with an associated electronic density ρ that is given by

$$\rho(\mathbf{r}) = \int |\Psi(\mathbf{r}, s_1, \mathbf{x}_2, \mathbf{x}_3, \ldots)|^2 ds_1 d\mathbf{x}_2 d\mathbf{x}_3 \ldots d\mathbf{x}_N \tag{1.46}$$

The first Hohenberg-Kohn theorem proves that a given electronic density uniquely determines the external potential. To prove this let a second potential $\nu'(r)$ define a Hamiltonian $\hat{H}'$. This Hamiltonian has an associated wavefunction Ψ' that we will assume generates the same density, ρ, as the first wavefunction Ψ. We know that $\Psi \neq \Psi'$ since they satisfy different Hamiltonians. By invoking the variational

principle we find

$$
\begin{aligned}
E' &= \langle \Psi' | \hat{H}' | \Psi' \rangle \\
&< \langle \Psi | \hat{H}' | \Psi \rangle \\
&< \langle \Psi | \hat{H} - \hat{V} + \hat{V}' | \Psi \rangle \\
&< E + \langle \Psi | \hat{V}' - \hat{V} | \Psi \rangle
\end{aligned}
\tag{1.47}
$$

and similarly

$$
E < E' + \langle \Psi | \hat{V} - \hat{V}' | \Psi \rangle
\tag{1.48}
$$

Summing equations (1.47) and (1.48) gives the contradiction

$$
E + E' < E + E'
\tag{1.49}
$$

proving that no two external potentials can generate the same electronic density.

The key consequence of this theorem is that there is no information lost by ignoring the wavefunction in favour of the density. This implies it is possible to extract any observable property of a quantum system from its density. A simple example of this would be a functional $E_\nu[\rho]$ which extracts the energy of a density ρ in the potential field of ν.

The second of Hohenberg and Kohn's theorems concerns such a functional, asserting that it is minimised by the density which corresponds to the given external potential. This is an extension of the variational principle of the wavefunction to the electronic density.

To prove this we define a Hamiltonian $\hat{H}$ with an external potential ν. This permits a wavefunction Ψ with an associated density ρ which we will compare with a trial density ρ'. From the first theorem we know that this trial density uniquely determines an external potential ν', which in turn defines a Hamiltonian for which a wavefunction Ψ' can be found. By applying the variational theorem to this wavefunction it follows that

$$
E_\nu[\rho'] = \langle \Psi' | \hat{H} | \Psi' \rangle \geq E_\nu[\rho] = \langle \Psi | \hat{H} | \Psi \rangle
\tag{1.50}
$$

While the Hohenberg-Kohn theorems provide an alternative route to approach the Schrödinger wave equation there is a significant barrier to its use. The theorems are non-constructive in nature, and so while we know an exact energy functional exists, we do not know what it is. Unfortunately deriving such a functional turns out to be a significantly difficult problem.

Of particular note is the difficulty in deriving a functional which correctly evaluates the kinetic energy, a quantity which would be considered trivial in wavefunction based methods. This problem is neatly sidestepped by working in the Kohn-Sham formalism, which introduces aspects of wavefunction theories in order to simplify the problem of optimising the electronic density.

1.4.2   The Kohn-Sham equations

The Kohn-Sham (KS) equations form the basis for the application of DFT in modern chemistry by providing a framework for optimising the electronic density self-consistently [18]. This is achieved by defining a set of non-interacting electron orbitals that have the same electronic density that generates the external potential of the desired system. The result bears a great deal of similarity to the approach of Hartree-Fock theory. It does, however, provide facility for the introduction of correlation effects by means of an appropriate density functional.

From the Hohenberg-Kohn theorems we know that the ground state energy of a Hamiltonian defined by the external potential ν can be written as

$$E_0 = \int \rho(\mathbf{r})\nu(\mathbf{r})dr + \frac{1}{2}\int\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1 d\mathbf{r}_2 + T[\rho] + F[\rho] \qquad (1.51)$$

where ρ is the ground state electronic density that produces the potential ν. The first and second term give the interaction of the electrons with the external potential and the mean-field Coulomb interaction respectively. The functional $T[\rho]$ gives the kinetic energy of the electrons while $F[\rho]$ describes the exchange and correlation effects.

We also know that this energy is minimised by ρ, and so we can determine via calculus of variations that

$$\int \delta\rho(\mathbf{r})\left(\nu_{\text{eff}}(\mathbf{r}) + \frac{\delta T[\rho]}{\delta\rho} + \nu_{\text{xc}}[\rho(\mathbf{r})]\right)d\mathbf{r} = 0 \qquad (1.52)$$

where we have defined

$$\nu_{\text{eff}}(\mathbf{r}) = \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \qquad (1.53)$$

$$\nu_{\text{xc}}(\rho) = \frac{\delta F[\rho]}{\delta\rho} \qquad (1.54)$$

Kohn and Sham observe that this same expression is obtained by applying the Hohenberg-Kohn theorems to a system of non-interacting particles that move within the external potential defined by $\nu_{\text{eff}}(\mathbf{r}) + \nu_{\text{xc}}[\rho(\mathbf{r})]$. Because of this it is possible to

obtain the density of the interacting system from the orbitals obtained by solving the one-particle equations

$$\left(-\frac{1}{2}\nabla^2 + \nu_{\text{eff}}(\mathbf{r}) + \nu_{\text{xc}}[\rho(\mathbf{r})] \right) \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{1.55}$$

and constructing the density with the following relationship

$$\rho = \sum_i |\phi_i|^2 \tag{1.56}$$

Constructing these orbitals, known as the Kohn-Sham orbitals, is achieved by introducing a basis set and a self-consistent algorithm in a similar manner to the Roothaan-Hall equations.

There are some notable aspects of KS theory. The introduction of the KS orbitals allows us to approximate the kinetic energy of the interacting system with that of the non-interacting orbitals. Deriving the true kinetic energy functional proves to be a difficult problem, and so this is a significant benefit of the method.

The error introduced by this approximation is then absorbed into the functional $F[\rho]$, which is the only remaining unknown quantity. Beyond accounting for the error in the kinetic energy, this functional is primarily responsible for describing exchange and correlation effects within the system. Unfortunately the KS orbitals are of little physical significance beyond reproducing the correct interacting density.

1.4.3 Local Density Approximations

Although the KS equations remove the need for finding a density functional representation of the kinetic energy, one is still needed for the exchange and correlation effects. Unfortunately, this is not an easy task. To begin working on this problem we turn to one of the great paradigms of modern physics: the uniform electron gas (UEG) [19].

Put simply, a UEG is a system where the electronic density is constant throughout. There are multiple ways to envisage the construction of a UEG. One possibility is to take a box of finite size that is filled with a constant background positive charge. Electrons are then placed in this box such that they achieve an overall neutrality with the background positive charge. These electrons will generate an average electron density ρ_{av}, which will be equal to the background charge, throughout the box. However the box boundaries will introduce oscillations into the actual electronic density ρ, and hence $\rho \neq \rho_{\text{av}}$.

If we expand the box while adding electrons so that charge neutrality is maintained and ρ_{av} remains constant then these oscillations will reduce in amplitude. At the limit of this process, when the box is infinite in size, the oscillations vanish and we are left with $\rho = \rho_{\mathrm{av}}$, which is to say the density is constant throughout space. This can be done for any value of ρ_{av}, and each choice will result in a UEG with unique properties.

This forms the basis of a simple first approximation to the true exchange and correlation density functional. Let ρ be an arbitrary electronic density. Within some small volume element $\Delta\mathbf{r}$ we can approximate ρ by a constant and, by extension, assume that it behaves like a UEG with a density equal to the average of ρ within $\Delta\mathbf{r}$. Of course, as this element shrinks in size this approximation improves.

Let us define a function $\varepsilon_{\mathrm{xc}}(\rho)$ which gives the reduced exchange and correlation energy of a UEG with density ρ. That is, it gives the exchange and correlation energy per electron within the UEG. Then we can state the above approximation more formally as

$$E_{\mathrm{xc}}[\rho] \approx \int \rho(\mathbf{r})\varepsilon_{\mathrm{xc}}[\rho(\mathbf{r})]d\mathbf{r} \tag{1.57}$$

This approximation is known as the Local Density Approximation (LDA), since it attempts to model the exchange and correlation energies based upon the local character of the density.

To make use of the LDA it is necessary to construct the function $\varepsilon_{\mathrm{xc}}(\rho)$. One typically starts by separating the exchange and correlation components

$$\varepsilon_{\mathrm{xc}}(\rho) = \varepsilon_{\mathrm{x}}(\rho) + \varepsilon_{\mathrm{c}}(\rho) \tag{1.58}$$

The exchange energy of a UEG with arbitrary density can be found analytically, which allows the term $\varepsilon_{\mathrm{x}}(\rho)$ to be constructed explicitly. The correlation energy is more difficult to model. The limiting high- and low-density behaviour can be found analytically, which gives useful boundary conditions for constructing this function. Constructing the function over the intermediate densities is done by a fitting procedure. This was first made possible when Ceperley and Alder published a set of correlation energies for UEGs of intermediate densities by using periodic Monte-Carlo techniques [20]. Correlation LDA functionals attempt to interpolate these energies and reproduce the correct limiting behaviour.

It should be evident that, due to the way it is constructed, applying an LDA functional to a UEG gives the correct results. On other systems its performance varies. When the electronic density is reasonably delocalised across the system, such as in metals and other solid state systems, the LDA can give quite respectable results.

Molecular systems are typically challenging for LDAs, where they tend to strongly over-bind the atoms. Fortuitous error cancellation can, however, lead to respectable reaction energy estimates.

1.4.4 Further developments

Although the LDA gives a good first approximation to a useful exchange and correlation functional, its accuracy leaves a lot to be desired for chemical systems. Most work on density functionals has focussed on augmenting the LDA in an attempt to achieve higher accuracy. Here we will give a brief overview of some of the notable classes of functionals that have been developed up to this point. For a far more in depth discussion of the state of DFT, Becke has published an excellent review of its history [21].

The failing of the LDA is the assumption that electrons behave like a UEG on a sufficiently small length scale regardless of the system. The clear next step is to model the effects of fluctuations in the density, i.e. to include the gradient of the density as a parameter of the functional.

By examining the effect of small perturbations to a UEG in a system known as the slowly varying electron gas it is possible to derive analytic gradient corrections to the LDA. The success of such corrections is only moderate, and so chemists have also turned to alternative constructions that achieve better accuracy. Collectively these functionals are known as Generalised Gradient Approximations (GGA), and constitute a significant advance in density functional design.

More can still be done, and moving beyond this we reach what are known as the meta-GGA functionals. These functionals include higher order density derivative quantities as well as more complicated objects such as the kinetic energy density

$$\tau = \sum_i |\nabla \psi_i|^2 \tag{1.59}$$

where ψ_i are the KS orbitals. These values are often included in order to achieve certain analytically derived corrections. Perhaps most notably they allow removal of the self-interaction error, a spurious LDA contribution which allows an electron to correlate its movement with itself. This is most clearly manifest in the non-zero correlation found when the LDA is applied to a one electron system such as the hydrogen atom. Meta-GGA functionals can achieve markedly improved accuracy, but oftentimes at a significant increase in computational cost and complexity.

Perhaps the greatest success for the application of DFT in chemistry, however, was the development of hybrid functionals. The pioneer in this respect was Becke's

B3PW91 functional from 1993 [22]. This functional combines the B88 exchange functional [23] with the PW91 [24] correlation functional, however it appears more commonly (to the exclusion of almost anything else in fact) as B3LYP, where the LYP correlation functional [25] is used in favour of PW91.

The main feature of this functional is its inclusion of a portion of the exact HF exchange energy alongside the exchange density functional

$$E_x^{\text{B3LYP}} = E_x^{\text{LDA}} + 0.20 \left(E_x^{\text{HF}} - E_x^{\text{LDA}} \right) + 0.72 \left(E_x^{\text{B88}} - E_x^{\text{LDA}} \right) \tag{1.60}$$

This combination allows for favourable error cancellation between the different exchange energies that leads to a highly successful description of chemical behaviour.

The development of hybrid functionals has lead to wide adoption of DFT methods by computational chemists. As a result there has been a significant proliferation of such functionals. There is a level of controversy associated with this, as the construction of these functionals is usually based on fitting to empirical data. This leads to questions of reliability across the wide array of possible chemical situations, as well as concern over a lack of physical insight.

Chapter 2

One-dimensional Chemistry

2.1 1D chemistry

Theoretical chemists typically concern themselves with electrons moving in three dimensional (3D) space. This is for obvious reasons, because electrons in reality are always moving within a 3D environment. There is no reason, however, that other dimensionalities cannot be considered, or that there is nothing to be gained by doing so. Indeed, it is possible to confine electrons experimentally as well, and one could attempt to model such situations. But there is also the possibility of other theoretical insights. In this chapter we will introduce and begin to study a new paradigm of one dimensional (1D) chemistry, wherein three dimensional (3D) particles are *strictly* confined to move along only a single spatial coordinate.

Experimentally, 1D systems can be realised in carbon nanotubes [26–30], organic conductors [31–35], transition metal oxides [36], edge states in quantum Hall liquids [37–39], semiconductor heterostructures [40–44], confined atomic gases [45–47] and atomic or semiconducting nanowires. Theoretically, Burke and coworkers [48, 49] have shown that 1D systems can be used as a "theoretical laboratory" to study strong correlation in "real" three-dimensional (3D) chemical systems within density-functional theory [50]. Herschbach and coworkers calculated the ground-state electronic energy of 3D systems by interpolating between exact solutions for the limiting cases of 1D and infinite-dimensional systems [51–53].

However, all these authors eschewed the Coulomb operator $1/|x|$. For example, Burke and coworkers [48, 49] used a softened version of the Coulomb operator $1/\sqrt{x^2 + 1}$ to study 1D chemical systems, such as light atoms (H, He, Li, Be, ...), ions (H^-, Li^+, Be^+, ...), and diatomics ($H_2{}^+$ and H_2). Herschbach and coworkers have worked intensively on the 1D He atom [54–57] replacing the usual Coulomb inter-particle interactions with the Dirac delta function $\delta(x)$ [58–61]. There are few

studies using the Coulomb operator because of its strong divergence at $x = 0$. Most of these focus on non-atomic and non-molecular systems [62–68]. Here, we prefer the Coulomb operator because, although it is not the solution of the 1D Poisson equation, it pertains to particles that are *strictly* restricted to move in a one-dimensional sub-space of three-dimensional space.

The first 1D chemical system to be studied was the H atom by Loudon [69]. Despite its simplicity, this model has been useful for studying the behavior of many physical systems, such as Rydberg atoms in external fields [70, 71] or the dynamics of surface-state electrons in liquid helium [72, 73] and its potential application to quantum computing [74, 75]. Most work since Loudon has focused on one-electron ions [69, 76–82] and, to the best of our knowledge, no calculation has been reported for larger chemical systems. In part, this can be attributed to the ongoing controversy concerning the mathematical structure of the eigenfunctions (parities and boundedness) [82–89].

According to more recent literature the underlying mathematics that drives this problem is the fact that the Coulomb operator is *not* self-adjoint in 1D as a result of the strength of its singularity [90, 91]. A correction to the operator must be applied to obtain the desired observable, however this correction cannot be uniquely determined by a mathematical analysis [92]. At the culmination of a series of papers [80, 93, 94], Oliveira and Verri have shown that, in the limit of a cylindrical confinement toward 1D, only one of the possible corrections permits a finite binding energy of an electron to a hydrogen nucleus. This is an extremely attractive property from a physical viewpoint.

On the basis of this evidence we adopt this correction in our model of 1D chemistry. It is equivalent to applying the Dirichlet boundary conditions to the wavefunction, which requires that the wavefunction vanish whenever two particles (electrons or nuclei) touch. Note that it has been previously established that this must occur when two electrons touch [95].

The Dirichlet boundary conditions carry three significant consequences on the structure of the solution to the Schrödinger equation. First, the system is spin-blind [62–64, 66], i.e. the energy is invariant under any change of spin coordinates. This means we are free to assume that all electrons have the same spin. Second, a Super-Pauli exclusion principle comes into effect where no two electrons may share the same spatial quantum state regardless of their spin state. In independent electron models such as Hartree-Fock (HF) theory [8] this results in orbitals having a maximum occupancy of one. Finally all particles are impenetrable to one another, in particular the electrons cannot tunnel through the nuclei [80, 87]. This has been

firmly established by an elegantly simple argument from Núñez-Yépez and coworkers [82] that shows the quantum flux to be zero at the nuclei.

One particularly potent ramification of the impenetrability of the nuclei is the separation of electrons into distinct regions of space. Since the electrons can never pass from one side of a nucleus to the other they become trapped by them, either occupying a ray (to the left or right of the molecule) or a line segment (between two nuclei). We refer to these intervals as domains, specifically infinite domains (for intervals outside the nuclei) and finite domains (for those between them).

In Sec. 2.3 and Sec. 2.4 of this chapter, we report electronic structure calculations for 1D atomic and molecular systems using the Coulomb operator $1/|x|$. Sec. 2.4 discusses several diatomic systems, the chemistry of H_3^+ and an infinite chain of 1D hydrogen atoms.

We have followed the methods developed by Hylleraas [96, 97] and James and Coolidge [98] to compute the exact or near-exact energies E_{exact} of one-, two-, and three-electron systems. Throughout this chapter we also make use of a collection of programs implemented in *Mathematica* [99] to obtain HF energies and Møller-Plesset perturbation energies. All of these are capable of highly accurate results (whether that be for exact energies or approximations) and, to the best of our knowledge, all the digits reported are correct. Unfortunately only small systems are within reach of these prototypes, in Chapters 3 and 4 we will discuss more elaborate programs with greater applicability.

2.2 Theory

2.2.1 Notation

Standard chemical notation is not sufficient for describing a 1D molecule because it has no way of expressing the relative positions of electrons and nuclei, a requirement given the impenetrability of the nuclei. Throughout this thesis we use a modified notation which does offer this description. Here the atoms are named individually as they appear from left to right. Subscripts are then placed to denote how many electrons occupy each domain (for an unoccupied domain the subscript is omitted). We only consider the ground state, so we assume that the electrons singly occupy the lowest energy orbitals available to them.

For example, the $_1\text{Li}_3\text{He}_1\text{H}_2\text{Be}_2$ molecule consists of four nuclei and nine electrons. There is a lithium nucleus on the left, followed by a helium nucleus, a hydrogen

nucleus and finally a beryllium nucleus on the right of the molecule. The electrons
are distributed as follows:

- One electron in the domain to the left of the Li nucleus

- Three in the domain between the Li and He nuclei

- One in the domain between the He and H nuclei

- Two in the domain between the H and Be nuclei

- Two in the domain to the right of the Be nucleus

As mentioned above, these electrons singly occupy the lowest energy orbitals of their
associated domains.

2.3   Atoms

2.3.1   Hydrogen-like ions

The electronic Hamiltonian of the 1D H-like ion with nucleus of charge Z at $x = 0$ is

$$\hat{H} = -\frac{1}{2}\frac{d^2}{dx^2} - \frac{Z}{|x|}, \tag{2.1}$$

and this has been studied in great detail [69, 76–78, 81, 82]. The eigenfunctions
which are consistent with the impenetrability of the nucleus are

$$\psi_n^+(x) = \frac{2}{n}\left(\frac{Z}{n}\right)^{3/2} xL_{n-1}^{(1)}(+2Zx/n)\exp(-Zx/n), \qquad x > 0, \tag{2.2}$$

$$\psi_n^-(x) = \frac{2}{n}\left(\frac{Z}{n}\right)^{3/2} xL_{n-1}^{(1)}(-2Zx/n)\exp(+Zx/n), \qquad x < 0, \tag{2.3}$$

where $L_n^{(a)}$ is a Laguerre polynomial [100] and $n = 1, 2, 3, \ldots$. All of these vanish
at the nucleus (which is counter-intuitive) and decay exponentially at large $|x|$.
Curiously the ground-state energy of the 1D hydrogen atom is $-1/2 E_\mathrm{h}$, identical to
that of the 3D hydrogen atom. Additionally, because of nuclear impenetrability, the
ground state of the 1D H atom has a dipole moment and $\langle x \rangle = \pm 1.5$.

2.3.2   Helium-like ions

The electronic Hamiltonian of the 1D He-like ion is

$$\hat{H} = -\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) - \frac{Z}{|x_1|} - \frac{Z}{|x_2|} + \frac{1}{|x_1 - x_2|} \tag{2.4}$$

Ion	Total energy		Correlation energy				HF property	
	$-E_{\text{exact}}$	$-E_{\text{HF}}$	$-E_c^{\text{MP2}}$	$-E_c^{\text{MP3}}$	$-E_c$	$-E_c^{\text{soft}}$	Gap	$\sqrt{\langle x^2 \rangle}$
$_1\text{H}_1^-$	0.646584	0.643050	1.713	2.530	3.534	39	0.170	2.296
$_1\text{He}_1$	3.245944	3.242922	2.063	2.688	3.022	14	1.265	0.985
$_1\text{Li}_1^+$	7.845792	7.842889	2.235	2.733	2.903	8	3.200	0.628
$_1\text{Be}_1^{2+}$	14.445725	14.442873	2.335	2.747	2.851	6	5.874	0.460
$_1\text{B}_1^{3+}$	23.045686	23.042864	2.401	2.751	2.822		9.294	0.364
$_1\text{C}_1^{4+}$	33.645661	33.642859	2.447	2.752	2.802		13.463	0.301
$_1\text{N}_1^{5+}$	46.245644	46.242855	2.481	2.751	2.789		18.382	0.256
$_1\text{O}_1^{6+}$	60.845631	60.842852	2.508	2.749	2.779		24.050	0.223
$_1\text{F}_1^{7+}$	77.445621	77.442849	2.529	2.748	2.772		30.468	0.198
$_1\text{Ne}_1^{8+}$	96.045613	96.042847	2.546	2.746	2.766		37.635	0.177

Table 2.1:   Total energies (in E_h), correlation energies (in mE_h), HOMO-LUMO gaps (in E_h) and radii (in a.u.) of the 1D helium-like ions. The softened Coulomb operator of Wagner et al. [48] has been used to obtain the $-E_c^{\text{soft}}$ values.

and two families of electronic states can be considered:

- The one-sided A_2^{Z-2} family where both electrons are on the same side of the nucleus;

- The two-sided $_1\text{A}_2^{Z-2}$ family where the electrons are on opposite sides of the nucleus.

Some of the properties of the first ten ions are gathered in Table 2.1.

One-sided or two-sided?

Because of the constraints of movement in 1D, electrons shield one another very effectively and, as a result, the outer electron lies far from the nucleus in the $\text{A}_2^{(Z-2)}$ state. Because of this, the $\text{A}_2^{(Z-2)}$ state is significantly higher in energy than the $_1\text{A}_1^{(Z-2)}$ state. For example, the HF energies of He_2 and $_1\text{He}_1$ are -2.107356 and -3.242922, respectively.

In the hydride anion H^- ($Z = 1$), the nucleus cannot bind the second electron in the H_2^- state and this species autoionizes. The corresponding state of the helium atom is bound but its ionization energy is only 0.1074. Whereas the minimum nuclear charge which can bind two electrons is $Z_{\text{crit}} \approx 1.1$ in the $\text{A}_2^{(Z-2)}$ state, it is $Z_{\text{crit}} \approx 0.65$ in the $_1\text{A}_1^{(Z-2)}$ state. In comparison, Baker *et al.* have reported [101] that the corresponding value in 3D is $Z_{\text{crit}} \approx 0.91$.

In the $_1\text{A}_1^{(Z-2)}$ state, each electron is confined to one side of the nucleus, and is perfectly shielded from the other electron by the nucleus, assuming that nuclei have the same shielding characteristics as electrons. As a result, the electron correlation

energy E_c is entirely of the dispersion type and is much smaller than in 3D atoms. For example, E_c in $_1\mathrm{He}_1$ is -3.022 while E_c in the ground state of 3D He is -42.024. It is interesting to note that, unlike the situation in 3D, the correlation energy of $_1\mathrm{H}_1{}^-$ is slightly larger than in $_1\mathrm{He}_1$ and approaches the large-Z limit from below.

Table 2.1 also shows that correlation energies E_c^{soft} arising from use of the softened Coulomb operator [48] are completely different from energies E_c from the unmodified operator. This qualitative change arises because the softened operator allows the electrons to share the same orbital.

Large-Z expansion

In the large-Z (i.e. high-density) limit, the exact and HF energies of the two-sided He-like ions can be expanded as a power series using Rayleigh-Schrödinger perturbation theory [102]

$$E_{\mathrm{exact}} = E^{(0)}\, Z^2 + E^{(1)}\, Z + E^{(2)} + \frac{E^{(3)}}{Z} + O(Z^{-2}) \tag{2.5}$$

$$E_{\mathrm{HF}} = E_{\mathrm{HF}}^{(0)}\, Z^2 + E_{\mathrm{HF}}^{(1)}\, Z + E_{\mathrm{HF}}^{(2)} + \frac{E_{\mathrm{HF}}^{(3)}}{Z} + O(Z^{-2}) \tag{2.6}$$

where

$$E^{(0)} = E_{\mathrm{HF}}^{(0)} = -1 \qquad\qquad E^{(1)} = E_{\mathrm{HF}}^{(1)} = 2/5 \tag{2.7}$$

For large Z, the limiting correlation energy is thus

$$E_c = E^{(2)} - E_{\mathrm{HF}}^{(2)} + \frac{E^{(3)} - E_{\mathrm{HF}}^{(3)}}{Z} + O(Z^{-2}) = E_c^{(2)} + \frac{E_c^{(3)}}{Z} + O(Z^{-2}) \tag{2.8}$$

The second- and third-order exact energies

$$E^{(2)} = -0.045545, \qquad\qquad E^{(3)} = -0.000650 \tag{2.9}$$

can be found by Hylleraas' approach [103], while the second- and third-order HF energies

$$E_{\mathrm{HF}}^{(2)} = -0.042832 \qquad\qquad E_{\mathrm{HF}}^{(3)} = -0.000495 \tag{2.10}$$

can be found by Linderberg's method [104, 105]. We conclude, therefore, that

$$E_c = -2.713 - \frac{0.155}{Z} + O(Z^{-2}) \tag{2.11}$$
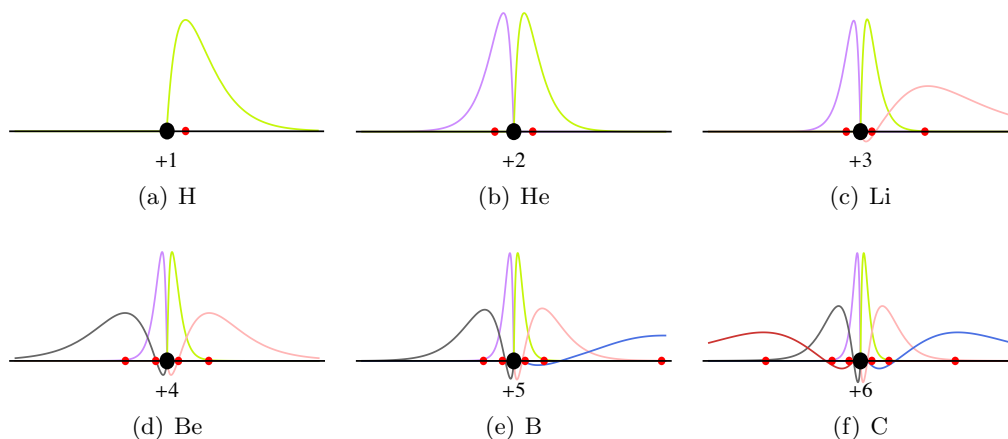
Figure 2.1:   HF ground state orbitals of the H, He, Li, Be, B and C atoms. The positions of the nucleus are represented by the black dots while the most likely position of the electrons are represented by red dots.

The negative sign of $E_{\mathrm{c}}^{(3)}$ explains the reduction in the correlation energy as Z increases.

It is interesting to note that the 2D and 3D values of $E_{\mathrm{c}}^{(2)}$ are -220.133 and -46.663, respectively [102, 106, 107], which are much larger than the corresponding 1D values.

2.3.3   Periodic Table

We have computed the ground-state energies of the 1D atoms from Li to Ne at the HF, MP2 and MP3 levels. We have also computed these energies for their cations and anions. To compute the exact energy of Li and Be^{+}, we have used a Hylleraas-type wave function containing a large number of terms. The results are reported in Table 2.2 and the HF ground state of the first six atoms are represented in Fig. 2.1.

Where exact energies are available, it appears that the MP2 and MP3 calculations recover a large proportion of the exact correlation energy. Their performance appears to improve rapidly as the atomic number grows and, for this reason, we consider the MP3 energies to be reliable benchmarks for the heavy atoms.

In view of the modest sizes of these atomic correlation energies, we conclude that it is likely that, for 1D systems, even the simple HF model is reasonably accurate and MP2 offers a very accurate theoretical model chemistry.

The accuracy of perturbative methods throughout Table 2.2 may be surprising given the small band gaps in some of the species, e.g. Li. Although a small gap is often an indicator of poor performance for perturbative corrections, the associated HOMO-LUMO excitations correspond to the movement of an electron from the

Ion	Energy		Correlation energy			HF property		
	$-E_{\text{exact}}$	$-E_{\text{HF}}$	$-E_c^{\text{MP2}}$	$-E_c^{\text{MP3}}$	$-E_{\text{c}}$	Gap	$\langle x \rangle$	$\sqrt{\langle x^2 \rangle}$
H^+	0	0	0	0	0	—	0	0
H_1	0.500000	0.500000	0	0	0	0.373	1.500	1.732
$_1H_1{}^-$	0.646584	0.643050	1.715	2.530	3.534	0.168	0	2.296
$He_1{}^+$	2.000000	2.000000	0	0	0	0.776	0.750	0.866
$_1He_1$	3.245944	3.242922	2.063	2.688	3.022	1.264	0	0.985
$_1He_2{}^-$			Autoionizes					
$_1Li_1{}^+$	7.845792	7.842889	2.235	2.733	2.903	3.200	0	0.628
$_1Li_2$	8.0119	8.007756	3.36	4.03	4.1	0.119	1.483	2.836
$_2Li_2{}^-$		8.059016	3.92	4.75		0.062	0	4.219
$_1Be_2{}^+$	15.0411	15.035639	4.77	5.48	5.5	0.220	0.829	1.599
$_2Be_2$		15.415912	6.68	7.69		0.386	0	2.111
$_2Be_3{}^-$			Autoionizes					
$_2B_2{}^+$		25.281504	8.75	9.80		0.897	0	1.437
$_2B_3$		25.357510	9.7	10.9		0.056	1.881	4.655
$_3B_3{}^-$		25.380955	9.97	11.33		0.036	0	7.042
$_2C_3{}^+$		37.918751	12.8	14.3		0.104	1.070	2.726
$_3C_3$		38.090383	14.6	16.5		0.176	0	3.684
$_3C_4{}^-$			Autoionizes					
$_3N_3{}^+$		53.528203	18.7	20.9		0.400	0	2.557
$_3N_4$		53.569533	19.1	21.5		0.031	2.423	7.139
$_4N_4{}^-$		53.582040	19.3	21.7		0.030	0	11.094
$_3O_4{}^+$		71.836884	23.8	26.6		0.059	1.382	4.267
$_4O_4$		71.929302	24.9	28.1		0.098	0	5.806
$_4O_5{}^-$			Autoionizes					
$_4F_4{}^+$		93.125365	30.5	34.2		0.217	0	4.048
$_4F_5$		93.149851	30.7	34.5		0.020	2.939	10.041
$_5F_5{}^-$		93.157319	31	35		0.037	0	15.538
$_4Ne_5{}^+$		117.256746	36.3	40.9		0.037	1.745	6.246
$_5Ne_5$		117.312529	37	42		0.067	0	8.586
$_5Ne_6{}^-$			Autoionizes					

Table 2.2: Total energies (in E_h), correlation energies (in mE_h), HOMO-LUMO gaps (in E_h), dipole moments $\langle x \rangle$ and radii $\sqrt{\langle x^2 \rangle}$ (in a.u.) of 1D atoms and ions

Atom	Ionization energy $A \longrightarrow A^+ + e^-$			Electron affinity $A + e^- \longrightarrow A^-$		
	HF	MP2	MP3	HF	MP2	MP3
H	13.606	13.606	13.606	3.893	3.939	3.961
He	33.822	33.878	33.895	0	0	0
Li	4.486	4.517	4.522	1.395	1.410	1.414
Be	10.348	10.400	10.408	0	0	0
B	2.068	2.09	2.098	0.643	0.651	0.655
C	4.670	4.719	4.733	0	0	0
N	1.125	1.14	1.14	0.340	0.35	0.35
O	2.515	2.54	2.56	0	0	0
F	0.666	0.67	0.67	0.203	0.21	0.2
Ne	1.518	1.5	1.5	0	0	0

Table 2.3: Ionization energies and electron affinities (in eV) of 1D atoms

outermost orbital on one side of the nucleus to the corresponding orbital on the other side, e.g. exciting from $_1\text{Li}_2$ to $_2\text{Li}_1$. However, such excitations are excluded from the perturbation sums because they involve the (physically forbidden) movement of an electron from one domain to another.

We have computed the ionization energy IE ($A \longrightarrow A^+ + e^-$) and the electron affinity EA ($A + e^- \longrightarrow A^-$) of each atom and these are summarised in Table 2.3. Our HF calculations revealed that anions of even-Z atoms (viz. He^-, Be^-, C^-, O^- and Ne^-) autoionize. The IEs display a clear zig-zag pattern as the atomic number grows, reminiscent of the IEs in 3D. However, in 1D the period is very short, viz. two.

The odd-Z atoms have a non-zero dipole moment, which allows reactivity with other odd-Z atoms via dipole-dipole interactions. In contrast, the even-Z atoms have only a quadrupole and would be expected to be more electrostatically inert. The combination of the periodic trends in the IEs and the pattern of atomic reactivities allows us to construct a periodic table for 1D atoms (Fig. 2.2). The 1D atoms H, Li, B, N and F are the analogs of the 3D alkali metals (i.e. H, Li, Na, K and Rb) and the 1D atoms He, Be, C, O and Ne are the analogs of the 3D noble gases (i.e. He, Ne, Ar, Kr and Xe).

Like their 3D analogs [108–113], the 1D IEs drop as the nuclear charge increases. However, this behaviour is more dramatic in 1D than in 3D because the strong shielding in 1D causes the outermost electrons to be very weakly attracted to the nucleus. This effect is so powerful that the third 1D noble gas (C) has an IE (4.733 eV) which is lower than the IE (5.139 eV) of the third 3D alkali metal (Na).

1D EAs also behave similarly to their 3D counterparts, decreasing as the nuclear charge increases. Because one side of the nucleus is completely unshielded, the EA

Figure 2.2: The periodic table in 1D

of 1D H (3.961 eV) is far larger than that of 3D H (0.754 eV). However, like the 1D IEs, shielding effects lead to a rapid reduction in EA as the nuclear charge increases. As a result, the fifth 1D alkali metal (F) has an EA (0.160 eV) which is considerably smaller than the EA (0.486 eV) of the fifth 3D alkali metal (Rb).

We have also computed $\sqrt{\langle x^2 \rangle}$ as a measure of atomic radius and compared these to the calculated values of Clementi *et al.* [114, 115] for 3D atoms. Whereas a 3D alkali metal atom is much larger than the noble gas atom of the same period, the 1D alkali metal atoms are only slightly larger than the noble gas within their period.

2.4   Molecules

2.4.1   One-electron diatomics

The electronic Hamiltonian of a one-electron diatomic $\mathrm{AB}^{(Z_A + Z_B - 1)}$ composed of two nuclei A and B of charges Z_A and Z_B located at $x = -R/2$ and $x = +R/2$ is

$$\hat{H} = -\frac{1}{2}\frac{d^2}{dx^2} - \frac{Z_A}{|x + R/2|} - \frac{Z_B}{|x - R/2|}. \tag{2.12}$$

For these systems, three families of states are of interest:

- The $_1\mathrm{AB}^{(Z_A + Z_B - 1)}$ and $\mathrm{AB}_1^{(Z_A + Z_B - 1)}$ families where the electron is outside the nuclei;[2]

---

[2] In the homonuclear case, i.e. $Z_A = Z_B$, these two families are equivalent.

	Molecule		Total energy		Correlation energy				
	State	Bond length	$-E_{\text{exact}}$	$-E_{\text{HF}}$	$-E_c^{\text{MP2}}$	$-E_c^{\text{MP3}}$	$-E_c$	Gap	ν
One-electron diatomics	H_1H^+	$R_{\text{eq}} = 2.581$	0.830710	0.830710	0	0	0	3.42	2470
	He_1H^{2+}	$R_{\text{eq}} = 2.182$	1.830303	1.830303	0	0	0	2.39	3553
		$R_{\text{ts}} = 3.296$	1.809411	1.809411	0	0	0	1.28	1914
	$\text{He}_1\text{He}^{3+}$	$R_{\text{eq}} = 1.793$	1.986928	1.986928	0	0	0	9.89	4267
		$R_{\text{ts}} = 4.630$	1.694543	1.694543	0	0	0	1.48	1028
Two-electron diatomics	H_1H_1	$R_{\text{eq}} = 2.639$	1.185948	1.184571	1.400	1.374	1.377	0.264	2389
	$_1\text{He}_1\text{H}^+$	$R_{\text{eq}} = 2.016$	3.444390	3.441957	2.457	2.438	2.433	1.220	3747
	He_1H_1^+	$R_{\text{eq}} = 2.037$	2.517481	2.516810	0.681	0.669	0.671	0.443	3939
	$\text{He}_1\text{He}_1^{2+}$	$R_{\text{eq}} = 1.668$	4.112551	4.110780	1.784	1.772	1.771	1.480	4755
		$R_{\text{ts}} = 3.989$	3.807432	3.807165	0.251	0.259	0.267	0.307	1286
Triatomics	$\text{H}_1\text{H}_1\text{H}^+$	$R_{\text{eq}} = 2.664$	1.570720	1.569820	0.918	0.897	0.900	1.557	1178[a]

[a]Symmetric vibrational mode.

Table 2.4: Structures (in a.u.), total energies (in E_h), correlation energies (in mE_h), gaps (in E_h) and vibrational frequencies ν (in cm^{-1}) of 1D molecules

- The $\text{A}_1\text{B}^{(Z_{\text{A}}+Z_{\text{B}}-1)}$ family where the electron is between the two nuclei.

Some of the properties of three such systems are reported in the upper half of Table 2.4.

H_2^+

The simplest of all molecules is the homonuclear diatomic H_2^+, in which $Z_{\text{A}} = Z_{\text{B}} = 1$. In 3D, this molecule was first studied by Burrau who pointed out that the Schrödinger equation is separable in confocal elliptic coordinates [116]. In 1928, Linus Pauling published a review summarizing the work of Burrau and many other researchers [117, 118].

Fortunately, it is possible to obtain some exact wave functions for H_1H^+ in 1D. The Schrödinger equation using the Hamiltonian in Eq. (2.12) can be solved for $Z_{\text{A}} = Z_{\text{B}}$ and $E = 0$, yielding

$$\psi_n(x) = (1-z^2) \begin{cases} x\,F\left(-\frac{n-1}{2}, \frac{n+4}{2}, 2, 1-z^2\right), & n \text{ odd}, \\ F\left(-\frac{n}{2}, \frac{n+3}{2}, 2, 1-z^2\right), & n \text{ even}, \end{cases} \tag{2.13}$$

where $F(a,b,c,x)$ is the Gauss hypergeometric function [100], $z = 2x/R$ and

$$R = \frac{(n+1)(n+2)}{2Z}. \tag{2.14}$$
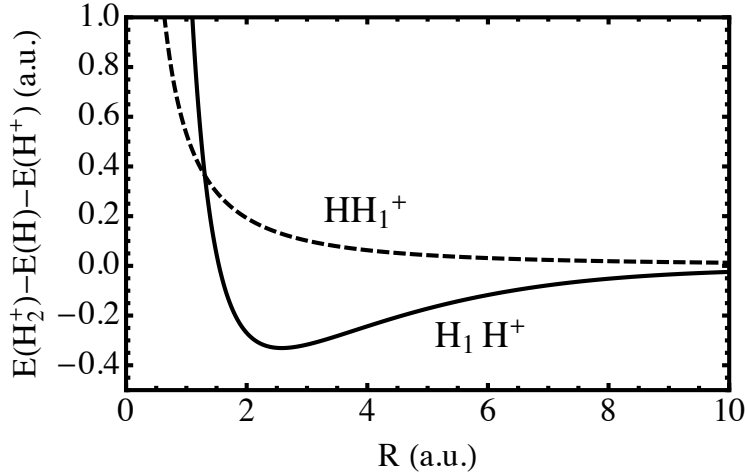
Figure 2.3: Potential energy curves of the H_1H^+ and HH_1^+ states of the hydrogen molecule cation

Hence for the H_2^+ molecule with, for example, bond length $R = 1$ the exact wave function will be $\psi_0(x) = (1 - 2x)(1 + 2x)$.

The near-exact potential energy curves of the H_1H^+ and HH_1^+ states are shown in Fig. 2.3. Beyond $R = 1.5$, the H_1H^+ state is lower in energy than the HH_1^+ state. However, when the bond is compressed, the kinetic energy of the trapped electron becomes so large that the H_1H^+ state rises above the HH_1^+ state. The bond dissociation energy ($0.3307\ E_h$) of H_1H^+ is large and its equilibrium bond length ($R_{eq} = 2.581$ bohr) is long. Both values are much larger than the corresponding 3D values ($0.1026\ E_h$ and 1.997 bohr) [119]. Whereas the H_1H^+ state is bound by a favorable charge-dipole interaction, the HH_1^+ state is repulsive because of a similar, but unfavorable, interaction. Using this simple electrostatic argument, one can predict that the H_1H^+ and HH_1^+ potential energy curves behave as $-\mu_H/R^2$ and $+\mu_H/R^2$ for large R, where $\mu_H = 3/2$. This charge-dipole model is qualitatively correct for $R \gtrsim 10$ for H_1H^+ and $R \gtrsim 5$ for HH_1^+.

HeH^{2+} and He$_2^{3+}$

The Hamiltonians of HeH^{2+} and He_2^{3+} are given by (2.12) for $Z_A = 1$ and $Z_B = 2$, and $Z_A = Z_B = 2$, respectively. As in H_2^+, we find that He_1He^{3+} is more stable than $HeHe_1^{3+}$, and He_1H^{2+} is more stable than HeH_1^{2+} and $_1HeH^{2+}$, except at short bond lengths.

In 3D, the molecules HeH^{2+} and He_2^{3+} are unstable except in strong magnetic fields [120]. However, as Fig. 2.4 shows, He_1H^{2+} and He_1He^{3+} are metastable species in 1D with equilibrium bond lengths of $R_{eq} = 2.182$ and 1.793, and transition structure bond lengths of $R_{ts} = 3.296$ and 4.630, respectively. Although these species
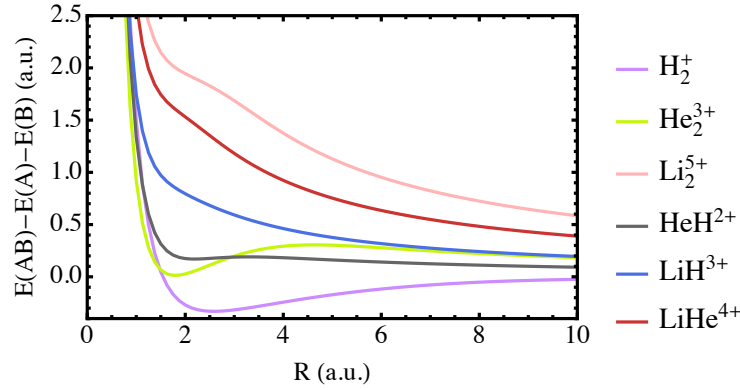
Figure 2.4:   Potential energy curves of the $A_1B^{(Z_A+Z_B-1)}$ states of several one-electron diatomics

are thermodynamically unstable with respect to $He^+ + H^+$ and $He^+ + He^{2+}$, they are protected from dissociation by barriers of 0.0209 and 0.2924, respectively. For large R, their dissociation curves behave as $1/R - \mu_{He^+}/R^2$ and $2/R - 2\mu_{He^+}/R^2$, respectively, where $\mu_{He^+} = 3/4$.

All the heavier one-electron diatomics have purely repulsive dissociation curves.

Chemical bonding in one-electron diatomics

Fig. 2.5 shows the electronic density $\rho(x)$ for H_1H^+ and He_1H^{2+} at their equilibrium bond lengths. Whereas the electron density in a typical 3D bond is greatest at the nuclei and reaches a minimum near the middle of the bond [118], the electron density in these 1D bonds vanishes at the nuclei and achieves a maximum in the middle of the bond. The bond in He_1H^{2+} is polarized towards the nucleus with the largest charge.

Harmonic vibrations

We have computed the harmonic vibrational frequencies of H_1H^+, He_1H^{2+} and He_1He^{3+} at their equilibrium bond lengths and these are shown in Table 2.4. The second derivative of the energy was obtained numerically using the three-point central difference formula and a stepsize of 10^{-2} bohr. The frequency of the 1D H_1H^+ ion (2470 cm^{-1}) is similar to that of the 3D ion (2321 cm^{-1}) [121] but this result is probably accidental. Although the barrier in He_1H^{2+} is small and its harmonic frequency relatively high (3553 cm^{-1}), the ion probably supports a vibrational state: the zero-point vibrational energy is only half the barrier height.
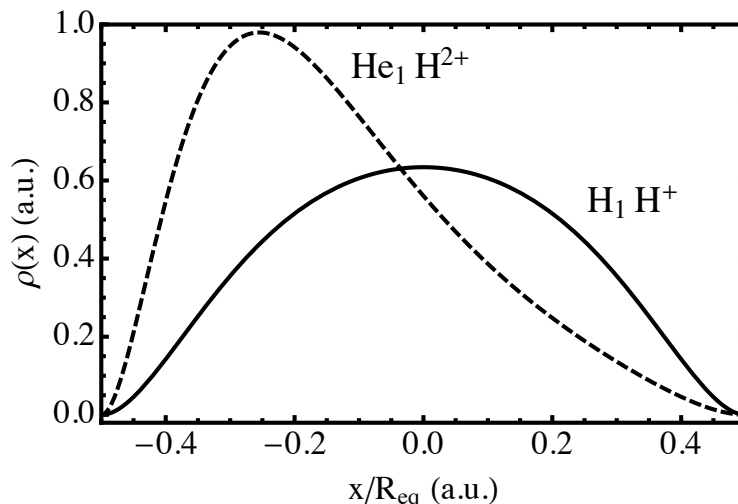
Figure 2.5: Electronic density $\rho(x)$ in H_1H^+ and He_1H^{2+} at their equilibrium bond lengths

2.4.2   Two-electron diatomics

The Hamiltonian of a two-electron diatomic $AB^{(Z_A+Z_B-2)}$ composed of two nuclei A and B of charges Z_A and Z_B located at $x = -R/2$ and $x = +R/2$ is

$$\hat{H} = -\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) - \frac{Z_A}{\left|x_1 + \frac{R}{2}\right|} - \frac{Z_A}{\left|x_2 + \frac{R}{2}\right|} - \frac{Z_B}{\left|x_1 - \frac{R}{2}\right|} - \frac{Z_B}{\left|x_2 - \frac{R}{2}\right|} + \frac{1}{\left|x_1 - x_2\right|}. \tag{2.15}$$

These systems possess six families of states:

- The $A_2B^{(Z_A+Z_B-2)}$ family;

- The $_1AB_1^{(Z_A+Z_B-2)}$ family;

- The $_1A_1B^{(Z_A+Z_B-2)}$ and $A_1B_1^{(Z_A+Z_B-2)}$ families;

- The $_2AB^{(Z_A+Z_B-2)}$ and $AB_2^{(Z_A+Z_B-2)}$ families;

Note that the last two items of the list are equivalent in the homonuclear case (i.e. $Z_A = Z_B$). Some of the properties of four such systems are reported in the lower half of Table 2.4.

H_2

The simplest two-electron diatomic is H_2 where $Z_A = Z_B = 1$. The 3D version of this molecule has been widely studied since the first accurate calculation of James and Coolidge [98] in 1933. The 1D ground state in each family has been calculated using Hylleraas-type calculations and is represented in Fig. 2.6. We note that the
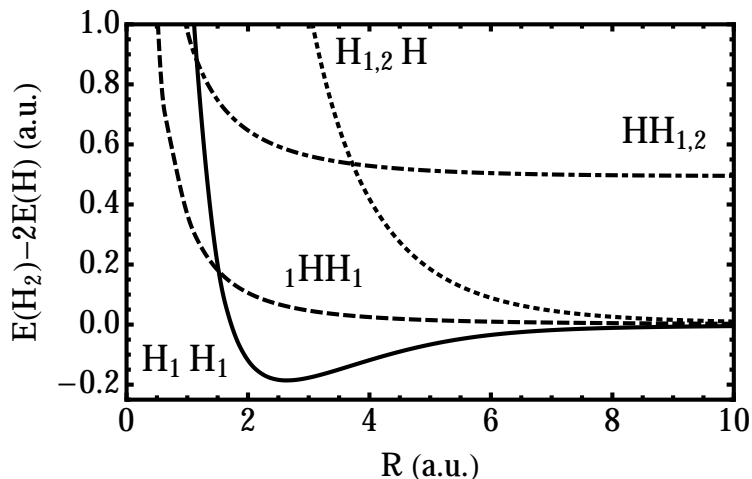
Figure 2.6: Potential energy curves of the H_1H_1, $_1HH_1$, H_2H and HH_2 states of the hydrogen molecule.

HF and Hylleraas curves are almost indistinguishable due to the small correlation energy in these systems (see Table 2.4).

As expected, HH_2 is high in energy due to shielding by the inner electron (see discussion on the He-like ions in Sec. 2.3.2), and dissociates into $H^+ + H_2^-$. The three other states dissociate into a pair of H atoms. As in H_2^+, the $_1HH_1$ state is the most stable at small bond lengths, but is higher in energy than H_1H_1 when $R > 1.5$ bohr. The H_1H_1 state is bound with an equilibrium bond length of 2.639 bohr and a dissociation energy of 0.1859 E_h. In comparison, the bond length of the 3D H_2 molecule is close to 1.4 bohr and has a similar dissociation energy (0.1745 E_h) [122]. The harmonic vibrational frequency of H_1H_1 (2389 cm^{-1}) is significantly lower than the 3D value (4401 cm^{-1}) [121]. The equilibrium bond lengths and vibrational frequencies of H_1H^+ and H_1H_1 are similar because of the efficient shielding in 1D. Finally, we note that H_1H_1 has a non-zero dipole moment and the two fragments H_1 are bound by a dipole-dipole interaction.

For those who are familiar with the traditional covalent two-electron bond in 3D chemistry, the instability of H_2H is probably surprising. However, this state is destabilized by two important effects: (a) the high kinetic energy of the electrons when trapped between nuclei (see discussion on H_2^+ in Sec. 2.4.1) and (b) the 1D exclusion principle, which mandates that the second electron occupy a higher-energy orbital than the first. For these reasons, 1D molecules are usually held together by one-electron bonds (sometimes called hemi-bonds).

Bonding in H_2^+, which is driven by the $H^+ + H$ charge-dipole interaction, is roughly twice as strong as the bonding in H_2, which arises from the much weaker
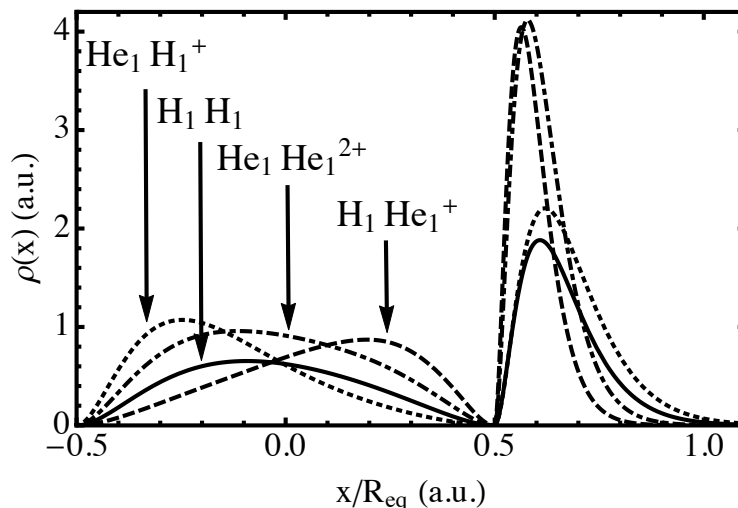
Figure 2.7: Electronic density $\rho(x)$ in H_1H_1, $H_1He_1^+$, $He_1H_1^+$ and $He_1He_1^{2+}$ at their equilibrium bond lengths

H + H dipole-dipole interaction. In contrast, in 3D, the H_2 bond is roughly twice as strong as that in H_2^+.

We expect that two-electron (or more) bonds exist in neutral species such as $1\,Li_2H_1$ because of favorable dipole-dipole interactions. However, such species are bound *despite* the two-electron bond, rather than because of it, and are probably very weakly bound. This will be explored further in the next chapter.

HeH^+ and He_2^{2+}

The Hamiltonian for HeH^+ and He_2^{2+} are given by (2.15) for $Z_A = 1$ and $Z_B = 2$, and $Z_A = Z_B = 2$, respectively. Like He_1He^{3+}, $He_1He_1^{2+}$ is metastable with a large energy barrier of 0.3051 E_h and a late transition structure with $R_{ts}/R_{eq} \approx 2.5$. In 3D, the He_2^{2+} dication is also metastable but with an earlier transition structure ($R_{ts}/R_{eq} \approx 1.5$) [123–127].

Like the 3D HeH^+ molecule [128], the 1D $_1He_1H^+$ and $He_1H_1^+$ ions are bound. The dissociation of $_1He_1H^+$ into $_1He_1 + H^+$ requires 0.1981 E_h and is much more endothermic than the dissociation of $He_1H_1^+$ into $He_1^+ + H_1$, which requires only 0.0174 E_h. Surprisingly, however, they have similar bond lengths and harmonic frequencies.

Chemical bonding in two-electron diatomics

Fig. 2.7 shows the electronic densities $\rho(x)$ in H_1H_1, $H_1He_1^+$, $He_1H_1^+$ and $He_1He_1^{2+}$ at their respective equilibrium bond lengths. The bonds in H_1H_1 and $He_1He_1^{2+}$ are polar because of the repulsion by the external electron. In $He_1H_1^+$, the bond

is highly polar because the repulsion by the external electron and the attraction of the He nucleus push in the same direction. In $H_1He_1{}^+$, the bond is polarized in the opposite direction because the repulsion by the external electron is dominated by the attraction of the He nucleus.

Correlation effects

Table 2.4 reports the MP2, MP3 and exact correlation energies at the equilibrium geometries of H_1H_1, $_1He_1H^+$, $He_1H_1{}^+$ and $He_1He_1{}^{2+}$. All these values are small compared to their 3D analogs because correlation energy in these 1D systems is entirely due to dispersion. As a result, correlation effects are pleasingly small and, for example, the HF bond length in H_1H_1 differs from the exact value by only 0.003 bohr. This re-emphasizes that the HF approximation is probably significantly more accurate in 1D than in 3D.

The range of E_c values (-2.434 in $H_1He_1{}^+$, -1.771 in $He_1He_1{}^{2+}$, -1.377 in H_1H_1, -0.671 in $He_1H_1{}^+$) can be rationalized by comparing the distance between the two electrons in each system (see Fig. 2.7): shorter distances yield larger correlation energies.

For the diatomics in Table 2.4, HF theory is again found to be accurate and the MPn series appears to converge rapidly towards the exact correlation energies. In particular, the MP3 and exact energies differ by only a few microhartrees.

2.4.3 Chemistry of $H_3{}^+$

The 3D $H_3{}^+$ ion was discovered by Thomson [129] in 1911 and plays a central role in interstellar chemistry [130–132]. In astrochemistry, the main pathway for its production is

$$H_2{}^+ + H_2 \longrightarrow H_3{}^+ + H \tag{2.16}$$

and this reaction is highly exothermic ($\Delta U = -0.0639\ E_h$) [132]. In 3D, the ion has a triangular structure [133] as first demonstrated by Coulson [134]. The proton affinity of H_2

$$H_2 + H^+ \longrightarrow H_3{}^+ \tag{2.17}$$

is also strongly exothermic ($\Delta U = -0.1613\ E_h$) [135].

In this Section, we study the 1D analogs of these two reactions, viz.

$$H_1H^+ + {}_1H_1H \longrightarrow H_1H_1H^+ + {}_1H \tag{2.18}$$

$$H_1H_1 + H^+ \longrightarrow H_1H_1H^+ \tag{2.19}$$
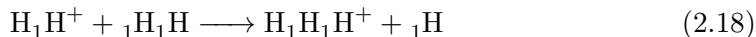
In 1D, the equilibrium structure of $H_1H_1H^+$ has $D_{\infty h}$ symmetry, a bond length of 2.664 bohr, and an energy of -1.570720 E_h (see Table 2.4). The correlation energy at this bond length is only 0.900 mE_h. Our calculations predict that reactions (2.18) and (2.19) are both exothermic ($\Delta U = -0.0541$ and -0.3848 E_h, respectively) and that reaction (2.19) is barrierless. It is interesting that the exothermicities of reactions (2.16) and (2.18) are close, and that the proton affinities (reactions (2.17) and (2.19)) are also broadly similar.

2.4.4   Hydrogen nanowire

Despite the fact that equi-spaced infinite H chain in 3D suffers from a Peierls instability [136], where the geometry distorts such that the atoms pair into dimers, this system has attracted considerable interest due to its strong correlation character and metal-insulator transition [137–141]. We have therefore used periodic HF calculations [142, 143] to compute the energy per atom of an infinite chain of equi-spaced 1D H atoms separated by a distance R. Motivated by our results for 1D H_2^+, H_2 and H_3^+, we have studied the state in which one electron is trapped between each pair of nuclei, i.e. $\cdots H_1H_1H_1H_1\cdots$.

We have expanded the HF orbital in the unit cell ($x \in [-R/2, R/2]$) as a linear combination of K even polynomials

$$\mathcal{E}_k^{AB}(x) = \sqrt{\frac{2/\pi^{1/2}}{R_{AB}}\frac{\Gamma(2k+3/2)}{\Gamma(2k+1)}}\,(1-z^2)^k \tag{2.20}$$

We find that, near the minimum-energy structure, $K = 4$ suffices to achieve convergence of the HF energy to within one microhartree and the resulting bond length is $R_{eq} = 2.763$, which is slightly longer than the values in H_2^+, H_2 and H_3^+. The corresponding energy is -0.734337 which yields a binding energy of 0.2343 per bond. In comparison, the binding energy in H_2 is roughly 80% of this value. This explains the particular stability of the equally-spaced H_∞ chain in 1D.

2.5   Concluding Remarks

We have studied the electronic structure of 1D chemical systems in which all nuclei and electrons are constrained to remain on a line. We have used the full Coulomb operator and our numerical results are strikingly different from those of previous studies [48, 49] in which a softened operator was used. We have explored atoms with up to 10 electrons, one- and two-electron diatomics, the chemistry of H_3^+ and an infinite chain of H atoms.

We find that, whereas atoms with odd numbers of electrons have non-vanishing dipole moments and are reactive, atoms with even numbers of electrons have zero dipole moments and are inert. Based on these results, we have concluded that the 1D version of the periodic table has only two groups: alkali metals and noble gases.

Our study of one- and two-electron diatomics has revealed that atoms in 1D are bound together by strong one-electron bonds. The Coulombic forces within such bonds can be accurately modelled by simple classical electrostatics, primarily as charge-dipole and dipole-dipole interactions. This leads to a variety of unexpected results, such as the discovery that the bond in H_2^+ is much stronger than the bond in H_2.

Chapter 3

Chem1D: a 1D electronic structure theory program

3.1 Introduction

In Chapter 2 we explored a model of one-dimensional (1D) chemistry for a small variety of atoms and molecules. This model applies the Dirichlet boundary condition to the wavefunction which requires a node whenever two particles touch [80, 93, 94]. There are three curious consequences of this condition: spin-blindness [62–64, 66], the Super-Pauli principle and particle impenetrability [80, 82, 87]. Section 2.1 has more details on both the use of the Dirichlet condition and its consequences.

Although limited in scope, this study was sufficient to come to some exciting conclusions, such as constructing the 1D periodic table in Sec. 2.3.3, and conjecture about the existence of larger molecules in Sec. 2.4.2. The technology used to achieve this was prototype in nature, and severely limited in the size of system which it was capable of handling. In this chapter we describe a program which we call CHEM1D. This is an electronic structure program which is capable of performing Hartree-Fock (HF) and Møller-Plesset (MP2 and MP3) calculations for (in principle) any 1D molecular system.

Due to the particle impenetrability in the Dirichlet model of 1D chemistry the nuclei separate the 1D space into regions which the electrons become trapped in. We refer to these regions as domains and one of CHEM1D's key features is its ability to exploit this separation. In this chapter we describe CHEM1D before using it to study larger chemical systems than were possible in Chapter 2. In Sec. 3.2 we describe the underlying theory and, particularly, our two-electron integral methodology. Sec. 3.3 sketches the overall structure of our program and Sec. 3.4 presents and discusses some

new results that it has yielded. We make continued use of the notation introduced in Sec. 2.2.1.

3.2   Theory

3.2.1   Basis sets

For this work we have developed a set of basis functions to describe the orbitals. The choices have been motivated by the exact wavefunctions found in Chapter 2 for the one electron systems H_1 and H_1H^+. A basis function $\mathbf{F}_\mu^p$ has index μ and resides entirely in domain p. When the domain superscript is redundant it is omitted. It is easy to see that the basis function pair $(\mathbf{F}_\mu^p \mathbf{F}_\nu^q|$, with $p \neq q$, must necessarily vanish everywhere.

In domain 0 we have a set of normalised exponentials

$$\mathcal{L}_\mu(x) = 2\mu^3 \alpha^{3/2} (A_1 - x) \exp[-\mu^2 \alpha (A_1 - x)] \tag{3.1}$$

in domain N we have an analogous set of exponentials

$$\mathcal{R}_\mu(x) = 2\mu^3 \alpha^{3/2} (x - A_N) \exp[-\mu^2 \alpha (x - A_N)] \tag{3.2}$$

and, in the finite domains, we have even polynomials

$$\mathcal{E}_\mu^p(x) = \sqrt{\frac{2/\pi^{1/2}}{R} \frac{\Gamma(2\mu + 3/2)}{\Gamma(2\mu + 1)}} \, (1 - z^2)^\mu \tag{3.3}$$

and odd polynomials

$$\mathcal{O}_\mu^p(x) = \sqrt{\frac{4/\pi^{1/2}}{R} \frac{\Gamma(2\mu + 5/2)}{\Gamma(2\mu + 1)}} \, z(1 - z^2)^\mu \tag{3.4}$$

where $z = (A_p + A_{p+1} - 2x)/(A_p - A_{p+1})$, $R = A_{p+1} - A_p$ and Γ is the Gamma function. We include only positive integer μ to ensure that the orbitals vanish at the nuclei.

All of the necessary one-electron integrals can be found in closed form and we list the relevant formulae below. We also present a discussion of the algorithms we have developed for evaluating the required two-electron quantities. We use chemist's notation [8] throughout.

3.2.2 One electron integrals

Define $\hat{T} = -\nabla^2/2$, $\hat{V}_p = |x - A_p|^{-1}$, $\zeta = \mu^2 + \nu^2$, $\eta = \lambda^2 + \sigma^2$ and

$$G(x, y) = \frac{\Gamma(x + y)}{\sqrt{\Gamma(2x)\,\Gamma(2y)}} \tag{3.5}$$

Then the necessary one electron integrals, which are relatively straightforward to evaluate, are given by the following expressions.

Overlap integrals

All overlap integrals vanish except the following

$$(\mathcal{L}_\mu|\mathcal{L}_\nu) = (\mathcal{R}_\mu|\mathcal{R}_\nu) = (2\mu\nu/\zeta)^3 \tag{3.6}$$

$$(\mathcal{E}_\mu^p|\mathcal{E}_\nu^p) = \frac{G(\mu + 1/2, \nu + 1/2)}{G(\mu + 3/4, \nu + 3/4)} \tag{3.7}$$

$$(\mathcal{O}_\mu^p|\mathcal{O}_\mu^p) = \frac{G(\mu + 1/2, \nu + 1/2)}{G(\mu + 5/4, \nu + 5/4)} \tag{3.8}$$

Kinetic integrals

All kinetic energy integrals vanish except the following

$$\frac{(\mathcal{L}_\mu|\hat{T}|\mathcal{L}_\nu)}{(\mathcal{L}_\mu|\mathcal{L}_\nu)} = \frac{(\mathcal{R}_\mu|\hat{T}|\mathcal{R}_\nu)}{(\mathcal{R}_\mu|\mathcal{R}_\nu)} = \frac{\alpha^2 \mu^2 \nu^2}{2} \tag{3.9}$$

$$\frac{(\mathcal{E}_\mu^p|\hat{T}|\mathcal{E}_\nu^p)}{(\mathcal{E}_\mu^p|\mathcal{E}_\nu^p)} = \frac{4\mu\nu(\mu + \nu + 1/2)}{(\mu + \nu)(\mu + \nu - 1)}\frac{1}{R^2} \tag{3.10}$$

$$\frac{(\mathcal{O}_\mu^p|\hat{T}|\mathcal{O}_\nu^p)}{(\mathcal{O}_\mu^p|\mathcal{O}_\nu^p)} = \frac{12\mu\nu(\mu + \nu + 3/2)}{(\mu + \nu)(\mu + \nu - 1)}\frac{1}{R^2} \tag{3.11}$$

where $R = A_{p+1} - A_p$.

Nuclear-attraction integrals

All nuclear-attraction integrals vanish except

$$\frac{(\mathcal{L}_\mu|\hat{V}_p|\mathcal{L}_\nu)}{(\mathcal{L}_\mu|\mathcal{L}_\nu)} = \frac{(\zeta\alpha R_p)^3}{R_p}U(3, 3, \zeta\alpha R_p) \qquad (R_p = A_p - A_1) \tag{3.12}$$

$$\frac{(\mathcal{R}_\mu|\hat{V}_p|\mathcal{R}_\nu)}{(\mathcal{R}_\mu|\mathcal{R}_\nu)} = \frac{(\zeta\alpha R_p)^3}{R_p}U(3, 3, \zeta\alpha R_p) \qquad (R_p = A_N - A_p) \tag{3.13}$$

$$\frac{(\mathcal{E}_\mu^q|\hat{V}_p|\mathcal{E}_\nu^q)}{(\mathcal{E}_\mu^q|\mathcal{E}_\nu^q)} = \frac{1}{|R_p|}F\left(1,\frac{1}{2},\mu+\nu+\frac{3}{2},\frac{a^2}{R_p^2}\right) \qquad (R_p = |A_q + a - A_p|) \qquad (3.14)$$

$$\frac{(\mathcal{O}_\mu^q|\hat{V}_p|\mathcal{O}_\nu^q)}{(\mathcal{O}_\mu^q|\mathcal{O}_\nu^q)} = \frac{1}{|R_p|}F\left(1,\frac{3}{2},\mu+\nu+\frac{5}{2},\frac{a^2}{R_p^2}\right) \qquad (R_p = |A_q + a - A_p|) \qquad (3.15)$$

$$\frac{(\mathcal{E}_\mu^q|\hat{V}_p|\mathcal{O}_\nu^q)}{(\mathcal{O}_\mu^q|\mathcal{O}_\nu^q)} = \frac{1}{|R_p|}F\left(1,\frac{3}{2},\mu+\nu+\frac{5}{2},\frac{a^2}{R_p^2}\right)\frac{a/R_p}{\sqrt{4\mu+3}} \qquad (R_p = |A_q + a - A_p|)$$
$$(3.16)$$

where U is the confluent hypergeometric function [100], F is the Gauss hypergeometric function [100] and $a = (A_{q+1} - A_q)/2$.

3.2.3   Two-domain two-electron integrals

Antisymmetrised ("double bar") electron repulsion integrals (ERIs) quantify interactions between an electron in domain p and another in domain q. If $p \neq q$, we have

$$(\mathbf{F}_\mu^p \mathbf{F}_\nu^p||\mathbf{F}_\lambda^q \mathbf{F}_\sigma^q) = (\mathbf{F}_\mu^p \mathbf{F}_\nu^p|\mathbf{F}_\lambda^q \mathbf{F}_\sigma^q) - (\mathbf{F}_\mu^p \mathbf{F}_\lambda^q|\mathbf{F}_\nu^p \mathbf{F}_\sigma^q)$$
$$= (\mathbf{F}_\mu^p \mathbf{F}_\nu^p|\mathbf{F}_\lambda^q \mathbf{F}_\sigma^q) \qquad (3.17)$$

and there are three cases: (1) Both domains are infinite; (2) One domain is infinite and the other is finite; (3) Both domains are finite.

Infinite/Infinite domains

If we let $\zeta = \mu^2 + \nu^2$, $\eta = \lambda^2 + \sigma^2$ and $R = A_N - A_1$ then

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu|\mathcal{R}_\lambda \mathcal{R}_\sigma)}{(\mathcal{L}_\mu|\mathcal{L}_\nu)(\mathcal{R}_\lambda|\mathcal{R}_\sigma)} = \frac{\alpha\zeta\eta}{2(\zeta-\eta)^5}\Bigg($$
$$\frac{1}{2}\left[5(\zeta+\eta)(\zeta-\eta)^3 - 3(\zeta+\eta)^3(\zeta-\eta) - 2R\alpha\zeta\eta(\zeta-\eta)^3\right]$$
$$- \zeta^2\eta^2\left[12 - 6R\alpha(\zeta-\eta) + (R\alpha(\zeta-\eta))^2\right]U(1,1,R\alpha\zeta)$$
$$+ \zeta^2\eta^2\left[12 + 6R\alpha(\zeta-\eta) + (R\alpha(\zeta-\eta))^2\right]U(1,1,R\alpha\eta)\Bigg) \quad (3.18)$$

or, in the limiting case where $\zeta = \eta$,

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu|\mathcal{R}_\lambda \mathcal{R}_\sigma)}{(\mathcal{L}_\mu|\mathcal{L}_\nu)(\mathcal{R}_\lambda|\mathcal{R}_\sigma)} = \frac{\alpha\zeta}{120}\left(24 - 6R\alpha\zeta + 2(R\alpha\zeta)^2 - (R\alpha\zeta)^3\right.$$
$$\left. + (R\alpha\zeta)^4 - (R\alpha\zeta)^5 U(1,1,R\alpha\zeta)\right) \quad (3.19)$$

When $R = 0$ (*i.e.* for atomic calculations), these two formulae reduce to

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{R}_\lambda \mathcal{R}_\sigma)}{(\mathcal{L}_\mu | \mathcal{L}_\nu)(\mathcal{R}_\lambda | \mathcal{R}_\sigma)} = \frac{\alpha \zeta \eta}{2(\zeta - \eta)^5} \left((\zeta^2 - \eta^2)(\zeta^2 - 8\zeta\eta + \eta^2) + 12\zeta^2 \eta^2 \ln \left(\frac{\zeta}{\eta} \right) \right) \quad (3.20)$$

and

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{R}_\lambda \mathcal{R}_\sigma)}{(\mathcal{L}_\mu | \mathcal{L}_\nu)(\mathcal{R}_\lambda | \mathcal{R}_\sigma)} = \frac{\alpha \zeta}{5} \quad (3.21)$$

Infinite/Finite domains

Unfortunately, we have not been able to obtain analytic and numerically satisfactory expressions for $(\mathcal{LL}|\mathcal{EE})$ integrals and others of this type. Fortunately, there are comparatively few of these integrals and if we define $a = (A_{p+1} - A_p)/2$, $R = (A_p + a) - A_1$, $\zeta = \mu^2 + \nu^2$ and $m = \lambda + \sigma$ and reformulate the desired ERIs in terms of the potential of the $\mathcal{LL}$ product, we obtain

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{E}_\lambda^p \mathcal{E}_\sigma^p)}{(\mathcal{L}_\mu | \mathcal{L}_\nu)(\mathcal{E}_\lambda^p | \mathcal{E}_\sigma^p)} = \frac{\zeta^3}{a^{2m+1}\sqrt{\pi}} \frac{\Gamma(m + 3/2)}{\Gamma(m + 1)}$$
$$\int_{R-a}^{R+a} (a^2 + (R - x)^2)^m x^2 U(3, 3, \zeta x) \, dx \quad (3.22)$$

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{E}_\lambda^p \mathcal{O}_\sigma^p)}{(\mathcal{L}_\mu | \mathcal{L}_\nu)(\mathcal{O}_\lambda^p | \mathcal{O}_\sigma^p)} = -\frac{2\zeta^3}{a^{2m+2}\sqrt{\pi}\sqrt{3 + 4\lambda}} \frac{\Gamma(m + 5/2)}{\Gamma(m + 1)}$$
$$\int_{R-a}^{R+a} (a^2 + (R - x)^2)^m (R - x) x^2 U(3, 3, \zeta x) \, dx \quad (3.23)$$

$$\frac{(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{O}_\lambda^p \mathcal{O}_\sigma^p)}{(\mathcal{L}_\mu | \mathcal{L}_\nu)(\mathcal{O}_\lambda^p | \mathcal{O}_\sigma^p)} = \frac{2\zeta^3}{a^{2m+3}\sqrt{\pi}} \frac{\Gamma(m + 5/2)}{\Gamma(m + 1)}$$
$$\int_{R-a}^{R+a} (a^2 + (R - x)^2)^m (R - x)^2 x^2 U(3, 3, \zeta x) \, dx \quad (3.24)$$

We have found that these can be evaluated satisfactorily by numerical quadrature. The $(\mathcal{RR}|\mathcal{EE})$ integrals and others can be found in the same way.

It is possible to obtain the $(\mathcal{L}_\mu \mathcal{L}_\nu | \mathcal{O}_\lambda^p \mathcal{O}_\sigma^p)$ integral by using the relation

$$\frac{(\mathcal{O}_\mu^p \mathcal{O}_\nu^p|}{(\mathcal{O}_\mu^p | \mathcal{O}_\nu^p)} = (3 + 2(\mu + \nu)) \frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p|}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)} - (2 + 2(\mu + \nu)) \frac{(\mathcal{E}_\mu^p \mathcal{E}_{\nu+1}^p|}{(\mathcal{E}_\mu^p | \mathcal{E}_{\nu+1}^p)} \quad (3.25)$$

but we do not make use of this here.

Finite/Finite domains

The final class of two-domain ERIs are those between two finite domains. We have constructed a four-term recurrence relation for the calculation of these integrals. Let $f^p = f(x - P)$ be a polynomial basis function pair in the domain p and centred

around the point P. Let $g^q = g(y - Q)$ be another function pair in the domain q centred around the point Q. Then we have the following

$$
\begin{aligned}
(f^p | g^q) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f(x - P)g(y - Q)}{y - x} dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - P) \left(\int_0^{\infty} e^{-s(y-x)} ds \right) g(y - Q) dx dy \\
&= \int_0^{\infty} \left(\int_{-\infty}^{\infty} f(x - P) e^{sx} dx \right) \left(\int_{-\infty}^{\infty} g(y - Q) e^{-sy} dy \right) ds \\
&= \int_0^{\infty} \mathcal{L}_f(-s) \mathcal{L}_g(s) e^{-(Q-P)s} ds
\end{aligned}
\tag{3.26}
$$

where $\mathcal{L}_f$ and $\mathcal{L}_g$ are the Laplace transforms of the basis functions pairs f^p and g^q translated so that they are centred around 0 respectively. Let $a = (A_{p+1} - A_p)/2$ be half the width of the domain p. The required Laplace transforms are

$$
\frac{\mathcal{L}[\mathcal{E}_\mu^p \mathcal{E}_\nu^p](s)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)} = {_0F_1}\left(\mu + \nu + \frac{3}{2}; \left(\frac{as}{2} \right)^2 \right)
\tag{3.27}
$$

$$
\frac{\mathcal{L}[\mathcal{E}_\mu^p \mathcal{O}_\nu^p](s)}{(\mathcal{O}_\mu^p | \mathcal{O}_\nu^p)} = -\frac{as}{\sqrt{4\mu + 3}} {_0F_1}\left(\mu + \nu + \frac{3}{2}; \left(\frac{as}{2} \right)^2 \right)
\tag{3.28}
$$

where $_0F_1(c; x)$ is the confluent hypergeometric limit function, which obeys the following three-term recurrence relation [100]

$$
{_0F_1}(c - 1; x) - {_0F_1}(c; x) = \frac{x}{c(c - 1)} {_0F_1}(c + 1; x)
\tag{3.29}
$$

Since we are interested in products of the functions in Eqs. (3.27) and (3.28), both of which contain the same argument to the hypergeometric function, we are able to apply the recurrence relation (3.29) twice and remove the additional argument dependence. Let $m = \mu + \nu$, $n = \lambda + \sigma$, $a = (A_{p+1} - A_p)/2$ and $b = (A_{q+1} - A_q)/2$, then

$$
\begin{aligned}
{_0F_1}\left(m - \frac{1}{2}; \frac{a^2 s^2}{4} \right) {_0F_1}\left(n + \frac{3}{2}; \frac{b^2 s^2}{4} \right) &= {_0F_1}\left(m + \frac{1}{2}; \frac{a^2 s^2}{4} \right) {_0F_1}\left(n + \frac{3}{2}; \frac{b^2 s^2}{4} \right) \\
&+ \frac{a^2 (n - \frac{1}{2})(n + \frac{1}{2})}{b^2 (m - \frac{1}{2})(m + \frac{1}{2})} {_0F_1}\left(m + \frac{3}{2}; \frac{a^2 s^2}{4} \right) \\
&\times \left[{_0F_1}\left(n - \frac{1}{2}; \frac{b^2 s^2}{4} \right) - {_0F_1}\left(n + \frac{1}{2}; \frac{b^2 s^2}{4} \right) \right]
\end{aligned}
\tag{3.30}
$$
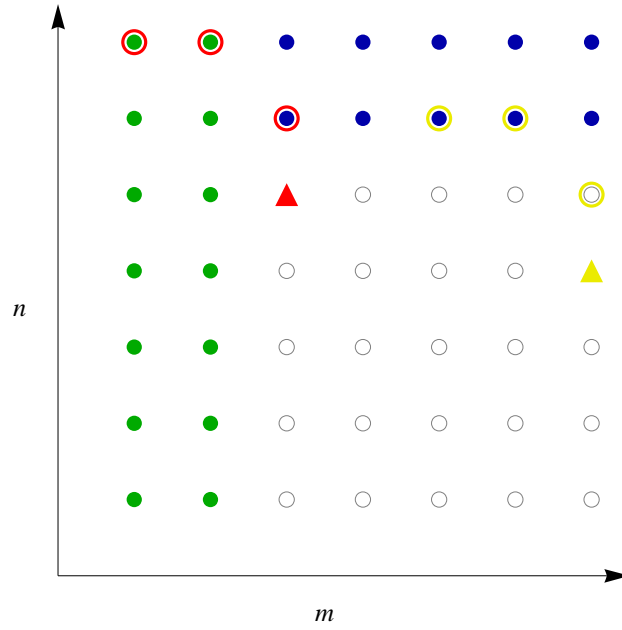
Figure 3.1: A graphical representation of the recurrence relation for computing electron repulsion integrals between two finite domains. Green dots represent integrals with one small parameter that can be computed by analytic expressions. Blue dots represent integrals that are evaluated from power series expansions. Grey circles are the integrals to be evaluated by recursion. The red and yellow circles denote the integrals that can be used to form the integrals shown by the red and yellow triangles respectively.

This recurrence relation can be applied to the integrals constructed from Eqs. (3.26), (3.27) and (3.28). As an illustrative example one possible recurrence relation is

$$
\frac{(\mathcal{E}_{\mu-1}^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_{\mu-1}^p | \mathcal{E}_\nu^p)(\mathcal{E}_\lambda^q | \mathcal{E}_\sigma^q)} = \frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{E}_\lambda^q | \mathcal{E}_\sigma^q)} + \frac{a^2(n - \frac{1}{2})(n + \frac{1}{2})}{b^2(m + \frac{1}{2})(m + \frac{3}{2})}
$$
$$
\left(\frac{(\mathcal{E}_{\mu+1}^p \mathcal{E}_\nu^p | \mathcal{E}_{\lambda-2}^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_{\mu+1}^p | \mathcal{E}_\nu^p)(\mathcal{E}_{\lambda-2}^q | \mathcal{E}_\sigma^q)} - \frac{(\mathcal{E}_{\mu+1}^p \mathcal{E}_\nu^p | \mathcal{E}_{\lambda-1}^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_{\mu+1}^p | \mathcal{E}_\nu^p)(\mathcal{E}_{\lambda-1}^q | \mathcal{E}_\sigma^q)} \right) \quad (3.31)
$$

Numerical experiments suggest that this backwards recurrence in μ is sufficiently stable for our purposes but that the forward recurrence is unstable. This necessitates the construction of starting values for the recurrence. Two sets of starting values are required, and these sets can be viewed as two rows where $m = \mu + \nu$ is small and two columns where $n = \lambda + \sigma$ is large. Fig. 3.1 shows a graphical representation of the recurrence.

Obtaining the two columns where n is large can be achieved adequately by using the power series definition of the $_0F_1$ hypergeometric function

$$_0F_1(a;z) = \sum_{k=1}^{\infty} \frac{1}{(a)_k} \frac{z^k}{k!} \tag{3.32}$$

$$(a)_k = a(a+1)(a+2)\dots(a+k-1)(a+k) \tag{3.33}$$

where $(a)_k$ is the Pochhammer symbol [100]. The terms in the sum decay rapidly when the hypergeometric parameter is of modest size and, as a result, the sum can be truncated after a few terms.

To construct the two rows with small values of $m = \mu + \nu$ we have obtained analytic expressions for the necessary integrals when $m = 0$ and $m = -1$. Let $m = \mu + \nu$, $n = \lambda + \sigma$, $a = (A_{p+1} - A_p)/2$, $b = (A_{q+1} - A_q)/2$ and $R = |(A_q + b) - (A_p + a)|$ be the distance between the centroids of the two domains p and q. Then for two pairs of even polynomials these expressions are

$$\frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{E}_\lambda^q | \mathcal{E}_\sigma^q)} = \frac{1}{2} \left[\frac{1}{R-a} {}_2F_1 \left(\frac{1}{2}, 1; n + \frac{3}{2}; \frac{b^2}{(R-a)^2} \right) \right.$$
$$\left. + \frac{1}{R+a} {}_2F_1 \left(\frac{1}{2}, 1; n + \frac{3}{2}; \frac{b^2}{(R+a)^2} \right) \right] \tag{3.34}$$

when $m = -1$ and

$$\frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{E}_\lambda^q | \mathcal{E}_\sigma^q)} = \frac{1}{8a} \left[\frac{2b^2}{(3+2n)(R-a)^2} {}_3F_2 \left(1, 1, \frac{3}{2}; 2, n + \frac{5}{2}; \frac{b^2}{(R-a)^2} \right) \right.$$
$$- \frac{2b^2}{(3+2n)(R+a)^2} {}_3F_2 \left(1, 1, \frac{3}{2}; 2, n + \frac{5}{2}; \frac{b^2}{(R+a)^2} \right)$$
$$\left. - 4 \left(\ln(R-a) - \ln(R+a) \right) \right] \tag{3.35}$$

when $m = 0$.

For a pair of even functions interacting with a pairing of an even and an odd polynomial the expressions are

$$
\frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{O}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{O}_\lambda^q | \mathcal{O}_\sigma^q)} = -\frac{b}{2\sqrt{3+4\lambda}} \left[\frac{(3+2n)[(3+2n)(R-a)^2 - b^2]}{((R-a)^2 - b^2)^2} \right.
$$
$$
- \frac{4(1+n)(2+n)(R-a)^2}{((R-a)^2 - b^2)^2} {}_2F_1\left(-\frac{1}{2}, 1; n+\frac{5}{2}; \frac{b^2}{(R-a)^2}\right)
$$
$$
+ \frac{(3+2n)[(3+2n)(R+a)^2 - b^2]}{((R+a)^2 - b^2)^2}
$$
$$
\left. - \frac{4(1+n)(2+n)(R+a)^2}{((R+a)^2 - b^2)^2} {}_2F_1\left(-\frac{1}{2}, 1; n+\frac{5}{2}; \frac{b^2}{(R+a)^2}\right) \right] \quad (3.36)
$$

when $m = -1$ and

$$
\frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{O}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{O}_\lambda^q | \mathcal{O}_\sigma^q)} = -\frac{b}{2a\sqrt{3+4\lambda}} \left[\frac{1}{R-a} {}_2F_1\left(\frac{1}{2}, 1; n+\frac{5}{2}; \frac{b^2}{(R-a)^2}\right) \right.
$$
$$
\left. + \frac{1}{R+a} {}_2F_1\left(\frac{1}{2}, 1; n+\frac{5}{2}; \frac{b^2}{(R+a)^2}\right) \right] \quad (3.37)
$$

when $m = 0$.

Note that when the two domains being integrated over are adjacent it is necessary to take the limit as $R - a \to 0$.

It is not necessary to consider any more integrals. Using Eq. (3.25) both bras and kets containing pairs of odd basis functions can be constructed from those over pairs of even functions. Additionally, integrals over two mixed pairs (one even and one odd function) can be formed from the integral over two pairs of even functions

$$
\frac{(\mathcal{E}_\mu^p \mathcal{O}_\nu^p | \mathcal{E}_\lambda^q \mathcal{O}_\sigma^q)}{(\mathcal{O}_\mu^p | \mathcal{O}_\nu^p)(\mathcal{O}_\lambda^q | \mathcal{O}_\sigma^q)} = -\frac{4b}{a} \frac{(m+1/2)(m+3/2)}{\sqrt{3+4\mu}\sqrt{3+4\lambda}}
$$
$$
\left(\frac{(\mathcal{E}_{\mu-1}^p \mathcal{E}_\nu^p | \mathcal{E}_{\lambda+1}^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_{\mu-1}^p | \mathcal{E}_\nu^p)(\mathcal{E}_{\lambda+1}^q | \mathcal{E}_\sigma^q)} - \frac{(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_{\lambda+1}^q \mathcal{E}_\sigma^q)}{(\mathcal{E}_\mu^p | \mathcal{E}_\nu^p)(\mathcal{E}_{\lambda+1}^q | \mathcal{E}_\sigma^q)} \right) \quad (3.38)
$$

Any other required integrals can be obtained via symmetry.

3.2.4   One-domain two-electron integrals

We now turn our attention to integrals where all four basis functions reside in the same domain. Here, one finds that the Coulomb and exchange components of the antisymmetrised integral diverge but their sum is finite. To treat these, we rewrite the Coulomb operator as the limit of a sequence of softened potentials $\mathrm{erf}\left(\omega(x-y)\right)/(x-y)$, where ω determines the degree of softening.

Taking the Fourier transform of this operator allows us to also bring the two basis function pairs into frequency space. Let $(\mathbf{F}^p_\mu \mathbf{F}^p_\nu | \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma)_\omega$ denote a two electron integral using the softened Coulomb operator, $\mathcal{F}_{\mu\nu} = \mathcal{F}[\mathbf{F}^p_\mu \mathbf{F}^p_\nu]$ be the Fourier transform of the product $\mathbf{F}^p_\mu \mathbf{F}^p_\nu$, $\Gamma(a, x)$ denote the incomplete gamma function and γ denote Euler's constant. Then we find

$$
\begin{aligned}
(\mathbf{F}^p_\mu \mathbf{F}^p_\nu | \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma)_\omega &= \left(\mathbf{F}^p_\mu \mathbf{F}^p_\nu \left| \frac{\mathrm{erf}\,(\omega(x-y))}{x-y} \right| \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma \right) \\
&= \left(\mathbf{F}^p_\mu \mathbf{F}^p_\nu \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma\left(0, \frac{k^2}{4\omega^2}\right) e^{ik(x-y)} dk \right| \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma \right) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma\left(0, \frac{k^2}{4\omega^2}\right) \left(\int_d \mathbf{F}^p_\mu \mathbf{F}^p_\nu e^{ikx} dx \right) \left(\int_d \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma e^{-iky} dy \right) dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma\left(0, \frac{k^2}{4\omega^2}\right) \mathcal{F}_{\mu\nu}(k) \mathcal{F}_{\lambda\sigma}(-k) dk \\
&= \frac{1}{2\pi} \left[\int_{-\infty}^{\infty} \left(2\sinh^{-1}(\sqrt{2}\omega) - \gamma - 3\ln 2 \right) \mathcal{F}_{\mu\nu}(k) \mathcal{F}_{\lambda\sigma}(-k)\, dk \right. \\
&\quad \left. + \int_{-\infty}^{\infty} \left(\Gamma\left(0, \frac{k^2}{4\omega^2}\right) - 2\sinh^{-1}(\sqrt{2}\omega) + \gamma + 3\ln 2 \right) \mathcal{F}_{\mu\nu}(k) \mathcal{F}_{\lambda\sigma}(-k)\, dk \right]
\end{aligned}
$$

$$(3.39)$$

$$
\begin{aligned}
(\mathbf{F}^p_\mu \mathbf{F}^p_\nu \| \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma) &= \lim_{\omega \to \infty} (\mathbf{F}^p_\mu \mathbf{F}^p_\nu | \mathbf{F}^p_\lambda \mathbf{F}^p_\sigma)_\omega - (\mathbf{F}^p_\mu \mathbf{F}^p_\lambda | \mathbf{F}^p_\nu \mathbf{F}^p_\sigma)_\omega \\
&= \lim_{\omega \to \infty} \left[\frac{1}{2\pi} \left(2\sinh^{-1}(\sqrt{2}\omega) - \gamma - 3\ln 2 \right) \right. \\
&\qquad \left(\int_{-\infty}^{\infty} \mathcal{F}_{\mu\nu}(k) \mathcal{F}_{\lambda\sigma}(-k) dk - \int_{-\infty}^{\infty} \mathcal{F}_{\mu\lambda}(k) \mathcal{F}_{\nu\sigma}(-k) dk \right) \\
&\qquad + \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\Gamma\left(0, \frac{k^2}{4\omega^2}\right) - 2\sinh^{-1}(\sqrt{2}\omega) + \gamma + 3\ln 2 \right) \\
&\qquad \left. (\mathcal{F}_{\mu\nu}(k) \mathcal{F}_{\lambda\sigma}(-k) - \mathcal{F}_{\mu\lambda}(k) \mathcal{F}_{\nu\sigma}(-k)) dk \right] \quad (3.40)
\end{aligned}
$$

We can use the following relationship, which is a consequence of Parseval's theorem [100], to remove the first term within the limit

$$\int_{-\infty}^{\infty} \mathcal{F}_{\mu\nu}(k)\mathcal{F}_{\lambda\sigma}(-k)dk = \int_{d} \mathbf{F}_{\mu}^{p}(r)\mathbf{F}_{\nu}^{p}(r)\mathbf{F}_{\lambda}^{p}(r)\mathbf{F}_{\sigma}^{p}(r)dr = \int_{-\infty}^{\infty} \mathcal{F}_{\mu\lambda}(k)\mathcal{F}_{\nu\sigma}(-k)dk \tag{3.41}$$

and obtain

$$(\mathbf{F}_{\mu}^{p}\mathbf{F}_{\nu}^{p}||\mathbf{F}_{\lambda}^{p}\mathbf{F}_{\sigma}^{p}) = \lim_{\omega \to \infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\Gamma\left(0, \frac{k^2}{4\omega^2}\right) - 2\sinh^{-1}(\sqrt{2}\omega) + \gamma + 3\ln 2 \right) \right.$$
$$\left. (\mathcal{F}_{\mu\nu}(k)\mathcal{F}_{\lambda\sigma}(-k) - \mathcal{F}_{\mu\lambda}(k)\mathcal{F}_{\nu\sigma}(-k))dk \right]$$
$$= \{\mathbf{F}_{\mu}^{p}\mathbf{F}_{\nu}^{p}|\mathbf{F}_{\lambda}^{p}\mathbf{F}_{\sigma}^{p}\} - \{\mathbf{F}_{\mu}^{p}\mathbf{F}_{\lambda}^{p}|\mathbf{F}_{\nu}^{p}\mathbf{F}_{\sigma}^{p}\} \tag{3.42}$$

where we have introduced the "quasi-integral"

$$\{\mathbf{F}_{\mu}^{p}\mathbf{F}_{\nu}^{p}|\mathbf{F}_{\lambda}^{p}\mathbf{F}_{\sigma}^{p}\} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}_{\mu\nu}(k)\mathcal{F}_{\lambda\sigma}(-k)\ln\left(\frac{k^2}{4}\right)dk \tag{3.43}$$

There are two cases: (1) the domain is infinite; (2) the domain is finite.

Infinite/Infinite quasi-integrals

It is easy to show that

$$\frac{\{\mathcal{L}_{\mu}\mathcal{L}_{\nu}|\mathcal{L}_{\lambda}\mathcal{L}_{\sigma}\}}{(\mathcal{L}_{\mu}|\mathcal{L}_{\nu})(\mathcal{L}_{\lambda}|\mathcal{L}_{\sigma})} = \frac{\{\mathcal{R}_{\mu}\mathcal{R}_{\nu}|\mathcal{R}_{\lambda}\mathcal{R}_{\sigma}\}}{(\mathcal{R}_{\mu}|\mathcal{R}_{\nu})(\mathcal{R}_{\lambda}|\mathcal{R}_{\sigma})} = \alpha\zeta\eta\left(\frac{\zeta^2 + 6\zeta\eta + \eta^2}{2(\zeta + \eta)^3} - \frac{6\zeta^2\eta^2 \ln \zeta\eta}{(\zeta + \eta)^5}\right) \tag{3.44}$$

Finite/Finite quasi-integrals

Because of symmetry, only the $\{\mathcal{E}_{\mu}^{p}\mathcal{E}_{\nu}^{p}|\mathcal{E}_{\lambda}^{p}\mathcal{E}_{\sigma}^{p}\}$, $\{\mathcal{E}_{\mu}^{p}\mathcal{E}_{\mu}^{p}|\mathcal{O}_{\lambda}^{p}\mathcal{O}_{\sigma}^{p}\}$, $\{\mathcal{O}_{\mu}^{p}\mathcal{O}_{\nu}^{p}|\mathcal{O}_{\lambda}^{p}\mathcal{O}_{\sigma}^{p}\}$ and $\{\mathcal{E}_{\mu}^{p}\mathcal{O}_{\nu}^{p}|\mathcal{E}_{\lambda}^{p}\mathcal{O}_{\sigma}^{p}\}$ integral classes are non-vanishing and, if the relationship (3.25) is exploited, it is necessary only to compute the first and last of these classes.

We have been unable to find a general expression for these quasi-integrals. However, because they scale inversely with the width of the domain, we have tabulated the necessary quasi-integrals for a domain on the interval $[-1, 1]$ and our program simply scales these, as required, on the fly.

3.3   Implementation

We have implemented the above theory in a program that we call CHEM1D. So that others may understand how the results of Sec. 3.2 are used practically we present an overview of the structure of CHEM1D. In the following section, we define $\mathcal{D}$ to be the set of all domains (both finite and infinite) in the molecule, $\mathcal{D}_I$ and $\mathcal{D}_F$ to be the set of all infinite and finite domains respectively, $N(p)$ to be the total number of basis functions in the domain p and $E(p)$ and $O(p)$ to be the number of even and odd basis functions, respectively, in the finite domain p.

3.3.1   Integral evaluation

The heart of any quantum chemistry program is its integral evaluation routines, particularly those for the two-electron integrals, for it is these quantities that typically represent the computational bottleneck and make the heaviest demand on computer memory. Our algorithms are not optimal but we have attempted to exploit some of the simplifications that arise as a consequence of the domain separation of electrons.

One-electron integrals

Because the product of two basis functions in different domains vanishes the one-electron integral matrices are block diagonal with each block corresponding to a domain. Therefore for each of the overlap, kinetic and nuclear attraction matrices, and for each domain $p \in \mathcal{D}$, we create a square matrix of size $N(p)$.

Because the integrals can be computed from analytic expressions, computing the content of these matrices is a relatively simple process. We use three subroutines: one for the overlap, one for the kinetic energy and one for the nuclear attraction energy. Each routine operates over one domain with each call.

Let S^p, T^p and V^p be the overlap, kinetic and nuclear attraction matrix for the domain p respectively and Z_n be the charge of the nth nucleus. Pseudocode describing our methods for computing these integrals can be seen in Algorithm 1

Two-domain two-electron integrals

We now turn our attention to the significantly more complicated two-electron integrals. As described in Sec. 3.2.3 these integrals can be separated into two classes. We begin with the case of non-overlapping integrals, which can be further separated into another three subclasses: integrals between two infinite domains, integrals between a finite and an infinite domain and integrals between two finite domains.

Algorithm 1 One-electron integrals

1: **procedure** OVERLAP(p)   ▷ Computes S^p
2:     **if** p is an infinite domain **then**
3:         **for** $i, j = 1 \rightarrow N(p)$ **do**
4:             $S^p_{i,j}$ = Eq. (3.6)
5:         **end for**
6:     **else**
7:         **for** $i, j = 1 \rightarrow E(p)$ **do**
8:             $S^p_{i,j}$ = Eq. (3.7)   ▷ Compute Γ ratios by recursion
9:         **end for**
10:         **for** $i, j = 1 \rightarrow O(p)$ **do**
11:             $S^p_{i+E(p),j+E(p)}$ = Eq. (3.8)
12:         **end for**
13:     **end if**
14: **end procedure**
15: **procedure** KINETIC(p)   ▷ Computes T^p
16:     **if** p is an infinite domain **then**
17:         **for** $i, j = 1 \rightarrow N(p)$ **do**
18:             $T^p_{i,j} = S^p_{i,j} \times$ Eq. (3.9)
19:         **end for**
20:     **else**
21:         **for** $i, j = 1 \rightarrow E(p)$ **do**
22:             $T^p_{i,j} = S^p_{i,j} \times$ Eq. (3.10)
23:         **end for**
24:         **for** $i, j = 1 \rightarrow O(p)$ **do**
25:             $T^p_{i+E(p),j+E(p)} = S^p_{i+E(p),j+E(p)} \times$ Eq. (3.11)
26:         **end for**
27:     **end if**
28: **end procedure**
29: **procedure** POTENTIAL(p)   ▷ Computes V^p
30:     **if** p is an infinite domain **then**
31:         **for** $i, j = 1 \rightarrow N(p)$ **do**
32:             $V^p_{i,j}$ = Eq. (3.12) or Eq. (3.13)
33:         **end for**
34:     **else**
35:         **for** $i, j = 1 \rightarrow E(p)$ **do**
36:             $V^p_{i,j}$ = Eq. (3.14)
37:         **end for**
38:         **for** $i = 1 \rightarrow E(p), j = 1 \rightarrow O(p)$ **do**
39:             $V^p_{i,j+E(p)}$ = Eq. (3.16)
40:             $V^p_{j+E(p),i} = V^p_{i,j+E(p)}$
41:         **end for**
42:         **for** $i, j = 1 \rightarrow O(p)$ **do**
43:             $V^p_{i+E(p),j+E(p)}$ = Eq. (3.15)
44:         **end for**
45:     **end if**
46: **end procedure**

Algorithm 2 Two-electron integrals between infinite domains

1: **procedure** INF-INF(p, q)
2:     **if** the system is an atom **then**
3:         **for** $\mu, \nu = 1, N(p)$ **do**
4:             $\zeta = \mu^2 + \nu^2$
5:             **for** $\lambda, \sigma = 1, N(q)$ **do**
6:                 $\eta = \lambda^2 + \sigma^2$
7:                 **if** $\zeta = \eta$ **then**
8:                     $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.21)
9:                 **else**
10:                    $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.20)
11:                **end if**
12:            **end for**
13:        **end for**
14:    **else**
15:        **for** $\mu, \nu = 1, \max(N(p), N(p))$ **do**
16:            Precompute $U\left(1, 1, R\alpha(\mu^2 + \nu^2)\right)$
17:        **end for**
18:        **for** $\mu, \nu = 1, N(p)$ **do**
19:            $\zeta = \mu^2 + \nu^2$
20:            **for** $\lambda, \sigma = 1, N(q)$ **do**
21:                $\eta = \lambda^2 + \sigma^2$
22:                **if** $\zeta = \eta$ **then**
23:                    $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.19)
24:                **else**
25:                    $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.18)
26:                **end if**
27:            **end for**
28:        **end for**
29:    **end if**
30: **end procedure**

For convenience, we store the full set of ERIs and this is feasible for all of the molecules reported here. The integrals are stored in a four-index array for each pair of domains (p, q) in the molecule. The first two indices of each array are of length $N(p)$, while the last two indices are of length $N(q)$.

When both domains are infinite we use the analytic expressions in Eqs. (3.18)–(3.21). When the system is not an atom the evaluation of a confluent hypergeometric function is required. We evaluate these functions before entering the loop structure for the integrals; the hypergeometric function is required twice for most evaluations, one call dependent on the two basis functions from p and one on the two basis functions from q. An overview of the procedure we use here can be found in Algorithm 2.

Algorithm 3 Two-electron integrals between infinite and finite domains

---

1: **procedure** INF_FIN(p, q)
2:     **for** $\mu, \nu = 1, N(p)$ **do**
3:         $\zeta = \mu^2 + \nu^2$
4:         **for** $i = 1, 500$ **do**
5:             Compute $U(3, 3, \zeta x_i)$
6:         **end for**
7:         **for** $\lambda, \sigma = 1, E(q)$ **do**
8:             $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.22)   ▷ Computed using trapezoidal rule
9:         **end for**
10:         **for** $\lambda = 1, E(q), \sigma = 1, O(q)$ **do**
11:             $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.23)   ▷ Computed using trapezoidal rule
12:             $(\mu\nu|\sigma\lambda) = (\mu\nu|\lambda\sigma)$
13:         **end for**
14:         **for** $\lambda, \sigma = 1, O(q)$ **do**
15:             $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times$ Eq. (3.24)   ▷ Computed using trapezoidal rule
16:         **end for**
17:     **end for**
18: **end procedure**

When p is an infinite domain and q is a finite domain (or vice versa), we use a 500-point trapezoidal rule to achieve the quadrature described above. This is simple to implement and affords sufficient accuracy and, at least as the number of domains increase, is not the bottleneck of the program's evaluation. The most expensive step in the quadrature is the evaluation of confluent hypergeometric functions. These can be evaluated within the loop structure for selecting a basis set pair in the infinite domain, but before the loop structure for the finite domain. They can then be kept in temporary storage and used for all three types of finite domain basis function pairs. Pseudocode for the evaluation of this class can be found in Algorithm 3; we assume p to be the infinite domain.

The final classes of two-domain integrals are those involving finite domains. This step requires the most complex code in CHEM1D due to the need to calculate starting values for the various recurrences from a variety of methods, execute the recurrences and correctly pack the resulting prototype integrals into the appropriate parts of the storage arrays. To handle these steps we use three levels of routines. The bottom level computes the necessary starting values for the recurrence relations used. The middle level computes the prototype integrals using the recurrence relations described in Sec. 3.2.3. The top level then transfers the completed prototype integrals from temporary storage into the final storage arrays.

The middle subroutines are straightforward implementations of our four-term recurrence relation. During the top level procedure, we make extensive use of the

symmetry allowed by the use of real-valued basis functions, which minimises the looping structures required to traverse the many possible arrangements of integrals. This requires a great deal of care to ensure that the integrals are scattered into the correct storage spaces, but is ultimately a straightforward process.

The most complicated routines for computing recurrence starting values are those for when $\mu + \nu = -1, 0$. The analytic expressions that must be evaluated contain two different hypergeometric functions. To evaluate these we employ backward recurrence, using the power series expansion to evaluate the starting points. For illustrative purposes we give an outline of the code for the computation of integrals of the types $(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{E}_\lambda^q \mathcal{E}_\sigma^q)$ and $(\mathcal{E}_\mu^p \mathcal{E}_\nu^p | \mathcal{O}_\lambda^q \mathcal{O}_\sigma^q)$ in Algorithm 4.

Quasi-integrals

Constructing the quasi-integrals for the single-domain case is significantly simpler than building the integrals for the two-domain cases. For the infinite domains we use the analytic expression in Eq. (3.44), which requires no special functions. As stated above we scale a set of standard integrals for the finite domain. We have used Mathematica [99] to compute the analytic solution to the integrals for a domain on the line segment $[-1, 1]$ and evaluate them to high precision. These values are appropriately scaled by CHEM1D as needed.

We use two subroutines to perform these tasks: QUASI_INF(p) gives the integrals for the infinite domain p and QUASI_FIN(q) those for the finite domain q.

Antisymmetrised integrals

The above routines are all called by a driver routine which handles the looping over the domain structures. In our implementation this routine is also responsible for managing the temporary storage space required and locating the appropriate areas of memory for storing the single-bar and quasi-integrals for each pairing of p and q. The contraction of quasi-integrals to form correct double bar integrals in the case of overlapping domains is handled by a separate subroutine. This operation is almost identical to that required for normal three-dimensional quantum chemical calculations; as such we have excluded its pseudocode. However the driver routine is described in Algorithm 5.

Algorithm 4 Two-electron integrals between finite domains

1: **procedure** FIN_FIN(p, q)
2:     MaxM $= 2 \max(E(p), E(q), O(p), O(q)) + 1$
3:     Allocate proto$(-1 : \text{MaxM}, -1 : \text{MaxM})$ ▷ Used to store prototype integrals
4:     EEEE_PROTOTYPES$(\text{MaxM}, \text{proto})$
5:     **for** $\mu, \nu = 1, E(p)$ **do**
6:         **for** $\lambda, \sigma = 1, E(q)$ **do**
7:             $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times \text{proto}(\mu + \nu, \lambda + \sigma)$
8:             $(\lambda\sigma|\mu\nu) = (\mu\nu|\lambda\sigma)$
9:         **end for**
10:        **for** $\lambda, \sigma = 1, O(q)$ **do**
11:            $(\mu\nu|\lambda\sigma) = S^p_{\mu\nu} S^q_{\lambda\sigma} \times [(3 + 2(\lambda + \sigma))\text{proto}(\mu + \nu, \lambda + \sigma) - (2 + 2(\lambda + \sigma))\text{proto}(\mu + \nu, \lambda + \sigma + 1)]$ ▷ Using Eq. (3.25)
12:            $(\lambda\sigma|\mu\nu) = (\mu\nu|\lambda\sigma)$
13:        **end for**
14:    **end for**
15:    $\vdots$
16: **end procedure**
17: **procedure** EEEE_PROTOTYPES(MaxM, proto)
18:     M_MINUS_ONE$(\text{proto}(:, -1))$
19:     M_ZERO$(\text{proto}(:, 0))$
20:     **for** $i = 1, \text{MaxM}$ **do**
21:         $\text{proto}(\text{MaxM}, i) = (3.34)$
22:         $\text{proto}(\text{MaxM} - 1, i) = (3.35)$ ▷ Compute from power series
23:     **end for**
24:     **for** $n = 1, \text{MaxM}$ **do**
25:         **for** $m = \text{MaxM} - 2, 1, -1$ **do**
26:             Generate proto(m, n) using Eq. (3.31)   ▷ $m = \mu + \nu, n = \lambda + \sigma$
27:         **end for**
28:     **end for**
29: **end procedure**
30: **procedure** M_MINUS_ONE$(\textit{output})$
31:     Compute hypergeometrics in (3.34) for $\mu + \nu = \text{MaxM}, \text{MaxM} - 1$
32:   ▷ From power series
33:     **for** $m = \text{MaxM} - 2, 1, -1$ **do**
34:         Compute hypergeometrics in (3.34) for $\mu + \nu = m$ via recurrence
35:     **end for**
36:     Combine hypergeometric functions and store in *output*   ▷ Using Eq. (3.34)
37: **end procedure**
38: **procedure** M_ZERO(output)
39:     Compute hypergeometrics in (3.35) for $\mu + \nu = \text{MaxM}, \text{MaxM} - 1$
40:   ▷ From power series
41:     **for** $m = \text{MaxM} - 2, 1, -1$ **do**
42:         Compute hypergeometrics in (3.35) for $\mu + \nu = m$ via recurrence
43:     **end for**
44:     Combine hypergeometric functions and store in *output*   ▷ Using Eq. (3.35)
45: **end procedure**

Algorithm 5 Electron repulsion integral driver

  1: **procedure** ERI
  2:     INF_INF
  3:     **for** $p \in \mathcal{D}_I$ **do**
  4:         QUASI_INF(p)
  5:         **for** $q \in \mathcal{D}_F$ **do**
  6:             INF_FIN(p, q)
  7:         **end for**
  8:     **end for**
  9:     **for** $p \in \mathcal{D}_F$ **do**
 10:         QUASI_FIN(p)
 11:         **for** $q \in \mathcal{D}_F$ **do**
 12:             FIN_FIN(p, q)
 13:         **end for**
 14:     **end for**
 15: **end procedure**

3.3.2   Self-consistent field calculations

Once the necessary integrals have been evaluated CHEM1D uses them to compute the Hartree-Fock self-consistent field (SCF) energy followed by the second- and third-order Møller-Plesset perturbation energies. The formulations used are those described in the text by Szabo & Ostlund [8]; however some modifications to the Hartree-Fock methods are made, mostly to exploit the domain separation of the electrons.

Like most basis sets used in quantum chemistry the basis sets described in Sec. 3.2.1 are not orthogonal. In order to perform the iterative SCF procedure for computing the Hartree-Fock energy it is necessary to transform the non-orthogonal basis to an orthogonal one. CHEM1D follows the procedure described by Szabo & Ostlund [8] for canonical orthogonalisation. We prefer this over symmetric orthogonalisation for numerical reasons. Performing either of these orthogonalisations requires finding the eigendecomposition for the overlap matrix. This matrix becomes poorly conditioned relatively quickly due to the basis sets we use. The canonical orthogonalisation allows us to account for this by removing columns of the matrix; the symmetric orthogonalisation does not, resulting in a large loss of precision and erratic SCF behaviour. Note that since this matrix is dependent upon the overlap matrix it inherits the same block structure.

Three quantities remain to be considered before an SCF iteration can be described. The first two of these are the density and orbital coefficient matrices. An initial guess at the orbitals is needed; CHEM1D diagonalises the core Hamiltonian matrix to

---

Algorithm 6 Construction of the two electron G^p matrix

1: **procedure** BUILD_G(p)
2:      **for** $q \in \mathcal{D}$ **do**   ▷ Find the interaction of domain p with every other domain
3:          **for** $\mu, \nu = 1, N(p)$ **do**
4:              $G_{\mu,\nu}^p = \sum_{\lambda,\sigma=1}^{N(q)} P_{\lambda\sigma}^q (\mu\nu||\lambda\sigma)$ ▷ P^q is the density in domain q
5:          **end for**
6:      **end for**
7: **end procedure**

---

generate a guess which we have found to be sufficient. It is then possible to generate the density matrix from the orbital coefficients as described by Szabo & Ostlund [8].

Each orbital must be constructed of basis functions in only one domain; constructing an orbital from basis functions in multiple domains implies that an electron occupying it would be able to move across nuclei, violating impenetrability. Both the density matrix and the orbital coefficient matrix therefore have the same block structure as the one electron matrices. We can store both of these and construct the density matrix domain by domain.

The final matrix that is needed to perform an SCF cycle is referred to as the G matrix by Szabo & Ostlund. This matrix measures the average repulsion of an electron in the field of all others and is given by the following formula

$$G_{\mu,\nu} = \sum_{\lambda\sigma} P_{\lambda,\sigma} (\mu\nu||\lambda\sigma) \tag{3.45}$$

where P is the density matrix. The reliance on P forces λ and σ to be in the same domain; the sum can then be separated further into another sum over each domain. For the integral $(\mu\nu||\lambda\sigma)$ to be non-zero μ and ν must then also be in the same domain, giving the matrix G the same block structure as all the previous matrices we have considered. We give our algorithm for the construction of G over a given domain p in Algorithm 6.

We have now described the necessary components for an SCF iteration to be evaluated. However, thanks to the domain separation present in all the matrices, a significant optimisation can be included: each iteration can be computed over each domain individually. The only exception to this complete separation of the SCF procedure is that the new density in each cycle must be calculated for each domain before the cycle can begin; this ensures the correct coupling between electrons in different domains. We note that this leads to a simple method for parallelisation of the calculation, with the limit that each thread will require race conditions to ensure the iteration over each domain proceeds in lockstep.

Algorithm 7 SCF iteration

1: **procedure** SCF_CYCLE
2:     **for** $p \in \mathcal{D}$ **do**
3:         BUILD_P(p) ▷ Update density matrix for all domains
4:     **end for**
5:     **for** $p \in \mathcal{D}$ **do**
6:         BUILD_G(p)
7:         $F_1^p = G^p + T^p + V^p$ ▷ F_1^p is the current Fock matrix
8:         Build and store new error matrix   ▷ F_i^p is from ith last cycle
9:         **if** Current cycle $< l$ **then**   ▷ Number of stored Fock matrices $= l$
10:             $F' = (X^p)^T F^p X^p$
11:         **else**
12:             Solve DIIS matrix   ▷ DIIS coefficients $\rightarrow c_i$
13:             $F' = (X^p)^T (\sum_{i=1}^{l} c_i F_i^p) X^p$
14:         **end if**
15:         Diagonalise F' ▷ Eigenvectors $\rightarrow C'$
16:         $C^p = X^p C'$
17:     **end for**
18:     **if** not all domains have converged **then**
19:         SCF_CYCLE   ▷ Recursive call
20:     **end if**
21: **end procedure**

We also employ the DIIS method of Pulay [12, 13] to accelerate the SCF convergence. It reduces both the build-up of roundoff error and the execution time. By default the SCF iteration terminates when the root-mean-square of the DIIS error matrix drops below 10^{-6}. We present pseudocode for the procedure we use to evaluate an SCF cycle in Algorithm 7.

3.3.3   Møller-Plesset perturbation theory

The final components of CHEM1D are the routines for calculating Møller-Plesset perturbation corrections to the Hartree-Fock energy. Both the second- and third-order corrections are currently implemented according to the procedure described by Szabo and Ostlund [8]. Our current implementation does not exploit the domain separation and because of this the calculation of MPn corrections is one of the bottlenecks of the program. Due to the standard nature of our algorithm for these calculations we have omitted a pseudocode description here.

3.3.4   Pseudocode overview

Now that we have defined how the various components operate we present an overview of the main CHEM1D program in Algorithm 8.

Algorithm 8 Chem1D overview

 1: **procedure** CHEM1D(*input*)
 2:     Read *input*
 3:     Allocate integral storage
 4:     **for** $p \in \mathcal{D}$ **do**   ▷ Build the one electron matrices
 5:         OVERLAP(*p*)
 6:         KINETIC(*p*)
 7:         POTENTIAL(*p*)
 8:     **end for**
 9:     ERI   ▷ Build the set of two electron integrals
10:     DOUBLE_BAR   ▷ Contract the quasi-integrals
11:     **for** $p \in \mathcal{D}$ **do**   ▷ Build the orthogonalisation matrices
12:         BUILD_X(*p*)
13:     **end for**
14:     SCF_CYCLE
15:     MO_TRANSFORM   ▷ Transform the double bar integrals to MO basis
16:     MPN_CORRELATION
17:     Print output
18: **end procedure**

3.4   Results

3.4.1   Atomic energies

We begin our tests of CHEM1D by studying atomic systems. In Chapter 2, we presented accurate total energies, ionisation energies and electron affinities of the first ten elements. As a preliminary test of the capabilities of CHEM1D we have performed the same study here.

Our experiments suggest that a basis set of nine exponential functions on each side of the nucleus (see Eqs. (3.1) and (3.2)) produces the best results within the confines of CHEM1D's limitations (which will be discussed shortly). CHEM1D's default value of $\alpha = Z/m^2$ has been used, where $m = 9$ and Z is the charge of the atom's nucleus. We remove one orbital during the canonical orthogonalisation process to facilitate convergence and the results are shown in Table 3.1.

For small atoms, the total energies produced by CHEM1D compare favourably with our previous results. Up to boron, the Hartree-Fock energies agree to within a millihartree. However, beyond this point, the results diverge significantly. These inaccuracies stem directly from the lack of diffuse functions in the basis set that we have used in the present study, because, as observed in Chapter 2, heavy atoms in 1D are surprisingly diffuse objects. Unsurprisingly, the Møller-Plesset correlation energies also suffer from basis set incompleteness because, as the number of electrons increases, the virtual space is progressively diminished. The total energy errors are

Atom	Atomic Energies			Ionisation Energies			Electron Affinities[a]		
	$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	HF	MP2	MP3	HF	MP2	MP3
H	0.500000	0.000	0.000	13.606	13.606	13.606	3.887	3.933	3.956
He	3.242552	1.998	2.636	33.812	33.866	33.883			
Li	8.007278	2.953	3.644	4.485	4.507	4.512	1.396	1.409	1.414
Be	15.415561	5.128	6.127	10.351	10.388	10.397			
B	25.356662	6.818	8.036	2.060	2.077	2.084	0.584	0.594	0.598
C	38.082819	9.613	11.420	4.621	4.654	4.666			
N	53.558680	11.175	12.967	1.054	1.056	1.057	0.349	0.350	0.351
O	71.890505	13.173	15.187	2.232	2.241	2.244			
F	93.055546	14.728	16.757	−0.208	−0.222	−0.227	−1.277	−1.294	−1.301
Ne	117.240157	14.951	16.670	1.075	1.069	1.067			

[a]The electron affinities for He, Be, C, O and Ne have been omitted. As described in Chapter 2, the anions of these species are auto-ionising

Table 3.1: Atomic energies, ionisation energies and electron affinities

inherited by the ionisation energies and electron affinities and, as before, the lack of diffuse functions is to blame.

Simply increasing the basis set beyond the somewhat limited size used here should greatly improve the results obtained. Unfortunately this is not an option here. The basis set used within CHEM1D rapidly develops linear dependencies as more diffuse functions are included. This causes problems during the orthogonalisation step of the Roothaan-Hall equations, which is highly dependent upon the condition of the overlap matrix.

This problem does appear commonly in 3D molecular calculations, where the atom centred Gaussian basis sets typically used can exhibit degrees of linear dependence. As such there are techniques for diminishing the effect of this problem, notably the canonical orthogonalisation procedure used in CHEM1D [8]. The level of linear dependence that appears here increases so rapidly that these techniques appear unable to overcome the issue.

In spite of the basis set deficiencies, we are still able to observe some of the unusual properties of 1D atoms. In particular, we are able to reconstruct the thin periodic table shown in Chapter 2 by looking at the periodic behaviour of the ionisation energies. In three-dimensional chemistry, the ionisation energy increases monotonically when moving across a row of the periodic table before dropping when moving to the next row. We find similar behaviour here, however the period of these trends is only two atoms, which suggests that the periodic table of 1D elements has only two groups.

Molecule	R_{eq}	$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	Bond Dissociation Energies		
					HF	MP2	MP3
H_1H_1	2.636	1.184572	0.844	1.154	0.185	0.185	0.186
$_1H_1He_1$	2.025	3.880304	2.300	2.978	0.138	0.138	0.138
$_1H_2Li_1$	5.076	8.679540	4.507	5.747	0.172	0.174	0.174
H_2Li_2	5.212	8.541236	3.565	4.352	0.034	0.035	0.035
$_1H_2Be_2$	3.895	16.074792	6.859	8.291	0.159	0.161	0.161
$_1He_2Li_2$	4.505	11.258911	5.398	6.757	0.009	0.010	0.010
$_1Li_3Li_2$	7.142	16.021014	7.107	8.694	0.006	0.008	0.008

Table 3.2: Equilibrium bond distances (in a.u.), Hartree-Fock energies (in E_h), correlation energies (in mE_h) and bond dissociation energies (in E_h) of diatomic molecules

3.4.2 Diatomic molecules

We now turn our attention to diatomic molecules. In Chapter 2 we investigated diatomics with a maximum of two electrons. We concluded that having two or more electrons in a finite domain is highly destabilising because of the large resulting kinetic energy. We also observed that atomic species with an odd number of electrons carry a permanent dipole due to the particle impenetrability trapping different numbers of electrons on each side of the nucleus. We speculated in Sec. 2.4.2 that, as a consequence of this, stable molecules could form with multiple electrons in a finite domain, despite the increase in kinetic energy, due to favourable dipole-dipole interactions. Thanks to the recent developments in CHEM1D we are now able to test this conjecture.

Results for a small collection of bound states of diatomic molecules can be found in Table 3.2. We use a basis set of 8 exponential functions in each infinite domain and a combination of 6 even and 6 odd functions in the finite domain. This choice is, again, intended to produce the best possible results within the limitations of CHEM1D.

These results show some interesting effects which reveal that our earlier prediction about the nature of bond formation in 1D was not entirely accurate. In fact, there are a number of striking results which are counter-intuitive from the viewpoint of three-dimensional chemistry.

First, we see that H has at least one favourable bonding configuration with other nuclei. The reason for this is simple: H can present an unshielded positive charge to the electrons of the other nucleus. This allows it to bind strongly even to the 1D elements with no permanent dipole. In fact, the bonding in $_1H_2Be_2$ is stronger than that of $_1H_1He_1$. The H nucleus interacts with the outermost electron of the other

atom, which in the case of Be is shielded from the charge of its own nucleus by the interior electron.

There are three configurations for LiH in 1D that seem likely to result in bonding: $_1Li_2H_1$, $_2Li_2H$ and $_2Li_1H_1$. The last of these does not result in binding, probably because the opposed dipoles of the H and Li atoms result in insufficient stabilising potential energy. The other two do form a bound molecule, although there is a very large difference in binding energy. The comparatively small bond dissociation energy of $_2Li_2H$ is a result of the second exterior electron being forced into a higher energy orbital; in $_1Li_2H_1$ both electrons on the outside of the molecule are able to occupy the lowest energy orbital in their domain.

The final noteworthy result relating to the binding in hydrogen-containing diatomics is the remarkable strength of the bond with respect to its length (excepting H_2Li_2). We noted in Chapter 2 that the length of the H_1H_1 bond is roughly double that of the 3D H_2 molecule, while they have a similar strength [122]. The bond in $_1H_2Li_1$, however, is roughly 90% of the strength of the bond in H_1H_1, yet the bond is almost double the length.

Diatomics of two heavy atoms produce more interesting and unexpected results, while also revealing a deficiency in CHEM1D.

We find that the molecule $_1He_2Li_2$ is very gently bound. Previously we would have expected this molecule to be unbound given the lack of stabilising dipole interactions. Given the bonding behaviour of the $_1H_1He_1$ and $_1H_2Be_2$ molecules however, we might have expected this bond to be reasonably strong as a result of similar electrostatic arguments. When the H atom is replaced by Li we introduce an extra electron into the finite bond. Now that the domain contains two electrons, one is forced to occupy a higher energy orbital. This carries a significant energy penalty which the favourable electrostatic interactions are only just able overcome.

We examined two possible configurations of the Li dimer: $_2Li_2Li_2$ and $_1Li_3Li_2$. The former configuration does not bind due to the opposition of the dipoles on the two individual atoms. The latter configuration does bind, but with an extremely long bond length. Using CHEM1D to determine the bonding energy suggests that this molecule has a smaller binding energy than $_1He_2Li_2$. This is unexpected given the dipole-dipole interactions that should stabilise the Li dimer.

Molecule	R_{eq}		$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	Atomisation Energies		
						HF	MP2	MP3
$H_1H_1H_1$	2.648	2.762	1.883315	1.538	2.099	0.383	0.385	0.385
$_1H_1H_1He_1$	2.705	2.033	4.565870	3.068	4.037	0.323	0.324	0.325
$_1H_1He_1H_1$	2.031	2.031	4.493606	2.637	3.417	0.251	0.252	0.252
$_1H_1H_2Be_2$	2.388	4.426	16.841024	6.985	8.480	0.425	0.427	0.428
$_1H_2Be_2H_1$	3.881	3.881	16.730968	8.060	9.862	0.315	0.318	0.319
$_1Li_2H_2Li_2$	4.129	8.458	16.766343	7.902	9.755	0.252	0.254	0.254

Table 3.3: Equilibrium bond distances, energies and atomisation energies of triatomic molecules

We have also examined this using a quadruple-precision version of CHEM1D. Using higher precision allows us to use larger basis sets before near linear dependence becomes an issue. The results from this program suggest that the equilibrium bond length of $_1Li_3Li_2$ is significantly longer (at least 8.5 bohrs). Around this configuration the binding energy increases by at least 40 millihartrees, which suggests the relative bonding behaviour we expect is more likely.

Unfortunately, the basis sets that can typically be used in double precision without incurring numerical problems have only limited ability to describe long bonds with multiple electrons. The basis functions we have used are high in amplitude around the middle of the bond and decay relatively quickly towards its edges. In a stretched bond, however, the electrons tend to localise near the nuclei; in a bond with many electrons and accordingly heavy nuclei this localisation rapidly becomes very restrictive. Describing these orbitals, which are built up near the edges of the bond, with the basis sets we have described requires large molecular orbital coefficients with oscillating signs, a standard recipe for numerical disasters.

In light of this, we have refrained from examining further systems containing 3 or more electrons in a finite domain. Additionally, we consider that a bond length of more than 6 bohrs is likely to be indicative of a significant drop in numerical accuracy.

3.4.3 Triatomic molecules

In addition to diatomics, we have also examined a small selection of triatomic molecules. We have used the same methods as were employed for the diatomic molecules. The results can be seen in Table 3.3

When compared to the behaviour of the diatomic molecules, this set of results reveals a great deal about how larger molecules are likely to form in 1D. The atomisation energies of each triatomic is roughly equal to the sum of the bond

dissociation energies in the appropriate two diatomic molecules. The difference can be rationalised as a consequence of the alignment of the permanent dipoles belonging to the constituent atoms.

Taking $H_1H_1H_1$ as a simple example we obtain an approximate atomisation energy of 0.370 hartrees by doubling the bond dissociation energy of H_1H_1. The dipoles of all three atoms are aligned however, and as result there is a bonus to the binding strength of about 15 millihartrees, giving a final atomisation energy of approximately 0.385.

Moving down the set of examples we find that the atomisation energy of $_1H_1H_1He_1$ is almost exactly equal to the strength of its two constituent bonds. The alignment of the two hydrogen dipoles is already accounted for in the bond dissociation energy of H_1H_1, and the helium atom lacks a dipole which would provide an extra contribution. The strength of the binding in $_1H_1He_1H_1$ (approximately 0.25 hartrees) is slightly lower than that of two $_1H_1He_1$ bonds (approximately 0.275 hartrees) due to the repulsion of the opposing hydrogen dipoles.

This destabilisation is dependent upon the distance between the two interacting dipoles. We see a loss of approximately 25 millihartrees in the bonding strength of $_1H_1He_1H_1$, but in $_1H_2Be_2H_1$, where the distance between the two hydrogen atoms has almost doubled, a loss of only approximately 5 millihartrees is observed. Additionally, the magnitude of the effect is dependent upon the types of dipoles interacting. As mentioned above, the $H_1H_1H_1$ molecule gains roughly 15 millihartrees from the additional dipole interaction, while $_1Li_2H_2Li_2$ gains approximately 50 millihartrees from the interaction between the two Li atoms.

It should be noted that this gain may actually be underestimated by our calculations due to the large equilibrium bond length we find in this molecule. As stated in Sec. 3.4.2, we believe that our basis set is insufficient for accurately describing bonds with lengths above 6 bohrs. It appears that the bonds in triatomic molecules are noticeably longer than those in diatomics, even when favourable dipole interactions increase the strength of the molecule's bonds. This leads to very long bond lengths for Li containing molecules, and due to concerns of numerical accuracy we have omitted results for other Li containing triatomic molecules.

3.5   Conclusion

We have written a program, called CHEM1D, for computing energies of 1D molecules at the Hartree-Fock level and beyond. This has enabled us to probe the behaviour of 1D molecules that were previously unaccessible and to develop a more sophisticated understanding of 1D chemistry.

In particular, we have shown that multi-electron bonds do exist in 1D molecules, notwithstanding our earlier conjecture to the contrary. We have also observed some strange effects that result from the limited dimensionality of the system, such as unusually strong bonding between distant atoms and the dominance of permanent dipoles upon reactivity.

Our experiments have highlighted the limitations of the CHEM1D program and, in particular, it appears that our choice of basis functions is non-optimal. This presents avenues for continued development that will be pursued in the following chapters in order to allow the examination of yet more complicated 1D systems.

Chapter 4

LegLag: an improved 1D electronic structure theory program

4.1 Introduction

In Chapter 3 we described a program for electronic structure calculations on one-dimensional (1D) molecules called CHEM1D. Unlike previous workers who used softened[48, 49] or otherwise altered[51–54, 58–60] interelectronic interactions in their studies of 1D chemical systems, CHEM1D employs the unadorned Coulomb operator $|x|^{-1}$. This potential introduces a non-integrable singularity which requires special treatment. CHEM1D avoids Coulombic divergences by solving the Schrödinger equation with Dirichlet boundary conditions that require the wavefunction to vanish wherever two particles touch. See Sec. 2.1 for more discussion of this strategy.

The Dirichlet conditions have three chemically interesting consequences. First, that molecular energies are spin-blind, i.e. they are invariant with respect to spin-flips. Second, that a "super-Pauli" exclusion rule applies, i.e. an orbital cannot be occupied by more than one electron. Third, that the nuclei become impenetrable, i.e. electrons are unable to tunnel through them [62–66, 68, 144].

Unfortunately, in Chapter 3 we identified that CHEM1D suffers from debilitating numerical stability issues. The CHEM1D program uses basis sets related to the exact wavefunctions of the hydrogen molecule cation H_2^+ but these quickly develop near-linear-dependence problems that prevent CHEM1D from achieving basis set convergence even for relatively modest molecular systems. In this chapter we describe
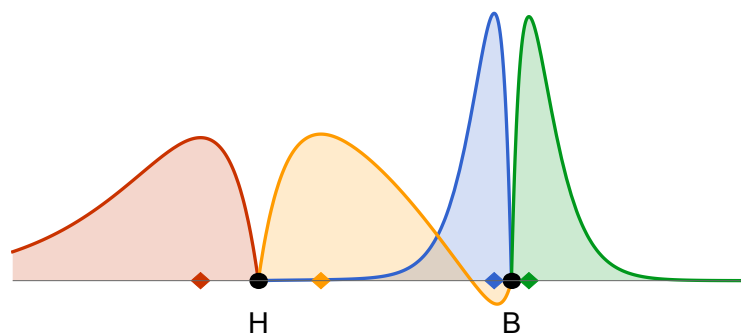
Figure 4.1: The ground state of the HLi molecule. Black circles represent the nuclei. Each coloured region represents a singly-occupied orbital and the corresponding coloured diamond shows the most likely position of the electron.

LegLag, a more numerically stable version of Chem1D, which can be applied to a wider range of molecules to gain deeper insight into 1D chemistry.

4.2   Theory

Under Dirichlet boundary conditions, nuclei are impenetrable to electrons and each electron in a molecule is therefore confined to a ray or line segment by the nuclei closest to it. In this way, the M nuclei divide 1D space into two semi-infinite domains and $M - 1$ finite domains. Each domain supports a set of orbitals that vanish at the boundaries of the domain and outside it. Figure 4.1 illustrates this for a small diatomic molecule.

In order to specify the constitution of a molecule, we employ a notation in which atomic symbols indicate nuclei and subscripts indicate the numbers of electrons in the intervening domains. For example, $_1\mathrm{Li}_4\mathrm{B}_3\mathrm{H}_1$ is a triatomic with a lithium, boron and hydrogen nucleus arranged from left to right in that order. There is one electron to the left of the lithium nucleus, four electrons between the lithium and the boron nuclei, three electrons between the boron and the hydrogen nuclei and one electron to the right of the hydrogen nucleus. In the present work, we consider only ground states and assume that the n electrons within a domain singly occupy the n lowest-energy orbitals.

4.3   Basis sets

There are three types of domain – left, right and middle – and we require a set of basis functions for each. The functions should vanish at the domain boundaries and form a complete set.
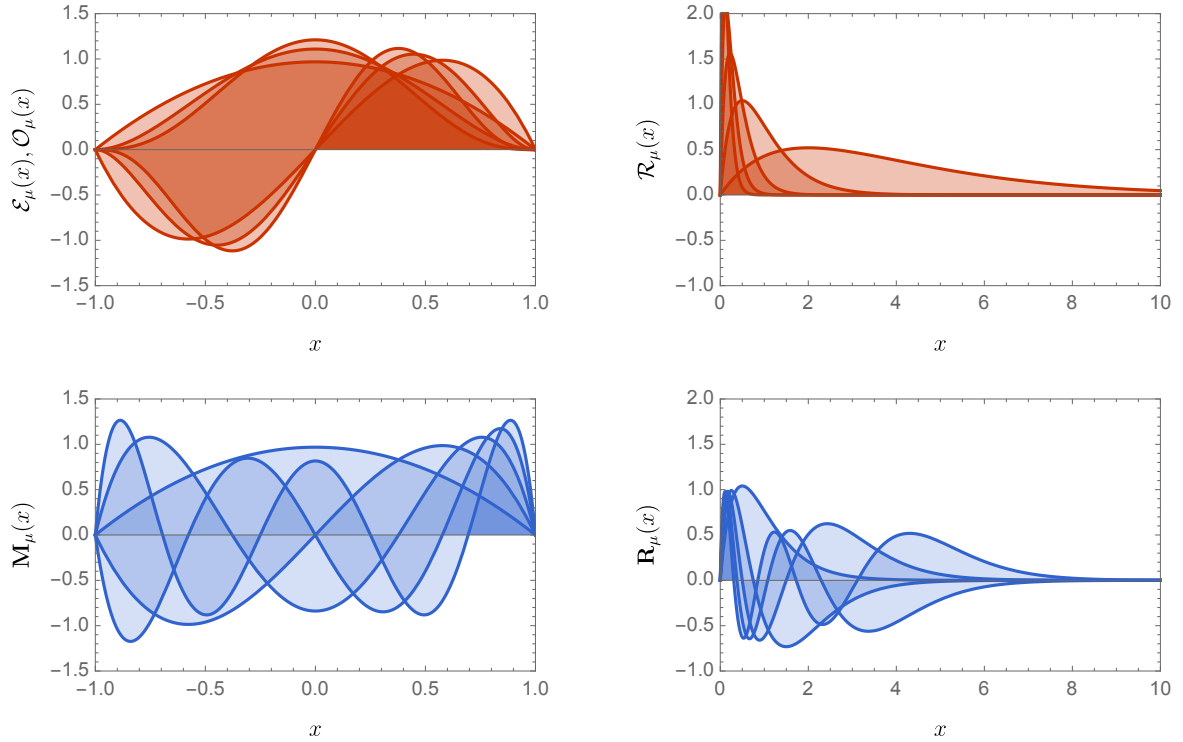
Figure 4.2:   A comparison of the basis functions used in CHEM1D (top row, in red) with those used in LEGLAG (bottom row, in blue). A finite domain is shown on the left and an infinite domain on the right.

The basis set used by CHEM1D is described in Sec. 3.2.1. We redefine the functions here with slightly different notation for comparison with the functions used by LEGLAG

$$\mathcal{L}_\mu(s) = 2\mu^3 \alpha^{1/2} s \exp(-\mu^2 s) \tag{4.1a}$$

$$\mathcal{R}_\mu(t) = 2\mu^3 \alpha^{1/2} t \exp(-\mu^2 t) \tag{4.1b}$$

$$\mathcal{E}_\mu(z) = \sqrt{\frac{(2\mu+1)_{1/2}}{\omega \pi^{1/2}}} \, (1 - z^2)^\mu \tag{4.1c}$$

$$\mathcal{O}_\mu(z) = \sqrt{\frac{2(2\mu+1)_{3/2}}{\omega \pi^{1/2}}} \, z(1 - z^2)^\mu \tag{4.1d}$$

where $s = \alpha(A - x)$, $t = \alpha(x - B)$ and $z = (x - C)/\omega$ are the reduced coordinates in the left, right and middle domains respectively, A and B are the positions of the leftmost and rightmost nuclei, and C and ω are the center and halfwidth of a middle domain, $\alpha > 0$ is an exponent, and $(a)_n$ is the Pochhammer symbol [100]. The $\mathbb{L}_\mu$ and $\mathbb{R}_\mu$ functions are used in the left and right domains, respectively. $\mathbb{E}_\mu$ and $\mathbb{O}_\mu$ are used in the middle domains.

One disadvantage of these functions is their increasing linear dependence as the size of the basis set grows, which creates numerical instability in the orthogonalisation step of the Hartree-Fock (HF) method [8]. This limits the size of basis set that can be employed before unacceptable numerical precision is lost.

A second disadvantage of this basis set is that, because the $\mathbb{E}_\mu$ and $\mathbb{O}_\mu$ functions are increasingly peaked around the middle of the domain, they struggle to describe details of the molecular orbitals near domain boundaries. As we discovered in Chapter 3 this becomes particularly problematic when the domain contains more than one electron.

In contrast, LegLag uses the basis functions

$$\mathbf{L}_\mu(s) = \sqrt{\frac{8\alpha}{(\mu)_2}}\; s L_{\mu-1}^2(2s) \exp(-s) \tag{4.2a}$$

$$\mathbf{R}_\mu(t) = \sqrt{\frac{8\alpha}{(\mu)_2}}\; t L_{\mu-1}^2(2t) \exp(-t) \tag{4.2b}$$

$$\mathbf{M}_\mu(z) = \sqrt{\frac{\mu+3/2}{\omega(\mu)_4}}\; P_{\mu+1}^2(z) \tag{4.2c}$$

where L_m^2 and P_m^2 are second-order associated Laguerre and Legendre polynomials [100]. (Our package's name stems from its use of Legendre and Laguerre polynomials.) The $\mathbf{L}_\mu$, $\mathbf{R}_\mu$ and $\mathbf{M}_\mu$ functions are used in the left, right and middle domains, respectively, and are mutually orthogonal. The $\mathbf{M}_\mu$ are evenly distributed across the domain, as Fig. 4.2 shows.

4.4   Integrals

In Chapters 2 and 3 we discovered that HF calculations [8] give unexpectedly accurate results in 1D. It also appears, in contrast to the situation in 3D [16, 145, 146], that the Møller-Plesset (MP) perturbation series [8] in 1D often converges rapidly to the exact energy. However, to perform such calculations it is necessary to evaluate the integrals

$$(\mathbf{F}_\mu|\mathbf{F}_\nu) = \int \mathbf{D}_{\mu\nu}(x)\,dx = \delta_{\mu\nu} \tag{4.3a}$$

$$(\mathbf{F}_\mu|\hat{T}|\mathbf{F}_\nu) = \frac{1}{2}\int \mathbf{F}'_\mu(x)\mathbf{F}'_\nu(x)\,dx \tag{4.3b}$$

$$(\mathbf{F}_\mu|\hat{V}|\mathbf{F}_\nu) = \int \frac{\mathbf{D}_{\mu\nu}(y)}{|x-y|}\,dy \tag{4.3c}$$

$$(\mathbf{F}_\mu\mathbf{F}_\nu|\mathbf{F}_\lambda\mathbf{F}_\sigma) = \iint \frac{\mathbf{D}_{\mu\nu}(x)\mathbf{D}_{\lambda\sigma}(y)}{|x-y|}\,dx\,dy \tag{4.3d}$$

where $\mathbf{F} \in \{\mathbf{L}, \mathbf{M}, \mathbf{R}\}$, $\mathbf{D}_{\mu\nu}(x) = \mathbf{F}_\mu(x)\mathbf{F}_\nu(x)$ is a density component, $\hat{T} = -\nabla^2/2$ is the kinetic energy operator and $\delta_{\mu\nu}$ is the Kronecker delta function [100].

If the four basis functions are in the same domain, the singularity of the Coulomb operator causes $(\mathbf{F}_\mu\mathbf{F}_\nu|\mathbf{F}_\lambda\mathbf{F}_\sigma)$ to diverge. However, the antisymmetrized integral

$$(\mathbf{F}_\mu\mathbf{F}_\nu||\mathbf{F}_\lambda\mathbf{F}_\sigma) = (\mathbf{F}_\mu\mathbf{F}_\nu|\mathbf{F}_\lambda\mathbf{F}_\sigma) - (\mathbf{F}_\mu\mathbf{F}_\sigma|\mathbf{F}_\lambda\mathbf{F}_\nu) \tag{4.4}$$

is finite and can be found from quasi-integrals, described in Sec. 3.2.4, using

$$(\mathbf{F}_\mu\mathbf{F}_\nu||\mathbf{F}_\lambda\mathbf{F}_\sigma) = \{\mathbf{F}_\mu\mathbf{F}_\nu|\mathbf{F}_\lambda\mathbf{F}_\sigma\} - \{\mathbf{F}_\mu\mathbf{F}_\sigma|\mathbf{F}_\lambda\mathbf{F}_\nu\} \tag{4.5}$$

Because $\mathbf{R}_\mu$ is the image of $\mathbf{L}_\mu$ under inversion through the molecular mid-point, formulae involving only $\mathbf{R}_\mu$ and $\mathbf{M}_\mu$ can be found from the equivalent formulae involving $\mathbf{L}_\mu$ and $\mathbf{M}_\mu$. We will therefore not discuss the former.

For the following definitions it will be convenient to define the parity function

$$\epsilon_n = \begin{cases} 1, & n \text{ is even} \\ 0, & n \text{ is odd} \end{cases} \tag{4.6}$$

4.4.1 One-electron integrals

If we assume $\mu \leq \nu$, the kinetic integrals are

$$(\mathbf{R}_\mu|\hat{T}|\mathbf{R}_\nu) = \frac{\alpha^2}{2}\left[\frac{(2\mu)_3}{6\sqrt{(\mu)_2(\nu)_2}} - \delta_{\mu\nu}\right] \tag{4.7}$$

and

$$(\mathbf{M}_\mu|\hat{T}|\mathbf{M}_\nu) = \epsilon_{\mu+\nu}\sqrt{\frac{(\mu)_4(\mu+\frac{3}{2})(\nu+\frac{3}{2})}{(\nu)_4}}\frac{\mu^2 + 3\mu - 1}{6\omega^2} \tag{4.8}$$

The potential to the left of $\mathbf{R}_\mu\mathbf{R}_\nu$ is

$$(\mathbf{R}_\mu|\hat{V}|\mathbf{R}_\nu) = \frac{2\alpha\, L^2_{\mu-1}(2t)}{\sqrt{(\mu)_2(\nu)_2}}\, (\nu+1)!\, U(\nu, -1, -2t) \tag{4.9}$$

where U is Tricomi's function [100].

The potentials to the left($+$) or right($-$) of $\mathbf{M}_\mu\mathbf{M}_\nu$ are

$$(\mathbf{M}_\mu|\hat{V}|\mathbf{M}_\nu) = \pm\frac{2}{\omega}\sqrt{\frac{(\mu+\frac{3}{2})(\nu+\frac{3}{2})}{(\mu)_4(\nu)_4}}\, P^2_{\mu+1}(z)\, Q^2_{\nu+1}(z) \tag{4.10}$$

where Q^2_m is a second-order associated Legendre function of the second kind [100].

4.4.2   Clebsch-Gordan expansions

Products of our basis functions have finite expansions

$$\mathbf{L}_\mu(s)\mathbf{L}_\nu(s) = \sum_n a_n^{\mu\nu}\mathbb{L}_n(s) \tag{4.11a}$$

$$\mathbf{R}_\mu(t)\mathbf{R}_\nu(t) = \sum_n a_n^{\mu\nu}\mathbb{R}_n(t) \tag{4.11b}$$

$$\mathbf{M}_\mu(z)\mathbf{M}_\nu(z) = \sum_n b_n^{\mu\nu}\mathbb{M}_n(z) \tag{4.11c}$$

where the expansion functions are

$$\mathbb{L}_n(s) = \frac{8\alpha}{(n)_2}\ s^2 L_{n-1}^2(2s)\exp(-2s) \tag{4.12a}$$

$$\mathbb{R}_n(t) = \frac{8\alpha}{(n)_2}\ t^2 L_{n-1}^2(2t)\exp(-2t) \tag{4.12b}$$

$$\mathbb{M}_n(z) = \frac{1}{2(n)_4\omega}\ (1-z^2)\ P_{n+1}^2(z) \tag{4.12c}$$

and n ranges from $|\mu - \nu| + 1$ to $\mu + \nu - 1$. For example,

$$\mathbf{R}_2\mathbf{R}_3 = 2\sqrt{2}\ \mathbb{R}_2 - 4\sqrt{2}\ \mathbb{R}_3 + 5\sqrt{2}\ \mathbb{R}_4 \tag{4.13a}$$

$$\mathbf{M}_2\mathbf{M}_3 = 10\sqrt{21}\ \mathbb{M}_2 - 0\mathbb{M}_3 + 35\sqrt{21}\ \mathbb{M}_4 \tag{4.13b}$$

We call these Clebsch-Gordan (CG) expansions and the coefficients are given by

$$a_n^{\mu\nu} = \int_0^\infty \frac{L_{\mu-1}^2(t)}{\sqrt{(\mu)_2}}\frac{L_{\nu-1}^2(t)}{\sqrt{(\nu)_2}}\frac{L_{n-1}^2(t)}{t^{-2}\exp(t)}\ dt \tag{4.14a}$$

$$b_n^{\mu\nu} = \int_{-1}^1 \frac{P_{\mu+1}^2(z)}{\sqrt{\frac{(\mu)_4}{\mu+3/2}}}\frac{P_{\nu+1}^2(z)}{\sqrt{\frac{(\nu)_4}{\nu+3/2}}}\frac{P_{n+1}^2(z)}{\frac{1-z^2}{2n+3}}\ dz \tag{4.14b}$$

4.4.3   Properties of the expansion functions

The Laplace transforms of $\mathbb{R}_n$ and $\mathbb{M}_n$ are

$$\int_0^\infty \mathbb{R}_n(t)\exp(-ut)\ dt = \frac{\alpha(u/2)^{n-1}}{(1+u/2)^{n+2}} \tag{4.15a}$$

$$\int_{-1}^1 \mathbb{M}_n(z)\exp(-uz)\ dz = \frac{(-1)^{n+1}i_{n+1}(u)}{\omega u^2} \tag{4.15b}$$

where i_n is a modified spherical Bessel function [100].

The moments of $\mathbb{R}_n$ and $\mathbb{M}_n$ are

$$\int_0^\infty \mathbb{R}_n(t)\, t^k \, dt = \frac{\alpha(-1)^{n+1}\, k!\, (k+2)!}{(n+1)!\, \Gamma(2+k-n)\, 2^k} \tag{4.16a}$$

$$\int_{-1}^1 \mathbb{M}_n(z)\, z^k \, dz = \frac{\epsilon_{n+k+1}\, \Gamma(\frac{k+1}{2})\, \Gamma(\frac{k+2}{2})}{8\omega\, \Gamma(\frac{k-n+3}{2})\, \Gamma(\frac{k+n+6}{2})} \tag{4.16b}$$

The kth moments of $\mathbb{R}_n$ and $\mathbb{M}_n$ vanish if $k < n - 1$. All of the higher moments of $\mathbb{R}_n$ have the same sign. All of the higher moments of $\mathbb{M}_n$ are positive.

The potential to the left of $\mathbb{R}_n$ is

$$\int_0^\infty \frac{\mathbb{R}_n(r)}{r-t}\, dr = 2\alpha\,\Gamma(n)\, U(n, -1, -2t) \tag{4.17}$$

and the potentials to the right$(+)$ or left$(-)$ of $\mathbb{M}_n$ are

$$\int_{-1}^1 \frac{\mathbb{M}_n(r)}{|z-r|}\, dr = \frac{\pm\Gamma(\frac{n}{2})\,\Gamma(\frac{n+1}{2})}{\Gamma(n+\frac{5}{2})\, 8\,\omega\, z^n}\, F\left[\frac{n}{2}, \frac{n+1}{2}, n+\frac{5}{2}, \frac{1}{z^2}\right] \tag{4.18}$$

where F is the Gauss hypergeometric function [100]. These potentials, which are illustrated in Fig. 4.3, are monotonically decreasing and behave asymptotically as $O(x^{-n})$.

The absolute contents of $\mathbb{R}_n$ and $\mathbb{M}_n$ satisfy

$$\int_0^\infty |\mathbb{R}_n(t)|\, dt < \frac{10}{9n^{5/4}} \tag{4.19a}$$

$$\int_{-1}^1 |\mathbb{M}_n(z)|\, dz < \frac{1}{10n^2} \tag{4.19b}$$

These quantities can be used to compute simple upper bounds to the Coulomb integrals in the next Section.
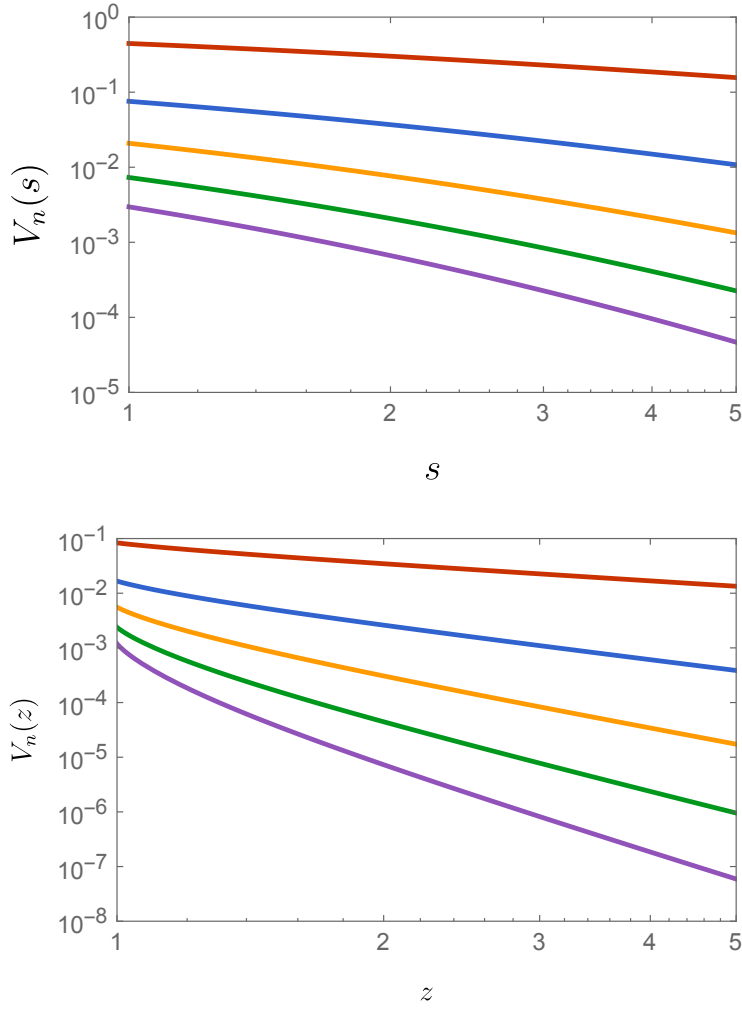
Figure 4.3:   Potential $V_n(x)$ to the right of $\mathbb{L}_n(s)$ with $\alpha = 1$ (top) and to the right of $\mathbb{M}_n(z)$ with $\omega = 1$ (bottom). From top to bottom, $n = 1, 2, 3, 4, 5$.

4.4.4   Coulomb integrals

The CG expansions yield

$$(\mathbf{L}_\mu \mathbf{L}_\nu | \mathbf{R}_\lambda \mathbf{R}_\sigma) = \sum_{mn} a_m^{\mu\nu} a_n^{\lambda\sigma} (\mathbb{L}_m | \mathbb{R}_n) \tag{4.20a}$$

$$(\mathbf{L}_\mu \mathbf{L}_\nu | \mathbf{M}_\lambda \mathbf{M}_\sigma) = \sum_{mn} a_m^{\mu\nu} b_n^{\lambda\sigma} (\mathbb{L}_m | \mathbb{M}_n) \tag{4.20b}$$

$$(\mathbf{M}_\mu \mathbf{M}_\nu | \mathbf{M}_\lambda \mathbf{M}_\sigma) = \sum_{mn} b_m^{\mu\nu} b_n^{\lambda\sigma} (\mathbb{M}_m | \mathbb{M}_n) \tag{4.20c}$$

The Coulomb integral between densities $f(x - X)$ and $g(y - Y)$ in different domains with $X \leq Y$ is given by

$$(f|g) = \int_0^\infty F(-u) G(u) \exp(-Ru) \, du \tag{4.21}$$

where F and G are the Laplace transforms of f and g and $R = Y - X$. In this way, we find that

$$(\mathbb{L}_m | \mathbb{R}_n) = 2\alpha \, \Gamma(m+n-1) \, U(m+n-1, -4, 2\alpha R) \tag{4.22a}$$

$$(\mathbb{L}_m | \mathbb{M}_n) = -\frac{(-\alpha \, \omega)^n \sqrt{\pi}}{4\omega}$$
$$\times \sum_{k=0}^{\infty} \frac{\Gamma(m+n-1+2k) \, U(m+n-1+2k, n-2+2k, 2\alpha R)}{\Gamma(n+\frac{5}{2}+k)} \frac{(\alpha^2 \omega^2)^k}{k!} \tag{4.22b}$$

$$(\mathbb{M}_m | \mathbb{M}_n) = \frac{\pi}{64R} \left(\frac{\omega_1}{2R}\right)^{m-1} \left(-\frac{\omega_2}{2R}\right)^{n-1}$$
$$\times \sum_{k=0}^{\infty} \frac{\Gamma(m+n-1+2k)}{\Gamma(m+\frac{5}{2}+k)\Gamma(n+\frac{5}{2}+k)} \left[\frac{\omega_2^2 - \omega_1^2}{4R^2}\right]^k P_k^{(m+\frac{3}{2}, n+\frac{3}{2})} \left[\frac{\omega_2^2 + \omega_1^2}{\omega_2^2 - \omega_1^2}\right] \tag{4.22c}$$

where U is Tricomi's function and $P_k^{(a,b)}$ is a Jacobi polynomial [100].

Each of the integrals (4.22a), (4.22) and (4.22) is $O(1/R^{m+n-1})$ for large R. Consequently, for domains that are far apart, many of the higher Coulomb integrals are negligible and can be safely neglected using the bound (4.28).

4.4.5 Quasi-integrals

The CG expansions yield

$$\{\mathbf{R}_\mu \mathbf{R}_\nu | \mathbf{R}_\lambda \mathbf{R}_\sigma\} = \sum_{mn} a_m^{\mu\nu} a_n^{\lambda\sigma} \{\mathbb{R}_m | \mathbb{R}_n\} \tag{4.23a}$$

$$\{\mathbf{M}_\mu \mathbf{M}_\nu | \mathbf{M}_\lambda \mathbf{M}_\sigma\} = \sum_{mn} b_m^{\mu\nu} b_n^{\lambda\sigma} \{\mathbb{M}_m | \mathbb{M}_n\} \tag{4.23b}$$

The quasi-integral, which have been described in Sec. 3.2.4, between densities $f(x)$ and $g(y)$ in the same domain is given by

$$\{f|g\} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} F(-ik)G(ik) \log k^2 \, dk \tag{4.24}$$

If we define the harmonic sum

$$H_n = \sum_{k=1}^{n} \frac{1}{2k-1} \tag{4.25}$$

assume $m \leq n$ and define $\Delta = n - m$, we find that

$$\{\mathbb{R}_m | \mathbb{R}_n\} = \alpha \frac{(-1)^{n+1} \Delta!}{(n+1)!} \frac{\sqrt{\pi}}{2} \sum_{k=0}^{\Delta/2} \frac{(2k+1)(2k+3)(H_{k+2} - H_{n-1-k})}{4^k \, \Gamma(3/2 - n + k) \, (\Delta - 2k)! \, k!} \qquad (4.26)$$

and

$$\{\mathbb{M}_m | \mathbb{M}_n\} = \frac{\epsilon_{m+n}}{64\omega \left(m + \frac{\Delta-1}{2}\right)_5} \times \begin{cases} +12(\frac{7}{12} - H_{m-1} - H_{m+4}), & \Delta = 0 \\ -8(\frac{2}{3} - H_m - H_{m+5}), & \Delta = 2 \\ +2(\frac{25}{24} - H_{m+1} - H_{m+6}), & \Delta = 4 \\ -24 \, / \left(\frac{\Delta}{2} - 2\right)_5, & \Delta \geq 6 \end{cases} \qquad (4.27)$$

4.5   Implementation

LegLag closely follows most of the algorithms employed in Chem1D, which have been described in Chapter 3. There are, however, some important distinctions between the two programs apart from the integral package, the main point of differentiation. LegLag also contains an efficient implementation of both the MP2 and MP3 levels of Møller-Plesset perturbation theory. This implementation takes full advantage of the domain structure of 1D molecules.

We let $\mathcal{D}$ be the set of domains in a molecule, $\mathcal{O}_d$ and $\mathcal{V}_d$ be the set of occupied and virtual orbitals in the domain d and ε_a be the energy of the orbital a. Then Alg. 9 shows the structure of the MP2 code in LegLag, in particular its separation over the domain structure of the molecule. This algorithm avoids pairing up occupied orbitals with virtual orbitals in other domains, eliminating large parts of the MP2 summation that have a strictly zero contribution to the total energy. The MP3 algorithm is significantly longer but not more complicated than that of the MP2 algorithm, and so we have not included it here.

Unlike Chem1D, which was written in Fortran 90, LegLag has been implemented using the Python programming language (version 3.5) in combination with the Cython language extension for optimising compute-intensive bottlenecks. It employs the external Numpy library for data structures and linear algebra operations and the Scipy library for computing some of the special functions which appear in Secs. 4.4.1, 4.4.4 and 4.4.5.

However, because we invariably need a range of values for the a and b parameters in the Tricomi confluent hypergeometric functions [100] $U(a, b, z)$ required for Coulomb integrals involving the $\mathbf{L}_\mu$ functions, it is more efficient to compute these functions recursively. It has been shown that backwards recurrence in the a parameter is

---

Algorithm 9 Domain based MP2 calculation

1: **procedure** MP2
2:     **for** $d \in \mathcal{D}$ **do**
3:         **for** $a, b \in \mathcal{O}_d$ **do**
4:             **for** $r, s \in \mathcal{V}_d$ **do**
5:                 Accumulate $\frac{\{ar|bs\} - \{as|br\}}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s}$
6:             **end for**
7:         **end for**
8:     **end for**
9:     **for** $p \in \mathcal{D}$ **do**
10:         **for** $q \in \mathcal{D}/p$ **do**
11:             **for** $a \in \mathcal{O}_p, b \in \mathcal{O}_q$ **do**
12:                 **for** $r \in \mathcal{V}_p, s \in \mathcal{V}_q$ **do**
13:                     Accumulate $\frac{(ar|bs)}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s}$
14:                 **end for**
15:             **end for**
16:         **end for**
17:     **end for**
18: **end procedure**

---

numerically stable and our algorithm exploits this [147]. To obtain the starting values for this recurrence we use an asymptotic expansion that is valid when $2a - b$ is large and positive [148, 149]. Our numerical experiments have shown that for arguments $z > 10$ this expansion converges at an unacceptable rate. Our algorithm therefore uses Miller's method [100, 150] when $z > 8$.

We also detect and avoid computing negligible integrals using the Coulomb upper bound [151]

$$|(P|Q)| \leq \min(V_P^* S_Q^*, S_P^* V_Q^*) \tag{4.28}$$

where V_P^* is the maximum potential of $P(x)$ in the domain of $Q(x)$ and S_Q^* is the integral of $|Q(x)|$. The V^* and S^* values can be found using expressions in Sec. 4.4.3.

4.6   Conclusion

In Chapter 3 we saw that CHEM1D was severely limited by the numerical instability of the basis orthogonalisation step in the Roothaan-Hall equations. By using an orthogonal basis set LEGLAG removes the requirement for this step and the numerical difficulties with it. We also argue that the Legendre functions used by LEGLAG for the finite domains will better model orbitals which localise close to the bordering nuclei. Validation of this conjecture will wait until the next chapter where we will use LEGLAG for further experiments on 1D molecular structure.

Chapter 5

Molecular structure in one dimension

5.1 Introduction

By changing to a more numerically stable basis set, LegLag promises to be a significant improvement over Chem1D. This allows us to perform a more extensive study of how molecules form in 1D when employing the unmodified Coulomb operator and the Dirichlet boundary condition.

The severe effects of the Dirichlet condition means that this model does not reflect the same type of experimental systems as the "quasi-1D" methods characterised by softened Coulomb interactions, which permit the nuclei to be penetrable and electrons to pair within spatial orbitals. This reflects situations where the 1D confinement is not strict, and so they are well suited to modelling confined experimental systems such as ultracold atoms confined within a 1D trap [47, 152, 153].

In contrast, the Coulomb interaction used by LegLag describes particles which are *strictly* restricted to move within a 1D sub-space of three-dimensional space. Early models of 1D atoms using this interaction have been used to study the effects of external fields upon Rydberg atoms [70, 71] and the dynamics of surface-state electrons in liquid helium [72, 73]. This description of 1D chemistry also has interesting connections with the exotic chemistry of ultra-high magnetic fields (such as those in white dwarf stars), where the electronic cloud is dramatically compressed perpendicular to the magnetic field [154–156]. In these extreme conditions, where magnetic effects compete with Coulombic forces, entirely new bonding paradigms emerge [154–161].
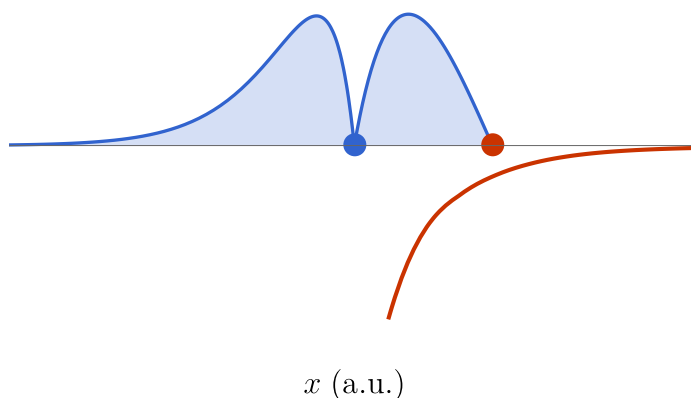
x (a.u.)

Figure 5.1: The exclusion potential (red) of a $_1$He$_1$ atom (blue). The blue regions show the occupied orbitals when computing the exclusion potential at the position of the red dot.

To begin to understand the nature of chemical bonding in 1D, we have performed an extensive search for stable molecules. After presenting accurate atomic energies in Sec. 5.3, we will discuss the structures of a wide variety of small molecules in Secs. 5.4, 5.5 and 5.6 along with a small set of polymeric systems in Sec. 5.7. This will allow us to deduce a set of simple rules that govern bonding in 1D molecules, which we present in Sec. 5.8.

We also present a generalisation of the electrostatic potential for application to the 1D molecules in this study, which we refer to as the exclusion potential. This will be used to aid in the understanding of why some molecules form. The theory of the exclusion potential is presented in Sec. 5.2.

As we determined in Chapter 2, the Periodic Table in 1D has only two groups and we will frequently refer to alkalis (H, Li, B, ... which have an odd number of electrons and a permanent dipole moment) and nobles (He, Be, C, ... which have an even number of electrons, are symmetrical and have no dipole.).

All of the calculations that we report use 30 basis functions in each of the left and right domains and 50 functions in each of the finite domains. We will refer to this as the (30,50) basis set. We report only the digits that have converged as the basis set is increased to the (30,50) set. For determining optimised geometries we use the numerical function minimisation routines available in Scipy.

Unless otherwise stated, atomic units are used throughout.

5.2 Exclusion potential

In 3D, the molecular electrostatic potential [162] (MESP) is the limit of the ratio of the Coulomb energy of a test particle to the magnitude of its charge, as that charge approaches zero. It is a potent tool for understanding chemical behaviour and can reveal, for example, electrophilic or nucleophilic regions. Unfortunately, however, the MESP diverges at all points in a 1D system except where the electron density vanishes [66]. Therefore, to define a meaningful potential in a 1D molecule, we must insist that the test particle create a new Dirichlet node at its position. We call the resulting potential the "exclusion potential" to emphasise that, in contradistinction to the 3D analog, the test particle in 1D excludes electrons from its neighbourhood and thereby significantly perturbs the system.

Note that this exclusion potential has a "path dependence" because the test particle creates new domains. We only ever consider the test particle to have approached from $\pm\infty$. This confines the electrons in the appropriate infinite domain into the new finite domain between the test particle and the molecule. This is to replicate what happens as atoms and molecules approach one another.

Figure 5.1 shows the exclusion potential for a $_1\text{He}_1$ atom as well as the perturbed orbitals for a given position of the test particle. Note how the Dirichlet node created by the test particle compresses the right orbital, and prevents the electron which occupies it from extending to the right.

5.3 Atoms

Our first task was to choose the exponent α in $\mathbf{L}_\mu(x)$ and $\mathbf{R}_\mu(x)$ that yields the best energies as the basis set is increased. We expected that α would be determined largely by the innermost (and lowest-energy) orbitals, and that therefore $\alpha \approx Z$, where Z is the nuclear charge of the atom in question. We were therefore surprised to find that this is not the case and that, except for hydrogen, the optimal exponent is always close to $\alpha = 2$, a compromise that attempts to describe both the compact inner orbitals and the diffuse outer orbitals. Figure 5.2 gives a graphical representation of some of the data which lead us to this conclusion. After this discovery, we used $\alpha = 2$ for all atoms.

In Chapter 2, our first foray into 1D chemistry, we used multiple-precision arithmetic in *Mathematica*[99] to compute the near-exact HF, MP2 and MP3 energies, ionisation energies and electron affinities of the ground-state atoms up to $_5\text{Ne}_5$. Our subsequent (double-precision) CHEM1D program was often unable to reproduce

Atom	Total energies			Ionization energies			Electron affinities[1]		
	$-E_{\mathrm{HF}}$	$-E_{\mathrm{MP2}}$	$-E_{\mathrm{MP3}}$	HF	MP2	MP3	HF	MP2	MP3
H	0.500000	0.500000	0.500000	13.606	13.606	13.606	3.893	3.939	3.961
He	3.242922	3.244986	3.245611	33.822	33.878	33.895	—	—	—
Li	8.007756	8.01112	8.01179	4.486	4.517	4.522	1.395	1.410	1.414
Be	15.415912	15.4226	15.4236	10.348	10.400	10.408	—	—	—
B	25.35751	25.3671	25.3684	2.068	2.09	2.099	0.64	0.65	0.65
C	38.09038	38.105	38.107	4.670	4.719	4.73	—	—	—
N	53.569	53.59	53.6	1.1	1.1	1.1	0.3	0.3	0.3
O	71.9293	71.95	71.96	2.516	2.548	2.556	—	—	—
F	93.1	93.2	93.2	0.5	0.5	0.5			
Ne	117.31	117.35	117.35	1.5	1.5	1.5	—	—	—

[a]The electron affinities of He, Be, C, O and Ne are omitted because the anions of these species are auto-ionising (Sec. 2.3.3).

Table 5.1: Total energies (in E_{h}), ionization energies and electron affinities (eV) of 1D atoms using the (30,50) basis set.

these energies, principally because of its inadequate basis functions (As discussed in Chapter 3). Table 5.1 shows that our (double-precision) LEGLAG calculations are much more successful in capturing the energies but the (30,50) basis still struggles for the largest atoms and, in particular, fails to yield any significant figures for the electron affinity of $_4\mathrm{F}_5$.

5.4   Diatomics

Notwithstanding the deficiencies of the (30,50) basis for the largest atoms, LEGLAG is able to treat a far wider variety of molecular systems than is possible in CHEM1D and we have surveyed the diatomics with atoms up to oxygen and with all possible electronic configurations that can be generated from the ground-state atoms. Table 5.2 reports the bond lengths and energies of the diatomics that we have found to be stable, i.e. lower in energy than their constituent atoms. This set of results, which greatly extends our previous efforts, allows us significantly more insight into the mechanics of 1D bonding.

Molecule	Bond length			Total energy			Dissociation energy		
AB	HF	MP2	MP3	$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	HF	MP2	MP3
H_1H_1	2.636	2.637	2.638	1.184572	1.185418	1.185728	184.572	185.418	185.728
$_1H_1He_1$	2.025	2.027	2.027	3.880313	3.882619	3.883301	137.39	137.633	137.691
H_2Li_2	5.345	5.323	5.320	8.544163	8.547920	8.548659	36.407	36.800	36.871
$_1H_2Li_1$	5.152	5.141	5.142	8.681782	8.686367	8.687589	174.025	175.25	175.80
$_1H_2Be_2$	3.966	3.961	3.962	16.079548	16.08707	16.08845	163.636	164.50	164.85
$_1H_3B_2$	8.880	8.810	8.806	26.020047	26.0310	26.0329	95.492	96.1	96.1
$_1H_2B_3$	3.298	3.296	3.296	25.957601	25.96793	25.96949	100.093	100.85	101.13
H_3B_3	10.349	10.238	10.235	25.863890	25.8736	25.8748	6.382	6.5	6.5
$_1H_3C_3$	6.666	6.635	6.633	38.756672	38.7721	38.7745	166.290	167.4	167.9
$_1H_4N_3$	14.316	14.268	14.257	54.22372	54.244	54.247	52.470	53	53
$_1H_3N_4$	5.407	5.392	5.392	54.218224	54.2379	54.2407	149.222	150.2	150.5
H_4N_4	19.20	18.168	18.131	54.0703	54.089	54.091	1.3	1	1
$_1H_4O_4$	10.468	10.383	10.378	72.590721	72.616	72.620	110.787	112	112
$_1He_2Li_2$	4.606	4.586	4.584	11.260655	11.266223	11.267543	9.977	10.118	10.144
$_1He_3B_3$	11.174	11.170	11.003	28.600892	28.6126	28.6145	0.461	0.5	0.5
$_1Li_3Li_2$	8.693	8.644	8.637	16.064647	16.07183	16.07326	49.134	49.59	49.69
$_2Li_3Be_2$	7.050	7.000	6.996	23.452479	23.46286	23.46460	28.811	29.16	29.22
$_2Li_4B_2$	13.330	13.228	13.157	33.418876	33.4323	33.4343	53.611	54.1	54.2
$_1Li_4B_3$	14.007	13.999	13.778	33.379031	33.3922	33.3941	13.766	14.0	14.0
$_2Li_4C_3$	10.435	10.358	10.336	46.140625	46.1588	46.1614	42.486	43.0	43
$_2Li_4N_4$	8.956	8.892	8.884	61.588547	61.6112	61.6143	11.788	12.3	12.4
$_2Li_5N_3$	19.552	19.258	19.229	61.63192	61.654	61.658	44.7	45	45
$_1Li_5N_4$	21.546	21.092	21.099	61.5802	61.602	61.605	3.5	3	3
$_2Li_5O_4$	14.943	14.769	17.748	79.987067	80.015	80.019	49.98	51	51
$_2Be_4B_3$	12.566	12.566	12.381	40.776885	40.7932	40.7955	3.464	3.6	3.6
$_2Be_5N_4$	20.571	19.884	19.869	68.9851	69.010	69.014	0.2	0	0
$_2B_5B_3$	19.349	19.233	19.003	50.733908	50.753	50.756	18.891	19	19
$_3B_5C_3$	16.040	16.009	15.779	63.457101	63.481	63.484	9.21	9	9
$_3B_6N_3$	26.480	25.946	25.912	78.949	78.977	78.981	22	22	22
$_2B_6N_4$	27.514	26.939	26.869	78.933	78.961	78.96	6	6	6
$_3B_6O_4$	21.138	20.799	20.735	97.3018	97.34	97.34	14.9	15	15
$_3C_6N_4$	22.880	24.906	24.801	91.660	91.69	91.70	1	1	1
$_3N_7N_4$	30.301	31.892	33.989	107.14	107.18	107.19	10	10	10
$_4N_7O_4$	29.583	28.780	28.727	125.50	125.54	125.55	0	0	0

Table 5.2: Equilibrium bond lengths (bohr), total energies (E_h) and dissociation energies (mE_h) of diatomic molecules.
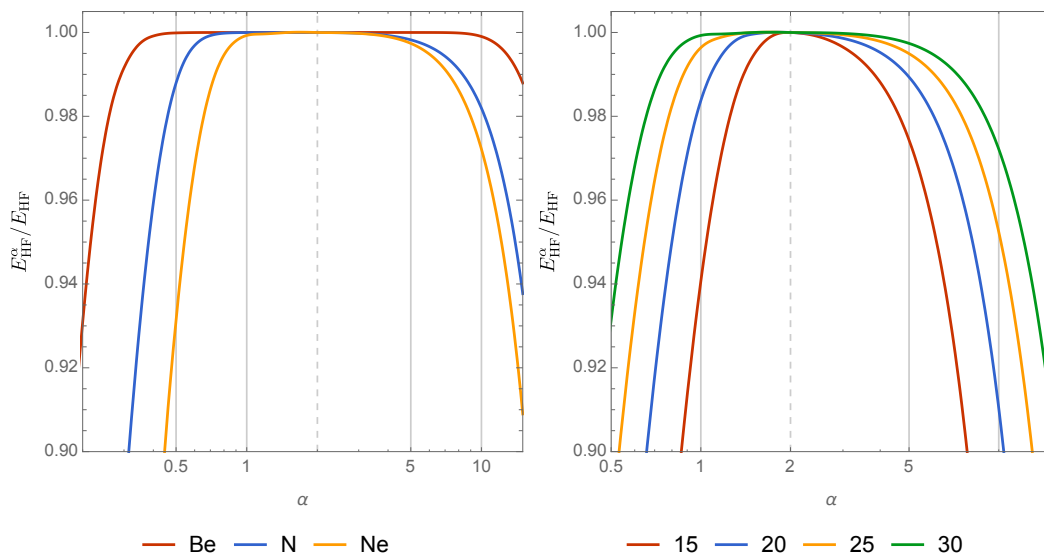
Figure 5.2: Ratios of the Hartree-Fock energy as a function of α (E_{HF}^{α}) to the lowest Hartree-Fock energy (E_{HF}). The left graph shows the behaviour with respect to α for three different atoms (with 30 basis functions in each domain). The right graph shows the same behaviour for four different basis set sizes (for the neon atom).

There appear to be four major factors that govern the binding between two atoms:

1. Valence attraction to an alkali nucleus

2. Nuclear shielding

3. Dipole interactions

4. Number of occupied domains

and we now discuss each of these in turn, and pictorial representations can be seen in Figure 5.4.

Valence-nucleus attraction (Fig. 5.4a) is strongest on the electron-deficient side of the alkali; on the other side of an alkali, or on either side of a noble, such an interaction is shielded too effectively. The four configurations of the HB molecule, viz. $_1H_3B_2$, $_1H_2B_3$, H_3B_3 and H_4B_2, illustrate this. The first three are bound and, in each, at least one of the atoms presents its electron-deficient side to the other atom. The fourth configuration, in which each atom presents its electron-rich side to the other, is unstable.

It follows that two nobles will not bind, as neither has an electron-deficient side. In 3D, noble gas atoms can bind weakly through dispersion interactions [163–166], but we have not seen evidence of this in 1D. This may be an artefact of the (30,50) basis but we believe that such binding, if it exists, is likely to be very weak.
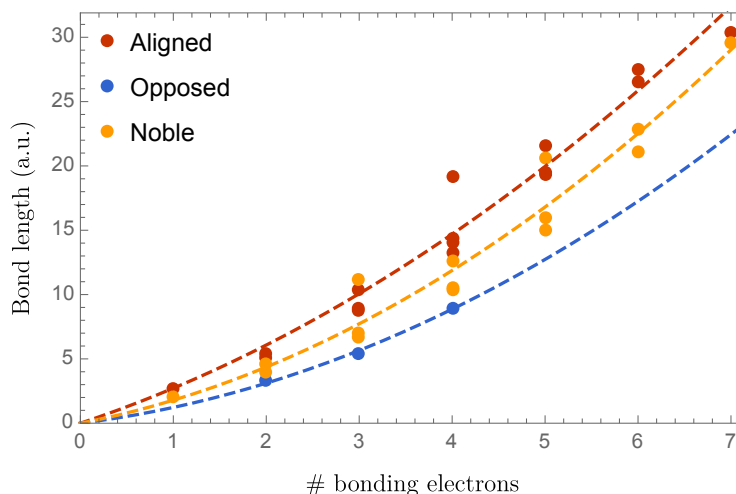
Figure 5.3: Variation of diatomic bond lengths (in bohrs) with the number of electrons in the middle domain. Data are grouped according to the character of the molecule: aligned alkali-alkali, opposed alkali-alkali, and alkali-noble. Quadratic least-square fits are shown as dotted lines.

Nuclear shielding (Fig. 5.4b) is also critical. Lighter atoms bind more strongly because their nuclei are less shielded and this is true *a fortiori* of the completely unshielded H atom. As the shielding increases, binding energies drop rapidly, and bond strengths in nitrogen-containing molecules like $_1\text{Li}_5\text{N}_4$ and $_3\text{C}_6\text{N}_4$ are in the millihartree ($\text{m}E_\text{h}$) range.

Dipole interactions (Fig. 5.4c) also influence bond strengths. The H_1H_1 and $_1\text{H}_1\text{He}_1$ molecules each hold a single electron in the internuclear domain and, naively, one might expect that $_1\text{H}_1\text{He}_1$ would be more strongly bound because of its shorter bond and greater attraction between the bonding electron and nuclei. However, the bond strength in H_1H_1, where the atomic dipoles are favourably aligned, exceeds that in $_1\text{H}_1\text{He}_1$ by roughly 50 $\text{m}E_\text{h}$. Similar arguments explain the relative strengths in H_2Li_2 and $_1\text{He}_2\text{Li}_2$. The rare instances, e.g. $_2\text{Li}_4\text{N}_4$, where a diatomic forms with opposed dipoles are driven by the attraction between two electron-deficient sides.

The number of occupied domains (Fig. 5.4d) is also relevant, as the very different bond energies in the HB diatomics show. Why, for example, is H_3B_3 so much more weakly bound than $_1\text{H}_3\text{B}_2$, despite both having favourable dipole alignments? The answer is that the six electrons are squeezed into two domains in H_3B_3, rather than three in $_1\text{H}_3\text{B}_2$. To form the former from the latter, the electron from the left domain is promoted into a high-energy orbital in the right domain and this incurs a large energy cost.

In Chapter 2 and 3 we discovered a few surprisingly long bonds. However, the results in Table 5.2 show that gargantuan bond lengths are not at all uncommon in
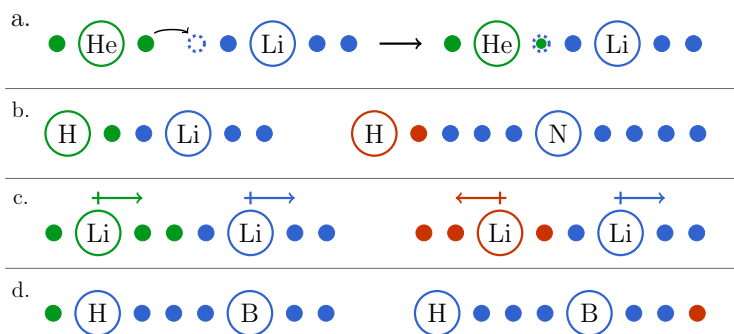
Figure 5.4:   1D analogues of Lewis dot diagrams representing the major factors governing diatomic bonding: a. valence shell interactions, b. nuclear shielding, c. dipole interactions and d. number of occupied domains. More stable configurations are represented in green, less stable in red. Dotted circles represent unoccupied orbitals.

1D. Figure 5.3 reveals a strong correlation between the length of the middle domain of a diatomic and the number of electrons occupying it. If the data are grouped into those with aligned dipoles, those with opposed dipoles and those with a noble atom, strong parabolic trends emerge. Similar behaviour was found for estimates of atomic radii in Chapter 2.

This similarity might suggest that there is no significant distortion of atomic densities during the formation of a diatomic molecule. Figure 5.5 depicts the difference in a selection of diatomic electron densities and the corresponding sum of atomic densities, showing that, in the majority of cases, this is true.

More specifically, this assertion is true when the bonding alkali, i.e. the alkali with its electron deficient side in the bonding domain, is heavier than H. In these cases we see the outermost electron of the other atom occupies the position of the LUMO of the bonding alkali. The orbital that this electron occupies does not appreciably change in shape, however. As a result the bond length is completely determined by the shape of the atomic species.

The exception to this are those diatomics where the bonding alkali is an H atom. The unshielded proton of the H atom is significantly more reactive than other species, an effect we also see when looking at bond strengths. This results in the outer electron of the other atom occupying an orbital similar to the LUMO of the H atom, rather than the HOMO of its parent atom. This creates a noticeable distortion of the atomic electron densities in such cases.

The fact that each electron is largely isolated within its local domain provides little opportunity for complicated interelectronic interactions. As a result, we find that the qualitative and quantitative effects of electron correlation are usually small.
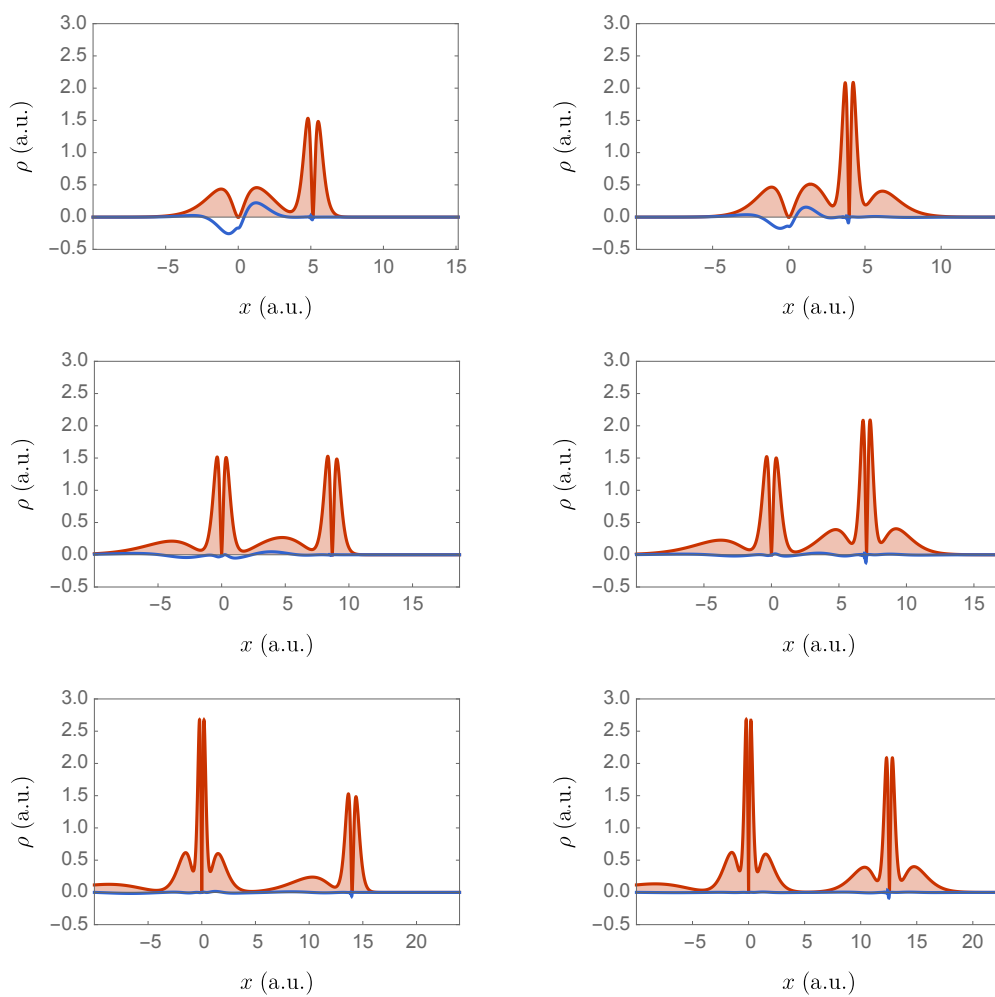
Figure 5.5: Electronic densities (red regions) and difference between molecular and atomic densities (blue lines) of six diatomic molecules. In the left column these are $_1H_2Li_1$, $_2Li_3Li_1$ and $_3B_4Li_1$. In the right column these are $_1H_2Be_2$, $_2Li_3Be_2$ and $_3B_4Be_2$.

The correlation energy in 1D constitutes a much smaller fraction (typically less than 0.1%) of the total energy than in 3D. Moreover, it largely cancels between reactants and products so that correlated bond energies are typically within 1 mE_h of their uncorrelated values. Correlated bond lengths are also similar to uncorrelated ones, especially in relative terms. Accordingly, we use HF structures henceforth.

5.5   Triatomics

We also undertook a systematic search for stable triatomic molecules, examining all possible electronic configurations generated by ground-state atoms up to, and including, carbon. Many stable species emerge, and we report bond lengths, total energies, atomisation energies and bond energies for some of these in Table 5.3. For the reasons discussed above, we report atomisation and bond energies only at the HF level.

In our exploration of 1D reactivity in Chapter 3, we concluded from a small set of atomisation energies that the bonds in a triatomic ABC are similar in strength to those in AB and BC. We argued that the small deviations could be rationalised by considering the A–C dipole interaction.

The results in Table 5.3 largely support this view. For example, the H–Li and Li–Li bond strengths in $H_2Li_3Li_2$ are 39 and 52 mE_h, which are slightly higher than those in H_2Li_2 (36 mE_h) and $_1Li_3Li_2$ (49 mE_h), and this increased stability can be ascribed to the favourable dipole alignment. In contrast, the Li–H and H–B bond strengths in $_2Li_2H_2B_3$ fall from 36 and 100 mE_h to 20 and 84 mE_h, respectively, because the boron dipole is opposed to those of the lithium and hydrogen atoms. The $_3B_5C_5B_3$ molecule also has opposed dipoles and the B–C bond energy drops from 9 mE_h in the diatomic to 7 mE_h in the triatomic.

We have also found two classes of triatomic that one might have expected to be unstable. The first consists of a noble flanked by two alkalis with aligned dipoles (e.g. $H_2He_3B_3$ and $_2Li_2He_4B_2$) and the second consists of two nobles on one side of an alkali (e.g. $_1He_2He_2Li_2$ and $_3B_3He_4C_3$). Such molecules contain bonded atoms that do not form stable diatomics, e.g. the H_2He_1 moiety in $H_2He_3B_3$ or the $_1He_4C_3$ moiety in $_3B_3He_4C_3$.

The exclusion potentials in Fig. 5.6 show the attractive force which binds these unusual ABC triatomics. In each case, the diatomic fragment BC generates a small positive potential in its left domain which can then interact favourably with the valence electron of A.

Molecule	Bond length		Total energy					
ABC	R_{AB}	R_{BC}	$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	E_{atom}	E_{AB}	E_{BC}
$_1$H$_2$H$_2$B$_3$	17.620	3.296	26.458006	26.468322	26.469874	100.497	0.404	—[a]
$_1$H$_1$H$_3$B$_2$	2.795	8.942	26.735055	26.74627	26.74818	310.500	215.008	89.366
H$_2$He$_3$B$_3$	13.090	9.665	29.101690	29.11340	29.11529	1.260	0.799	—[a]
$_1$H$_1$He$_4$C$_3$	2.025	16.294	41.972376	41.9891	41.9917	139.071	—[a]	1.680
H$_2$Li$_3$Li$_2$	5.336	8.860	16.604027	16.611622	16.61313	88.514	39.380	52.107
$_1$H$_2$Li$_3$Be$_2$	5.243	7.048	24.126866	24.13849	24.14078	203.198	174.387	29.172
$_1$H$_2$Li$_4$B$_2$	5.387	13.356	34.094432	34.1091	34.11168	229.167	175.556	55.142
$_1$H$_2$Be$_3$Li$_2$	3.946	7.047	24.111979	24.12320	24.12529	188.310	159.500	24.674
H$_3$B$_4$Li$_2$	10.289	13.443	33.925825	33.9393	33.9414	60.560	6.949	54.178
$_1$H$_3$B$_4$Li$_1$	9.099	14.048	34.040693	34.0553	34.0579	108.382	82.704	12.890
H$_3$B$_5$B$_3$	10.281	19.510	51.24102	51.2605	51.2631	26.006	7.115	19.62
$_1$H$_3$B$_5$B$_2$	9.205	19.492	51.39508	51.4159	51.4192	113.018	74.956	17.53
$_1$H$_3$C$_3$H$_1$	6.649	6.649	39.422734	39.4393	39.44217	332.352	166.062	166.062
$_1$H$_3$C$_5$B$_3$	6.632	16.058	64.12276	64.1480	64.1517	174.867	165.657	8.577
$_1$He$_1$H$_3$B$_3$	2.027	10.290	29.244835	29.25684	29.25879	144.404	138.023	7.014
$_1$He$_2$He$_2$Li$_2$	11.009	4.601	14.503682	14.511311	14.513255	10.081	0.104	—[a]
$_1$He$_2$Li$_3$Li$_2$	4.601	8.737	19.317942	19.327327	19.329411	59.507	10.373	49.530
$_1$He$_3$B$_4$Li$_2$	10.949	13.320	36.662290	36.6778	36.6805	54.103	0.492	53.642
$_1$Li$_2$H$_2$Li$_2$	5.210	5.524	16.749370	16.75602	16.759498	233.857	197.450	59.832
$_1$Li$_2$H$_3$B$_3$	5.191	10.017	34.055957	34.07025	34.07275	190.692	184.310	16.667
$_2$Li$_2$H$_2$B$_3$	5.619	3.338	33.985753	33.999903	34.002215	120.488	20.395	84.081
$_2$Li$_2$H$_3$C$_3$	5.533	6.789	46.820115	46.83917	46.84223	221.976	55.686	185.569
$_2$Li$_2$He$_3$Be$_2$	4.574	12.299	26.677407	26.68967	26.69200	10.817	—[a]	0.840
$_2$Li$_2$He$_4$B$_2$	4.514	20.723	36.623610	36.6388	36.6414	15.423	—[a]	5.45
$_1$Li$_3$Li$_3$Li$_2$	8.69	8.966	24.127495	24.13851	24.140705	104.226	55.092	55.092
$_2$Li$_3$Li$_3$Be$_2$	8.833	7.04	31.51126	31.52547	31.52797	79.832	51.021	30.698
$_1$Li$_3$Li$_4$B$_3$	8.715	14.190	41.440738	41.457727	41.46047	67.717	53.951	18.583
$_2$Li$_3$Li$_4$C$_3$	8.948	10.461	54.201813	54.2238	54.2272	95.918	53.432	46.784
$_2$Li$_3$Be$_3$Li$_2$	7.077	7.075	31.482743	31.49682	31.49927	51.318	22.507	22.507
$_1$Li$_4$B$_5$B$_3$	13.937	19.704	58.75720	58.7802	58.7835	34.426	15.54	20.66
$_2$Be$_2$H$_3$B$_3$	3.983	10.32	41.445882	41.46312	41.46577	172.461	166.08	8.825
$_2$Be$_4$B$_5$B$_3$	12.580	19.444	66.15358	66.1798	66.1834	22.652	3.761	19.19
$_2$B$_4$H$_2$B$_3$	25.12	3.288	51.3170	51.3367	51.3397	10.198	1.9	—[a]
$_3$B$_3$H$_3$C$_3$	10.164	6.747	64.126951	64.15210	64.15578	179.060	12.770	172.678
$_3$B$_3$He$_4$C$_3$	9.794	18.02	66.69221	66.7183	66.7221	1.39	—[a]	0.93
$_2$B$_5$Be$_4$B$_3$	30.000	12.358	66.1349	66.1608	66.1644	4.00	0.5	—[a]
$_3$B$_5$C$_5$B$_3$	16.126	16.129	88.82132	88.8553	88.8597	15.916	6.71	6.71

[a]After breaking this bond, the remaining diatomic is unstable.

Table 5.3: Equilibrium bond lengths (bohr), total energies (E_h), HF atomisation energies (E_{atom}, mE_h) and HF bond dissociation energies (E_{AB} and E_{BC}, mE_h) of triatomic molecules.
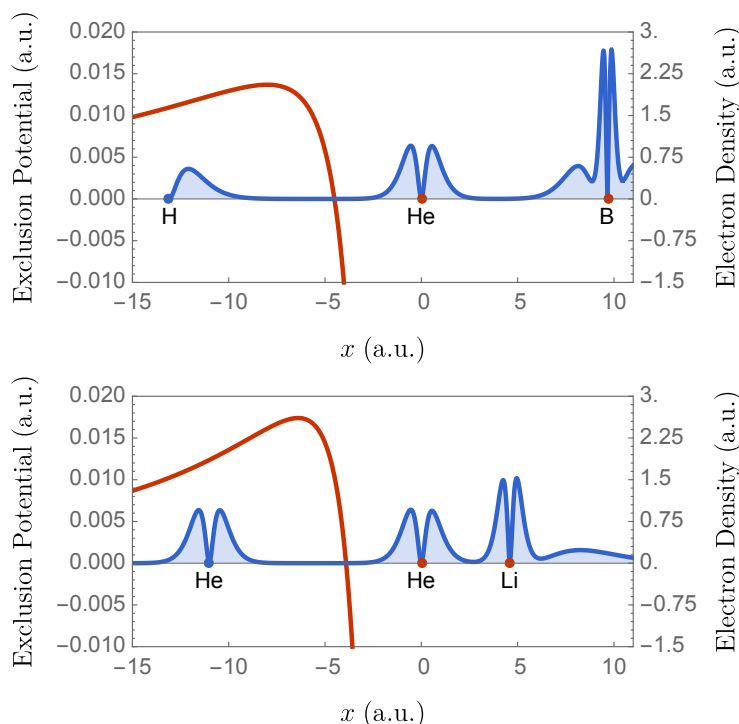
Figure 5.6: Left exclusion potential (red) of $_1\text{He}_3\text{B}_3$ (top) and $_1\text{He}_2\text{Li}_2$ (bottom) and electron density (blue) of $\text{H}_2\text{He}_3\text{B}_3$ (top) and $_1\text{He}_2\text{He}_2\text{Li}_2$ (bottom).

5.6   Tetra-atomics

Table 5.4 presents results for the stable tetra-atomic molecules formed by dimerising $_1\text{H}_3\text{B}_2$, $_1\text{H}_2\text{B}_3$ and H_3B_3. These results confirm that the length and strength of a bond are largely independent of its environment but, as in the triatomic study, we find significant increases in some bond strengths as a result of favorable dipole interactions. For example, in $_2\text{B}_3\text{H}_1\text{H}_3\text{B}_3$, the central H−H bond is approximately 50 mE_h stronger than that in H_1H_1, and the right bond is almost four times as strong as in H_3B_3. However, this effect is not universal. For example, the dipoles in $_1\text{H}_3\text{B}_5\text{B}_3\text{H}$ are all aligned, yet the individual bonds are not strengthened and, indeed, the central bond is weaker than in the corresponding diatomic. This is because the central domain houses five electrons, forcing the bond to be long (> 18 bohrs) and greatly reducing the dipole stabilisation.

In most of the stable species ABCD, the central pair BC is also a stable diatomic. However, this is not the case in $_3\text{B}_2\text{H}_4\text{B}_3\text{H}$ and $_2\text{B}_3\text{H}_2\text{H}_2\text{B}_3$. In both of these, the central bond is significantly weaker than in the other tetra-atomics and they are therefore analogous to the loosely associated triatomics in Table 5.3.

We also found two molecules, $_3\text{B}_2\text{H}_3\text{B}_3\text{H}_1$ and $_1\text{H}_3\text{B}_2\text{H}_3\text{B}_3$, where each individual bond is present in a stable diatomic but the overall tetra-atomic is not bound. The

Molecule	Bond length			Total energy						
ABCD	AB	BC	CD	$-E_{HF}$	$-E_{MP2}$	$-E_{MP3}$	E_{atom}	E_{AB}	E_{BC}	E_{CD}
$_3B_2H_4B_3H$	3.284	23.316	10.349	51.822463	51.842	51.845	107.48	—[a]	1.0	5.5
$_2B_3H_2H_2B_3$	8.880	18.08	3.290	51.980101	52.0013	52.0048	198.034	90.492	2.449	—[a]
$_2B_3H_1H_3B_3$	9.085	2.795	9.828	52.115557	52.1361	52.1392	333.182	75.51	231.308	22.682
$_3B_2H_1H_3B_3$	3.356	2.715	10.35	51.99337	52.0142	52.0173	278.35	82.682	171.879	2.84
$_1H_3B_5B_2H_1$	8.979	19.70	3.297	51.993816	52.015	52.019	211.6	77.3	16.1	98.6
$_1H_3B_5B_3H$	9.253	18.48	10.340	51.901493	51.923	51.926	119.429	73.090	17.56	6.41
$_1H_3B_3H_3B_2$	8.880	9.99	8.88	52.06129	52.083	52.087	212.18	74.54	21.20	81.1
$_1H_2B_3H_3B_3$	3.298	8.854	10.155	51.999949	52.0210	52.0244	217.728	99.060	88.595	21.238
$H_3B_3H_3B_3$	10.348	9.128	9.825	51.907673	51.9280	51.9311	125.255	6.587	75.332	23.818

[a]After breaking this bond, the remaining molecule is unstable.

Table 5.4: HF equilibrium bond lengths (bohr), total energies (E_h), HF atomisation energies (E_{atom}, mE_h) and HF bond dissociation energies (E_{AB}, E_{BC} and E_{CD}, mE_h) of tetra-atomic molecules.

exclusion potential of the fragment $_3B_2H_1$, which is present in both $_3B_2H_4B_3H$ and $_3B_2H_3B_3H_1$, becomes positive beyond 8 bohr to the right of the H atom and this provides the driving force for bonding in the molecule $_3B_2H_4B_3H$. However, $_3B_2H_3B_3H_1$ remains unbound because of the unfavourable dipole interactions between the left boron atom and the other atoms. Similar dipole interactions also appear in $_1H_3B_2H_3B_3$. It is possible that the $_3B_2H_1$ and $_2B_3H_1$ fragments in $_3B_2H_3B_3H_1$, or the $_1H_3B_2H_1$ and $_2B_3$ fragments in $_1H_3B_2H_3B_3$ associate weakly at a very large separation but that the (30,50) basis cannot adequately describe this.

5.7   Polymers

In Chapter 2, we examined the bond length and energy within the hydrogen nanowire—an infinite chain of alternating protons and electrons—using a periodic HF calculation. Using LegLag, we can study the same system as the extrapolated limit of a sequence of finite chains and we can also examine other homogeneous, or heterogenous, polymers.

For each polymer, we studied a range of short oligomers and fit their properties to the functions of the type $\sum_{k=0}^{2} a_k n^{-k}$, where n is the number of monomer units in the oligomer. We then extrapolated these functions to the infinite polymer, i.e. $n \to \infty$. For computational efficiency, we used the (30,30) basis set, rather than the (30,50) set used above. Our results are summarised in Table 5.5 and we report only the digits that have converged as the basis set is increased to the (30,30) set.

AB	Monomer		Polymer			
	R_{AB}	$-E_{HF}$	R_{AB}	R_{BA}	$-E_{HF}$	E_{stab}
H_1H_1	2.636	1.184572	2.798	2.798	1.420210	235.638
$_1Li_3Li_2$	8.693	16.064647	9.0	9.0	16.13383	69.18
$_1H_2Li_1$	5.152	8.681782	5.46	5.46	8.752991	71.209
$_1H_3B_2$	8.880	26.020047	9.6	9.6	26.0474	27.4
$_2Li_4B_2$	13.330	33.418876	14.0	14.0	33.447	28
$_1H_1He_1$	2.025	3.880313	2.025	10	3.882261	1.948
$_1H_2B_3$	3.298	25.957601	3.3	20	26.0	0

Table 5.5: Equilibrium bond lengths (R_{AB} and R_{BA}, bohr), energies per monomer (E_{HF}, E_h) and stabilisation energy per monomer (E_{stab}, mE_h) of 1D polymers at the HF level of theory.

Results for the hydrogen polymer agree with the earlier periodic calculations and confirm that the H−H bond becomes longer (stretching from around 2.6 bohrs to 2.8 bohrs) and stronger upon polymerisation. The lengthening / strengthening trend is ubiquitous and results from a competition between growing numbers of repulsive interelectronic interactions (which are reduced if the polymers expand by a few percent) and an accumulation of favourable dipole interactions (which stabilise the polymer relative to the monomers).

However, not all of the polymers in Table 5.5 follow this pattern and the $(_1H_1He_1)_n$ and $(_1H_2B_3)_n$ polymers are notable exceptions. In both of these, the inter-monomer bonds are exceptionally long and the resulting stabilisation is small. Because these new bonds do not arise in stable diatomics, these "polymers" are better viewed as loose aggregates.

5.8   Rules of 1D bonding

Our studies reveal that chemistry in 1D is largely local. The combination of particle impenetrability and strong shielding causes distant particles to have very little effect on one another and, as a result, the functional groups in 1D chemistry are essentially the diatomic units within a molecule. This reduction requires us to understand the bonding in diatomic molecules and has led us to three simple rules which describe all of the bound diatomics reported in Table 5.2:

- Two alkalis with aligned dipoles bind

- Two alkalis with unaligned dipoles bind if their nuclear charges differ by at least two

- A noble binds to an alkali's electron-deficient side

Strong bonds result from three ingredients:

- Light atoms

- Aligned atomic dipoles

- Low electron occupations in each domain

The first ingredient improves electron-nuclear attraction (because of reduced shielding); the last also enhances Coulombic attraction and also reduces kinetic energy.

In general, a polyatomic is strongly bound if all of its constituent diatomics are separately stable. There are interesting exceptions (such as the stable triatomic $H_2H_2B_3$ and the unstable tetra-atomic $_3B_2H_3B_3H_1$) but it is true for all of the tightly bound triatomics that we have identified. Curiously, the rule incorrectly predicts $_3B_2H_3B_3H_1$ and $_1H_3B_2H_3B_3$ to be strongly bound when, in fact, it turns out that one of their constituent bonds is insufficiently strong to overcome the unfavourable dipole interactions. Fortunately, this is the only example that we have found where the rule fails.

5.9 Conclusion

Using LegLag we have performed an extensive survey of 1D chemistry. By adopting improved basis functions, we have been able to identify and characterise a wide variety of stable molecules and a small set of polymers. Many of these are novel structures and, prior to this work, would not have been expected to exist.

We have also developed an understanding of the bonding interactions in these molecules and we have identified the most significant factors that contribute to their stability. This has allowed us to formulate a set of simple rules which predict whether a putative 1D molecule is stable.

Chapter 6

The generalised local density approximation

6.1 Local Density Approximation

The Local Density Approximation (LDA), unlike most of the "sophisticated" density functional approximations in widespread use today, is truly a first-principles quantum mechanical method [50]. It is entirely non-empirical, depending instead on the properties of one of the great paradigms of modern physics: the infinite homogeneous electron gas (HEG) [167, 168]. Application of the LDA is straightforward, at least in principle. Although the electronic charge density $\rho(\mathbf{r})$ in any real system is non-uniform, the LDA proceeds by assuming that the charge in an infinitesimal volume element around the point $\mathbf{r}$ behaves like a locally homogeneous gas of density $\rho(\mathbf{r})$, and adds all of the resulting contributions together. This implicitly assumes that the infinitesimal contributions are independent (which is undoubtedly not the case) but then requires only that the properties of the HEG be known for all values of ρ.

The density of a HEG is commonly given by ρ (the number of electrons per unit volume) or the Seitz radius r_s and these equivalent parameters are related by

$$r_s^{\mathcal{D}} \rho = \pi^{-\mathcal{D}/2} \Gamma(\mathcal{D}/2 + 1) \tag{6.1}$$

where $\mathcal{D}$ is the dimensionality of the space in which the electrons move. In terms of these, the LDA correlation functional is

$$E_c^{\text{LDA}} = \int \rho(\mathbf{r}) \varepsilon_c(r_s(\mathbf{r})) \, d\mathbf{r} \tag{6.2}$$

where the correlation kernel $\varepsilon_c(r_s)$ is the reduced (*i.e.* per electron) correlation energy of the HEG with Seitz radius r_s.

In high-density HEGs (*i.e.* $r_s \ll 1$), the kinetic energy dominates the Hamiltonian and the Coulomb repulsion between the electrons can be treated via perturbation theory. This has facilitated investigations of $\varepsilon_c(r_s)$ in 3D [169–184] and 2D [185–194] but, because the Coulomb operator is so strong in 1D that two electrons cannot touch, the 1D gas has received less attention [62, 65, 195].

In low-density HEGs (*i.e.* $r_s \gg 1$), the potential energy dominates, the electrons localize into a Wigner crystal and strong-coupling methods can be used to find asymptotic expansions of $\varepsilon_c(r_s)$. Here, too, the 3D [196–198], 2D [199, 200], and 1D [195] HEGs have all been studied.

For intermediate densities, the best estimates of $\varepsilon_c(r_s)$ come from Quantum Monte Carlo (QMC) calculations, as pioneered by Ceperley and refined by several other groups [20, 95, 201–214]. By combining these with the high- and low-density results, various groups [24, 215–217] have constructed interpolating functions that allow $\varepsilon_c(r_s)$ to be estimated rapidly for any value of r_s.

Unfortunately this approach is flawed, for the correlation energy of a uniform electron gas depends on more than just its r_s value [218]. In this chapter we argue that $\varepsilon_c(r_s)$ should be generalised to $\varepsilon_c(r_s, \eta)$, where the parameter η measures the *two-electron* density. Although not mathematically mandated [219], we prefer that η, like r_s, be a *local* quantity. In Sec. 6.2, we propose a definition for η inspired by a number of previous researchers [220–224].

To learn more about the two-parameter kernel, we have turned to the finite uniform electron gases (UEGs) formed when n electrons are confined to a $\mathcal{D}$-sphere [64, 66, 106, 107, 225–228]. In Sec. 6.3, we report accurate values of η and $\varepsilon_c(r_s, \eta)$ for electrons on a 1-sphere, systems that we call "n-ringium". In Sec. 6.4, we devise three functionals to approximate these results and in Sec. 6.5, we test two of these on small 1D systems.

6.2   Hole curvature

Suppose that an electron lies at a point $\mathbf{r}$. The probability $P(\mathbf{u}|\mathbf{r})$ that a second electron lies at $\mathbf{r} + \mathbf{u}$ is given [220, 221, 229–238] by the conditional intracule

$$P(\mathbf{u}|\mathbf{r}) = \rho_2(\mathbf{r}, \mathbf{r} + \mathbf{u})/\rho(\mathbf{r}) \ = [\rho(\mathbf{r} + \mathbf{u}) + \rho_{\mathrm{xc}}(\mathbf{r}, \mathbf{r} + \mathbf{u})]/2 \qquad (6.3)$$

where ρ_{xc} is the exchange-correlation hole [50] and

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = n(n-1) \int |\Psi|^2 \, ds_1 ds_2 d\mathbf{r}_3 \dots d\mathbf{r}_n \tag{6.4}$$

is the spinless second-order density matrix [239]. For fixed $\mathbf{r}$, we have the normalization

$$\int P(\mathbf{u}|\mathbf{r}) \, d\mathbf{u} = n - 1 \tag{6.5}$$

Because the Laplacian $\nabla_{\mathbf{u}}^2 P(\mathbf{0}|\mathbf{r})$ measures the tightness of the hole around the electron at $\mathbf{r}$ and has dimensions of $1/(\text{Length})^{\mathcal{D}+2}$, we can use the dimensionless hole curvature

$$\eta(\mathbf{r}) = C_{\mathcal{D}} \, r_s(\mathbf{r})^{\mathcal{D}+2} \, \nabla_{\mathbf{u}}^2 P(\mathbf{0}|\mathbf{r}) \tag{6.6}$$

to measure the proximity of other electrons to one at $\mathbf{r}$. (We will fix the coefficient $C_{\mathcal{D}}$ in the next Section.) It is difficult to find this Laplacian for the exact wave function but, at the Hartree-Fock (HF) level, it involves simple sums over the occupied orbitals, *viz.*

$$\nabla_{\mathbf{u}}^2 P(\mathbf{0}|\mathbf{r}) = 2 \sum_i^{\text{occ}} |\nabla\psi_i|^2 - \frac{|\nabla\rho|^2}{2\rho} = 2\tau - \tau_w \tag{6.7}$$

where $\tau = \sum_i^{\text{occ}} |\nabla\psi_i|^2$ is the HF kinetic energy density and $\tau_w = \frac{|\nabla\rho|^2}{2\rho}$ is the Weizsäcker kinetic energy density. We will therefore employ HF curvatures henceforth.[1]

This expression for the curvature can be rewritten in terms of the "kinetic energy density" [222]. These terms are essential in removing the self-interaction error, and appear here in the appropriate configuration for achieving this task [240, 241].

6.3 Calculations on n-Ringium

6.3.1 Density and curvature

The HF orbitals of the ground state of n electrons on a ring of radius R are complex exponentials [64, 66]

$$\psi_m(\theta) = (2\pi R)^{-1/2} \exp(im\theta) \tag{6.8}$$

$$m = -\frac{n-1}{2}, -\frac{n-3}{2}, \dots, +\frac{n-3}{2}, +\frac{n-1}{2} \tag{6.9}$$

---

[1]We have calculated the exact and HF curvatures in 2-ringium for several values of the ring radius R. The HF curvature is $\eta_{\text{HF}} = 3/4$ for all R. In contrast, the exact curvature diminishes from $\eta = 3/4$ at $R = 0$, to $\eta = 9\pi/(32 + 12\pi) \approx 0.406$ at $R = 1/2$, to $\eta = 18\pi/(64\sqrt{6} + 51\pi) \approx 0.178$ at $R = \sqrt{3/2}$, and finally to $\eta = 0$ at $R = \infty$.

and, because of the symmetry of the system, the density and Seitz radius

$$\rho = n/(2\pi R) \tag{6.10}$$

$$r_s = \pi R/n \tag{6.11}$$

do not depend on θ. The hole curvature is also constant and, using (6.6) and (6.7), one finds

$$\eta = 2C_1(\pi R/n)^3 \sum_m^{\text{occ}} m^2/(2\pi R^3)$$

$$= C_1(\pi^2/12)(1 - 1/n^2) \tag{6.12}$$

If we choose $C_1 = 12/\pi^2$ so that $\eta = 1$ for the 1D HEG (*i.e.* ∞-ringium), we obtain

$$\eta = 1 - 1/n^2 \tag{6.13}$$

In general, requiring that $\eta = 1$ in the $\mathcal{D}$-dimensional HEG leads (via Fermi integration) to

$$C_{\mathcal{D}} = \frac{(1 + 2/\mathcal{D})\pi^{\mathcal{D}/2}/8}{\Gamma(1 + \mathcal{D}/2)^{1+4/\mathcal{D}}} \tag{6.14}$$

and the particular values $C_2 = \pi/4$ and $C_3 = (10/27)(4\pi/3)^{1/3}$.

6.3.2   Correlation energy

The Hamiltonian for n electrons on a ring is

$$\hat{H} = \hat{T} + \hat{V} = -\frac{1}{2}\sum_{i=1}^{n} \nabla_i^2 + \sum_{i<j}^{n} r_{ij}^{-1} \tag{6.15}$$

where r_{ij} is the distance (across the ring) between electrons i and j. As noted previously by Loos and Gill [66], the energy is independent of the spin-state and so we assume that all electrons are spin-up. The exact wave function can then be written as $\Psi = F\Phi$, where the correlation factor

$$F = \sum_{a=1}^{\infty} x_a f_a \tag{6.16}$$

		Degree 0		Degree 1		Degree 2		Degree 3	
	f_1	m_1	f_2	m_2	$f_3 \ldots f_5$	$m_3 \ldots m_5$	$f_6 \ldots f_{13}$	$m_6 \ldots m_{13}$	
	1	1	$\sum r_{ij}$	$_n\mathrm{C}_2$	$\sum r_{ij}^2$	$_n\mathrm{C}_2$	$\sum r_{ij}^3$	$_n\mathrm{C}_2$	
					$\sum r_{ij}r_{ik}$	$6\,_n\mathrm{C}_3$	$\sum r_{ij}r_{ik}r_{jk}$	$6\,_n\mathrm{C}_3$	
					$\sum r_{ij}r_{kl}$	$6\,_n\mathrm{C}_4$	$\sum r_{ij}^2 r_{ik}$	$18\,_n\mathrm{C}_3$	
							$\sum r_{ij}^2 r_{kl}$	$18\,_n\mathrm{C}_4$	
							$\sum r_{ij}r_{ik}r_{il}$	$24\,_n\mathrm{C}_4$	
							$\sum r_{ij}r_{ik}r_{jl}$	$72\,_n\mathrm{C}_4$	
							$\sum r_{ij}r_{ik}r_{lm}$	$180\,_n\mathrm{C}_5$	
							$\sum r_{ij}r_{kl}r_{mn}$	$90\,_n\mathrm{C}_6$	
Total	1	1	$\sum r_{ij}$	$_n\mathrm{C}_2$	$(\sum r_{ij})^2$	$(_n\mathrm{C}_2)^2$	$(\sum r_{ij})^3$	$(_n\mathrm{C}_2)^3$	

Table 6.1: Definitions f_a and number of terms m_a in the correlation factors of degree 0, 1, 2, 3

is a sum of functions f_a which are m_a-term symmetric polynomials in the r_{ij} (see Table 6.1) and Φ is the HF wave function [66]

$$\Phi = \frac{1}{\sqrt{n!(2\pi)^n}} \prod_{i<j}^{n} \hat{r}_{ij} \tag{6.17}$$

where $\hat{r}_{ij} = 2R\sin\frac{\theta_i - \theta_j}{2}$ is the signed distance.

Judicious integration by parts allows us to partition the total energy

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{6.18}$$

into the HF energy [66]

$$E_{\mathrm{HF}} = T_{\mathrm{HF}} + V_{\mathrm{HF}} = \frac{n(n^2-1)}{24R^2} + \frac{1}{4\pi R}\left(\sum_{k=1}^{n} \frac{4n^2-1}{2k-1} - 3n^2\right) \tag{6.19}$$

and the correlation energy

$$E_c = \frac{\langle \Phi \mid \frac{1}{2}\nabla F \cdot \nabla F + (\hat{V} - V_{\mathrm{HF}})F^2 \mid \Phi \rangle}{\langle \Phi | F^2 | \Phi \rangle} \tag{6.20}$$

E_c can be minimised either by QMC methods [213] or via the secular equation

$$(\mathbf{T} + \mathbf{V})\mathbf{x} = E_c\,\mathbf{S}\,\mathbf{x} \tag{6.21}$$

		r_s										
n	η	0	1/10	1/5	1/2	1	2	5	10	20	50	100
1	0	0	0	0	0	0	0	0	0	0	0	0
2	3/4	13.212	12.985	12.766	12.152	11.250	9.802	7.111	4.938	3.122	1.533	0.848
3	8/9	18.484	18.107	17.747	16.755	15.346	13.179	9.369	6.427	4.030	1.965	1.083
4	15/16	21.174	20.700	20.250	19.027	17.324	14.765	10.391	7.087	4.425	2.150	1.184
5	24/25	22.756	22.216	21.706	20.332	18.444	15.648	10.947	7.441	4.636	2.249	1.237
6	35/36	23.775	23.190	22.638	21.161	19.148	16.196	11.285	7.655	4.774	2.307	1.268
7	48/49	24.476	23.855	23.273	21.723	19.618	16.557	11.509	7.795	4.844	2.345	1.289
8	63/64	24.981	24.328	23.729	22.122	19.951	16.813	11.664	7.890	4.901	2.370	1.302
9	80/81	25.360	24.686	24.067	22.415	20.199	17.001	11.777	7.960	4.941	2.389	1.312
10	99/100	25.651	24.960	24.327	22.644	20.386	17.143	11.857	8.013	4.973	2.404	1.320
∞	1	27.416	26.597	25.91	23.962	21.444	17.922	12.318	8.292	5.133	2.476	1.358

Table 6.2: η and $-\varepsilon_c(r_s, \eta)$ (mE_h per electron) for the ground state of n electrons on a ring

where the overlap, kinetic and Coulomb matrix elements

$$S_{ab} = \langle \Phi | f_a f_b | \Phi \rangle \tag{6.22a}$$

$$T_{ab} = \frac{1}{2} \langle \Phi | \nabla f_a \cdot \nabla f_b | \Phi \rangle \tag{6.22b}$$

$$V_{ab} = \langle \Phi | f_a \hat{V} f_b | \Phi \rangle - V_\text{HF} S_{ab} \tag{6.22c}$$

can be found analytically in Fourier space (See Appendix 6.A). We have used the CASINO QMC package [242] and, where possible, the Knowles–Handy Full CI program to confirm results [243, 244].

Table 6.2 shows the resulting near-exact correlation energies for ground-state n-ringium. The fact that the ε_c values in a given column are not equal demonstrates that the correlation energy of a UEG is not determined by its r_s value alone [218]. Moreover, the variations in ε_c for a given r_s are large: the $n = 2$ values, for example, are only about half of the $n = \infty$ values, implying that the correlation energy of a few-electron system is grossly overestimated by the LDA functional which is based on the HEG.

6.4   Generalised Local Density Approximation

In the LDA, the correlation contribution is estimated from r_s alone, according to Eq. (6.2). However, the fact that UEGs with the same r_s, but different η, have different energies compels us to devise a Generalised Local Density Approximation

(GLDA) wherein we write

$$E_c^{\text{GLDA}} = \int \rho(\mathbf{r}) \, \varepsilon_c(r_s(\mathbf{r}), \eta(\mathbf{r})) \, d\mathbf{r} \tag{6.23}$$

where the correlation kernel $\varepsilon_c(r_s, \eta)$ is the reduced correlation energy of a UEG with Seitz radius r_s and curvature η. For present purposes, we will use r_s and η values from the HF, rather than the exact, wave function.

One might think that the kernel could be constructed by fitting the results in Table 6.2 but these data allow us to construct $\varepsilon_c(r_s, \eta)$ only for $\eta \leq 1$. To construct the rest of the kernel will require accurate correlation energies for uniform gases with high curvatures ($\eta > 1$). For now we will concern ourselves with the construction of GLDA in the (ρ, η) region that the current data covers, we will discuss extensions beyond this region in Chapter 8 (Sec. 8.2).

6.4.1 High densities

Rayleigh-Schrödinger perturbation theory for n-ringium yields the high-density expansion

$$\varepsilon_c(r_s, n) = \alpha_2(n) + \alpha_3(n)r_s + \alpha_4(n)r_s^2 + \ldots, \qquad (r_s \ll 1) \tag{6.24}$$

The leading coefficient [66] is

$$\begin{aligned}
\alpha_2(n) &= -\frac{1}{n} \sum_{a<b}^{\text{occ}} \sum_{r=r_{\min}}^{\infty} \frac{V_{r-a,r-b}^2}{(r-a)(r-b)} \\
&= -\frac{\pi^2}{360} + \frac{a \ln^2 n + b \ln n + c}{n^2} + \ldots
\end{aligned} \tag{6.25}$$

but, if we fit a truncated version of this series, while ensuring that α_2 vanishes for one electron, we obtain the approximation

$$\tilde{\alpha}_2(n) = -\frac{\pi^2}{360} \left(1 - \frac{1}{n^2} \right) + \frac{\ln^2 n + 3 \ln n}{87 n^2} \qquad (1 \leq n < \infty) \tag{6.26}$$

which can be rewritten in terms of the curvature, using Eq. (6.13) to obtain

$$\tilde{\alpha}_2(\eta) = -\frac{\pi^2}{360}\eta + (1 - \eta)\frac{\ln^2(1 - \eta) - 6 \ln(1 - \eta)}{348} \qquad (0 \leq \eta \leq 1) \tag{6.27}$$

The accuracy of this approximation is shown in columns 2 and 3 of Table 6.3.

6.4.2   Low densities

Strong-coupling perturbation theory for n-ringium yields the low-density expansion

$$\varepsilon_c(r_s, n) = \frac{\beta_2(n)}{r_s} + \frac{\beta_3(n)}{r_s^{3/2}} + \frac{\beta_4(n)}{r_s^2} + \dots \qquad (r_s \gg 1) \qquad (6.28)$$

The leading coefficient is the difference between the Wigner crystal Coulomb coefficient

$$
\begin{aligned}
E_{\mathrm{V}}^{\mathrm{W}}(n) &= \frac{\pi}{4n} \sum_{k=1}^{n-1} \csc(k\pi/n) \\
&= \frac{1}{2} \int_0^1 \frac{1 - x^{n-1}}{1 - x} \frac{dx}{1 + x^n} \\
&= \frac{\ln n}{2} + \frac{\gamma + \ln(2/\pi)}{2} - \frac{\pi^2}{144n^2} + \dots
\end{aligned}
\qquad (6.29)
$$

and the HF Coulomb coefficient

$$
\begin{aligned}
E_{\mathrm{V}}^{\mathrm{HF}}(n) &= \left(1 - \frac{1}{4n^2}\right) \sum_{k=1}^{n} \frac{1}{2k - 1} - \frac{3}{4} \\
&= \left(1 - \frac{1}{4n^2}\right) \left(\frac{\ln n}{2} + \frac{\gamma + 2\ln 2}{2} + \frac{1}{48n^2} + \dots\right) - \frac{3}{4}
\end{aligned}
\qquad (6.30)
$$

It follows that

$$\beta_2(n) = \frac{3}{4} - \frac{\ln 2\pi}{2} + \frac{\ln n}{8n^2} + \frac{18\gamma + 36\ln 2 - 3 - \pi^2}{144n^2} + \dots \qquad (6.31)$$

but, if we truncate this series after the n^{-2} term and modify it to ensure that β_2 vanishes for one electron, we obtain the approximation

$$\tilde{\beta}_2(n) = \left(\frac{3}{4} - \frac{\ln 2\pi}{2}\right)\left(1 - \frac{1}{n^2}\right) + \frac{\ln n}{8n^2} \qquad (1 \le n < \infty) \qquad (6.32)$$

which can be rewritten in terms of the curvature, using Eq. (6.13) to obtain

$$\tilde{\beta}_2(\eta) = \left(\frac{3}{4} - \frac{\ln 2\pi}{2}\right)\eta - \frac{(1 - \eta)\ln(1 - \eta)}{16} \qquad (0 \le \eta \le 1) \qquad (6.33)$$

The accuracy of this approximation is shown in columns 4 and 5 of Table 6.3.

6.4.3   Intermediate densities

How can we model $\varepsilon_c(r_s, \eta)$ for fixed η? Ideally, we would like a function that reproduces the behaviors of Eqs (6.24) and (6.28) and interpolates accurately between these limits. However, for practical reasons, we will content ourselves with a function

						Max errors	
n	η	$-\alpha_2$	$-\tilde{\alpha}_2$	$-\beta_2$	$-\tilde{\beta}_2$	$\tilde{\gamma}$	%	Abs
2	3/4	0.01321	0.01321	0.1073	0.1050	1.9792	1.0	0.10
3	8/9	0.01848	0.01862	0.1361	0.1349	2.1375	0.9	0.13
4	15/16	0.02117	0.02133	0.1483	0.1475	2.2054	0.8	0.16
5	24/25	0.02276	0.02291	0.1546	0.1541	2.2431	0.8	0.18
6	35/36	0.02378	0.02391	0.1584	0.1580	2.2670	0.7	0.14
7	48/49	0.02448	0.02460	0.1608	0.1605	2.2837	0.8	0.16
8	63/64	0.02498	0.02509	0.1624	0.1622	2.2958	0.7	0.16
9	80/81	0.02536	0.02546	0.1636	0.1635	2.3051	0.7	0.15
10	99/100	0.02565	0.02574	0.1645	0.1644	2.3125	0.8	0.20
∞	1	0.02742	0.02742	0.1689	0.1689	2.3750	0.8	0.13

Table 6.3: Application of the $\tilde{\varepsilon}_c(r_s, \eta)$ approximation to the data in Table 6.2. Absolute errors in mE_{h}.

that approaches $\tilde{\alpha}_2(\eta)$ for small r_s, behaves like $\tilde{\beta}_2(\eta)/r_s$ for large r_s, and changes monotonically between these.

Although we could use robust interpolation [245], the hypergeometric function [100]

$$f(r) = \alpha F\left(1, \frac{3}{2}, \gamma, \frac{2\alpha(1-\gamma)}{\beta} r\right) \tag{6.34}$$

$$\sim \begin{cases} \alpha + O(r) & r \ll 1 \\ \beta/r + O(r^{-3/2}) & r \gg 1 \end{cases} \tag{6.35}$$

possesses all of the desired features and we therefore adopt the approximate kernel

$$\tilde{\varepsilon}_c(r_s, \eta) = \tilde{\alpha}_2(\eta) F\left(1, \frac{3}{2}, \tilde{\gamma}(\eta), \frac{2\tilde{\alpha}_2(\eta)(1 - \tilde{\gamma}(\eta))}{\tilde{\beta}_2(\eta)} r_s\right) \tag{6.36}$$

Table 6.3 shows that this kernel models the energies in Table 6.2 well if we choose

$$\tilde{\gamma}(n) = \frac{19}{16}\left(\frac{4n-3}{2n-1}\right) \qquad (1 \le n < \infty) \tag{6.37}$$

or, equivalently,

$$\tilde{\gamma}(\eta) = \frac{19}{16}\left(\frac{4 - 3\sqrt{1-\eta}}{2 - \sqrt{1-\eta}}\right) \qquad (0 \le \eta \le 1) \tag{6.38}$$

reproduces the Table 6.2 data to within a relative error of 1% and absolute error of 0.20 mE_{h}.

6.4.4   The LDA1, GLDA1 and gLDA1 functionals

We can now consider three approximate kernels for correlation in 1D systems. The first is the LDA1 kernel, which is defined by

$$\varepsilon_c^{\text{LDA1}}(r_s) = \alpha F\left(1, \frac{3}{2}, \tilde{\gamma}, \frac{2\alpha(1-\tilde{\gamma})}{\beta}r_s\right) \tag{6.39}$$

where $\alpha = -\pi^2/360$, $\beta = 3/4 - (\ln 2\pi)/2$ and $\tilde{\gamma} = 19/8$. This underpins the traditional LDA and, by construction, it is exact (within fitting errors) for the 1D HEG or, equivalently, for ∞-ringium. It is independent of the hole curvature η.

The second is the GLDA1 kernel, which is defined by

$$\varepsilon_c^{\text{GLDA1}}(r_s, \eta) = \tilde{\alpha}_2(\eta)F\left(1, \frac{3}{2}, \tilde{\gamma}(\eta), \frac{2\tilde{\alpha}_2(\eta)(1-\tilde{\gamma}(\eta))}{\tilde{\beta}_2(\eta)}r_s\right) \tag{6.40}$$

where $\tilde{\alpha}_2(\eta)$, $\tilde{\beta}_2(\eta)$ and $\tilde{\gamma}(\eta)$ are defined in Eqs (6.27), (6.33) and (6.38). Unfortunately, because of a lack of information about high-curvature UEGs, these three equations are not defined for $\eta > 1$ and thus, at this time, the GLDA1 is defined only for systems where $\eta \leq 1$ at all points. Completing the definition of the GLDA1 is an important topic for future work.

The third is the gLDA1 kernel, a partially corrected LDA, which is defined by

$$\varepsilon_c^{\text{gLDA1}}(r_s, \eta) = \begin{cases} \varepsilon_c^{\text{GLDA1}}(r_s, \eta) & \eta < 1 \\ \varepsilon_c^{\text{LDA1}}(r_s) & \eta \geq 1 \end{cases} \tag{6.41}$$

When applied to UEGs with $\eta \geq 1$, the gLDA1 and LDA1 kernels are, of course, identical. However, when applied to gases with $\eta < 1$, they behave differently and, by construction, the gLDA1 kernel is exact (within fitting errors) for any n-ringium.

The gLDA1 kernel defaults back to the LDA1 kernel at points where $\eta > 1$ but we cannot predict *a priori* whether this will cause it to under-estimate or to over-estimate the GLDA. If the monotonic increase in the magnitude of the kernel between $\eta = 0$ and $\eta = 1$ continues beyond $\eta = 1$, then the gLDA1 kernel (which assumes that the kernel is constant beyond $\eta = 1$) will underestimate the GLDA1 kernel and consequently underestimate the magnitude of the correlation energies in systems with high-curvature regions.

Until the true kernel for $\eta > 1$ is known, we cannot draw any firm conclusions about the accuracy of GLDA1. However, it is reasonable to conjecture that even the imperfect gLDA1 may be superior to LDA1 for Density Functional Theory (DFT)

calculations on inhomogeneous 1D systems and we now explore this through some preliminary validation studies.

6.5 Validation

Having defined the gLDA1 functional, we turn now to its validation. The functional is exact by construction for any n-ringium, so we require systems with non-uniform densities. Here we choose the ground states of n electrons in a 1D box of length $L = \pi$ (a family that we call the n-boxiums) and of n electrons in a 1D harmonic well with force constant $k = 1$ (a family that we call the n-hookiums). We will apply the functional to 1D molecules in the next chapter. The HOMO–LUMO gap in n-boxium increases roughly linearly with n while in n-hookium it slowly decreases. We therefore regard them as "large-gap" and "small-gap" systems, respectively.

Given that the fitting errors (Table 6.3) in the gLDA1 functional can be of the order of 0.1 mE_h, we aimed to obtain the energies of the n-boxium and n-hookium to within 0.1 mE_h of their complete basis set (CBS) limits. This is easily achieved for the HF, LDA1 and gLDA1 energies, because they converge exponentially [246–248] with the size M of the one-electron basis, but it is less straightforward for traditional post-HF energies.

We analysed the convergence behavior (see Appendix 6.B) of Møller-Plesset perturbation (MP2 and MP3) and full configuration interaction (FCI) energies in 2-ringium, 2-boxium and 2-hookium and our results are summarised in Table 6.4. From these, we devised appropriate extrapolation formulae and applied these to the energies obtained with our largest basis sets. We also used QMC calculations [242] to assess the accuracy of our extrapolated FCI energies.

Tables 6.5 and 6.6 show the energies obtained for 5-boxium and 5-hookium, respectively, as the basis set size increases from $M = 5$ to $M = 30$. The three components of the third-order energy are separated because of their different convergence behaviors. Table 6.7 summarizes our best estimates of the HOMO–LUMO gaps, together with the HF, LDA1, gLDA1, MP2, MP3 and FCI energies, for n-boxium and n-hookium with $n = 2, 3, 4$ or 5.

6.5.1 n-Boxium

The 2-boxium system (albeit with length $L = 3$) was studied in a basis of delta functions by Salter *et al.* [249] and, using 804609 basis functions, they obtained energies within roughly 10 μE_h of the exact values. The present work is the first study of n-boxium with $n \geq 3$.

	MP2	MP3	FCI
2-ringium	$O(M^{-3})$	$O(M^{-3})$	$O(M^{-3})$
2-boxium	$O(M^{-3})$	$O(M^{-3})$	$O(M^{-3})$
2-hookium	$O(M^{-3/2})$	$O(M^{-3/2}/\ln M)$	$O(M^{-3/2})$

Table 6.4: Basis set truncation errors ΔE_M for the energies in two-electron systems
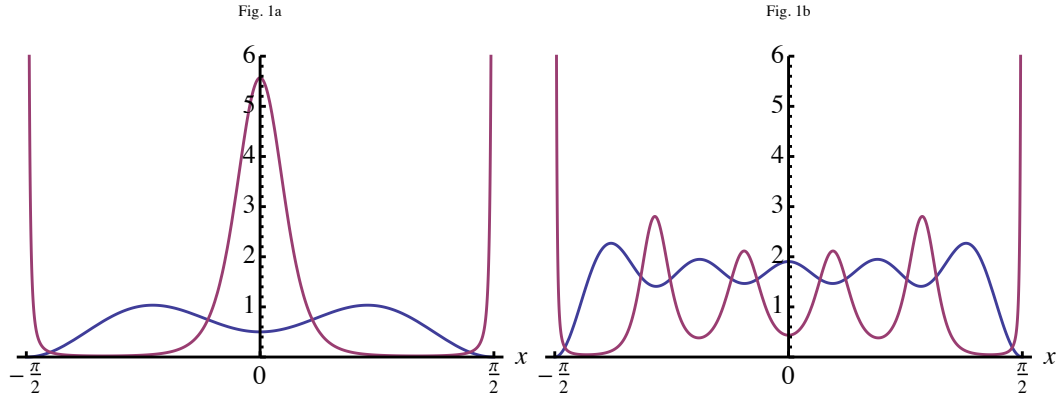
Figure 6.1: HF density $\rho(x)$ (blue) and curvature $\eta(x)$ (red) in 2-boxium (left) and 5-boxium (right)

The orbitals of 1-boxium are

$$\phi_m(x) = \begin{cases} \sqrt{2/\pi}\cos mx & m \text{ is odd} \\ \sqrt{2/\pi}\sin mx & m \text{ is even} \end{cases} \qquad (m = 1, 2, 3, \ldots) \qquad (6.42)$$

and the first M of these form a convenient orthonormal basis for expanding the HF orbitals in n-boxium. The antisymmetrised two-electron integrals $\langle \mu\sigma||\nu\lambda\rangle$ can be found in terms of the Sine and Cosine Integral functions [100] and we have used these to perform SCF calculations with up to $M = 30$ basis functions. Our convergence criterion was $\max |[\mathbf{P}, \mathbf{F}]| < 10^{-5}$.

We first discuss 2-boxium. Choosing $M = 8$ yields the HF orbitals

$$\psi_1(x) = 0.994844\,\phi_1(x) - 0.101256\,\phi_3(x) - 0.005729\,\phi_5(x) - 0.000044\,\phi_7(x)$$
$$(6.43a)$$

$$\psi_2(x) = 0.999715\,\phi_2(x) - 0.023850\,\phi_4(x) + 0.000728\,\phi_6(x) - 0.000176\,\phi_8(x)$$
$$(6.43b)$$

and Fig. 6.1 reveals that the density ρ has maxima at $x \approx \pm\pi/4$, indicating that an electron is likely to be found in these regions. LDA1 interprets these maxima as the most strongly correlated regions in the well and, through Eqs (6.2) and (6.39),

predicts the correlation energy

$$E_c^{\text{LDA1}} = \int_{-\pi/2}^{\pi/2} \rho(x)\varepsilon_c^{\text{LDA1}}(r_s)\,dx = -46.1 \text{ m}E_{\text{h}} \qquad (6.44)$$

In contrast, because the hole curvature η is strongly peaked at the center and edges of the box and is small near the density maxima, gLDA1 identifies the *center* of the box as the most correlated region and Eqs (6.23) and (6.41) predict the much smaller correlation energy

$$E_c^{\text{gLDA1}} = \int_{-\pi/2}^{\pi/2} \rho(x)\varepsilon_c^{\text{gLDA1}}(r_s,\eta)\,dx = -11.0 \text{ m}E_{\text{h}} \qquad (6.45)$$

LDA1 and gLDA1 offer very different qualitative and quantitative descriptions of 2-boxium, but both perturbation theory ($E_c^{\text{MP2}} = -8.33 \text{ m}E_{\text{h}}$ and $E_c^{\text{MP3}} = -9.45 \text{ m}E_{\text{h}}$) and near-exact calculations ($E_c^{\text{FCI}} = -9.82 \text{ m}E_{\text{h}}$) support the gLDA1 picture.

We have also performed HF, LDA1, gLDA1, MP2, MP3 and FCI calculations on 3-, 4- and 5-boxium and the density and curvature for 5-boxium are shown on the right of Fig. 6.1. Both functions oscillate much more rapidly but with much smaller amplitude than in 2-boxium, and it is easy to foresee that, as the number of electrons becomes large, both the density and the curvature will become increasingly uniform.

The convergence of the 5-boxium energies is shown in Table 6.5 and confirms the theoretical predictions of Table 6.4. The LDA1 energies, which depend only on the density $\rho(x)$, converge rapidly, changing by less than 1 μE_{h} beyond $M = 11$. The HF and gLDA energies, which depend on the orbitals (rather than the density) converge more slowly, achieving 1 μE_{h} convergence around $M = 20$. Because the occupied orbitals converge more rapidly than the virtual ones [250], the $\mathcal{O}^4\mathcal{V}^2$ component of MP3 converges almost as fast as HF, the $\mathcal{O}^3\mathcal{V}^3$ component (which is negative) converges more slowly, and the $\mathcal{O}^2\mathcal{V}^4$ component (which is positive) even more slowly.[2] Because of the resulting differential cancellation [251, 252], the total 3rd-order contribution initially becomes more negative, reaches a minimum at $M = 13$ and rises thereafter. The MP2 energy is the most slowly converging, and changes by 60 μE_{h} between $M = 29$ and $M = 30$. It is interesting to note the almost perfectly linear growth of the third-order energies. Because the n-boxiums are large-gap systems, MP2 and MP3 work well, recovering more than 92% and 99% of the correlation energy in 5-boxium.

---

[2]The symbols $\mathcal{O}$ and $\mathcal{V}$ refer to the number of occupied and virtual orbitals, respectively. The $\mathcal{O}^4\mathcal{V}^2$ component, for example, involves four sums over occupied orbitals and two over virtual orbitals.

M	E_{HF}	$-E_c^{\text{LDA1}}$	$-E_c^{\text{gLDA1}}$	$-E_c^{\text{MP2}}$	$E^{(3)}$ components			$-E_c^{\text{FCI}}$
					O^4V^2	O^3V^3	O^2V^4	
5	40.990 531	126.517	68.858	0	0	0	0	0
6	40.855 806	126.499	63.929	0	0	0	0	0
7	40.807 556	126.486	63.678	16.020	1.129	-3.512	0.804	17.840
8	40.798 066	126.482	63.314	28.753	1.728	-6.379	1.683	32.157
9	40.793 901	126.478	63.208	38.619	2.085	-8.570	2.475	43.234
10	40.793 518	126.478	63.207	45.564	2.276	-10.046	3.104	50.937
11	40.792 520	126.477	63.067	49.972	2.371	-10.871	3.577	55.640
12	40.792 237	"	63.024	53.055	2.426	-11.394	3.932	58.850
13	40.792 064	"	63.026	55.272	2.458	-11.729	4.203	61.102
14	40.792 057	"	63.031	56.876	2.478	-11.939	4.411	62.682
15	40.792 051	"	63.019	58.059	2.491	-12.071	4.572	63.815
16	40.792 051	"	63.017	58.946	2.499	-12.157	4.699	64.646
17	40.792 049	"	63.026	59.624	2.505	-12.214	4.799	65.271
18	40.792 049	"	63.027	60.151	2.509	-12.254	4.880	65.750
19	40.792 049	"	63.028	60.568	2.512	-12.282	4.945	66.124
20	40.792 048	"	63.028	60.901	2.514	-12.302	4.998	66.420
21	"	"	63.029	61.170	2.515	-12.317	5.042	66.658
22	"	"	"	61.391	2.516	-12.328	5.079	66.852
23	"	"	"	61.574	2.517	-12.337	5.110	67.011
24	"	"	"	61.726	2.517	-12.343	5.136	67.143
25	"	"	"	61.854	2.518	-12.349	5.158	67.253
26	"	"	"	61.963	2.518	-12.353	5.176	67.346
27	"	"	"	62.055	2.519	-12.356	5.193	67.425
28	"	"	"	62.134	"	-12.358	5.207	67.493
29	"	"	"	62.203	"	-12.360	5.219	67.551
30	"	"	"	62.262	"	-12.362	5.230	67.601

Table 6.5: Basis set convergence of E_{HF} (in E_{h}) and E_c energies (in mE_{h}) in 5-boxium

Our best estimates of the CBS limit HF and correlation energies are summarised in the left half of Table 6.7. Because LDA1 operates without the benefit of curvature information, it gravely overestimates the correlation energy, by between a factor of five (for 2-boxium) and a factor of just under two (for 5-boxium). In contrast, gLDA1 is within 12% of the true correlation energy for all n-boxiums studied.

6.5.2 n-Hookium

Electrons in 3D harmonic wells have been studied by numerous authors [102, 253–266] but this is the first investigation of n electrons in a 1D harmonic well. The orbitals of 1-hookium are

$$\phi_m(x) = \frac{H_{m-1}(x)\exp(-x^2/2)}{\sqrt{\pi^{1/2}2^{m-1}(m-1)!}} \qquad (m = 1, 2, 3, \ldots) \qquad (6.46)$$
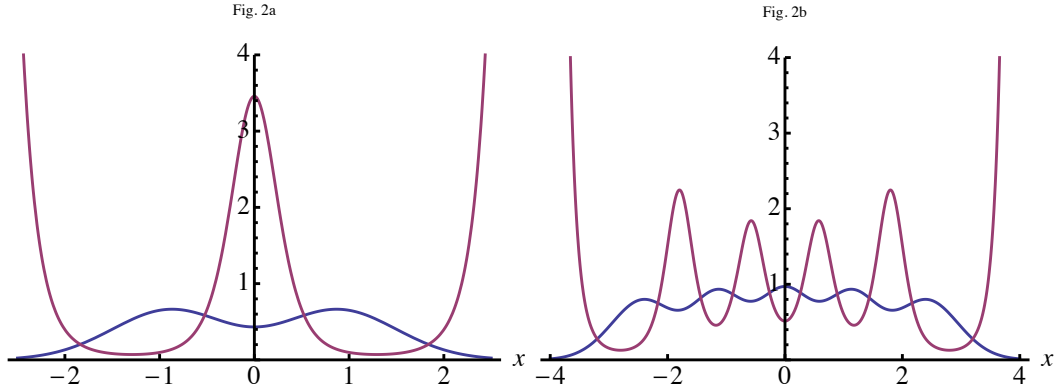
Figure 6.2: HF density $\rho(x)$ (blue) and curvature $\eta(x)$ (red) in 2-hookium (left) and 5-hookium (right)

and the first M of these form a convenient orthonormal basis for expanding the HF orbitals in n-hookium. The antisymmetrised two-electron integrals $\langle \mu\sigma || \nu\lambda \rangle$ can be found in closed form (*e.g.* see Appendix 6.A) and we have used these to perform SCF calculations with up to $M = 30$ basis functions. Our convergence criterion was $\max |[\mathbf{P}, \mathbf{F}]| < 10^{-5}$.

We first discuss 2-hookium. Choosing $M = 8$ yields the HF orbitals

$$\psi_1(x) = 0.989962 \, \phi_1(x) + 0.139577 \, \phi_3(x) - 0.021464 \, \phi_5(x) + 0.005740 \, \phi_7(x) \tag{6.47a}$$

$$\psi_2(x) = 0.997679 \, \phi_2(x) + 0.067586 \, \phi_4(x) - 0.008026 \, \phi_6(x) + 0.001894 \, \phi_8(x) \tag{6.47b}$$

and Fig. 6.2 reveals that the density and curvature are softened versions of those in 2-boxium. As before, LDA1 interprets the density maxima as regions of strong correlation, predicting

$$E_c^{\mathrm{LDA1}} = \int_{-\infty}^{\infty} \rho(x)\varepsilon_c^{\mathrm{LDA1}}(r_s) \, dx = -42.2 \ \mathrm{m}E_{\mathrm{h}} \tag{6.48}$$

whereas gLDA1 finds that almost all of the correlation comes from a narrow region near the middle of the well and predicts

$$E_c^{\mathrm{gLDA1}} = \int_{-\infty}^{\infty} \rho(x)\varepsilon_c^{\mathrm{gLDA1}}(r_s, \eta) \, dx = -12.7 \ \mathrm{m}E_{\mathrm{h}} \tag{6.49}$$

As for 2-boxium, LDA1 and gLDA1 offer entirely different pictures of electron correlation but both perturbation theory ($E_c^{\mathrm{MP2}} = -10.78 \ \mathrm{m}E_{\mathrm{h}}$ and $E_c^{\mathrm{MP3}} = -12.66 \ \mathrm{m}E_{\mathrm{h}}$) and near-exact calculations ($E_c^{\mathrm{FCI}} = -13.55 \ \mathrm{m}E_{\mathrm{h}}$) agree that gLDA1 is closer to the truth.

				$E^{(3)}$ components			
M	E_{HF}	$-E_c^{\mathrm{LDA1}}$	$-E_c^{\mathrm{gLDA1}}$	$-E_c^{\mathrm{MP2}}$	O^4V^2	O^3V^3	O^2V^4	$-E_c^{\mathrm{FCI}}$
5	19.649 014	116.419	75.381	0	0	0	0	0
6	19.353 767	115.709	60.013	0	0	0	0	0
7	19.180 033	114.892	64.207	18.983	2.783	−7.833	1.952	23.103
8	19.171 222	114.736	63.602	27.047	3.466	−11.252	2.972	33.352
9	19.167 260	114.619	63.990	33.786	4.058	−14.364	4.077	42.140
10	19.165 782	114.658	63.679	37.870	4.298	−16.063	4.812	47.219
11	19.165 244	114.680	63.512	41.400	4.488	−17.523	5.434	51.621
12	19.165 079	114.681	63.381	44.276	4.633	−18.697	5.973	55.159
13	19.164 701	114.684	63.163	46.478	4.729	−19.539	6.417	57.776
14	19.164 677	114.685	63.238	48.459	4.813	−20.301	6.807	60.137
15	19.164 499	114.687	63.059	49.999	4.870	−20.854	7.137	61.905
16	19.164 467	114.687	63.063	51.368	4.919	−21.340	7.428	63.459
17	19.164 400	114.687	62.993	52.493	4.956	−21.717	7.679	64.704
18	19.164 370	114.687	62.957	53.476	4.987	−22.036	7.901	65.768
19	19.164 342	114.687	62.940	54.317	5.012	−22.297	8.097	66.658
20	19.164 323	114.687	62.917	55.049	5.032	−22.515	8.271	67.417
21	19.164 309	114.688	62.899	55.688	5.049	−22.698	8.427	68.066
22	19.164 299	”	62.897	56.250	5.063	−22.851	8.567	68.623
23	19.164 291	”	62.885	57.746	5.075	−22.981	8.692	69.105
24	19.164 287	”	62.889	57.187	5.085	−23.091	8.806	69.525
25	19.164 283	”	62.885	57.579	5.094	−23.185	8.909	69.891
26	19.164 281	”	62.888	57.931	5.101	−23.266	9.003	70.214
27	19.164 279	”	62.897	58.248	5.108	−23.335	9.088	70.500
28	19.164 278	”	62.898	58.534	5.113	−23.396	9.167	70.754
29	19.164 278	”	62.903	58.793	5.118	−23.448	9.239	70.982
30	19.164 277	”	”	59.029	5.123	−23.495	9.305	71.186

Table 6.6: Basis set convergence of E_{HF} (in E_{h}) and E_c energies (in mE_{h}) in 5-hookium

We have also performed HF, LDA1, gLDA1, MP2, MP3 and FCI calculations on 3-, 4- and 5-hookium and the density and curvature for 5-hookium are shown on the right of Fig. 6.2. As before, both functions oscillate more rapidly but with smaller amplitude than in 2-hookium and it is clear that, as the number of electrons becomes large, both functions will become increasingly uniform [267].

The convergence of the 5-hookium energies is shown in Table 6.6. As in 5-boxium, the LDA1 energies converge most rapidly, followed by the HF and gLDA1 energies, then the O^4V^2, O^3V^3 and O^2V^4 components of the third-order energy, and finally the MP2 energy. However, each of these energies converges significantly more slowly than its 5-boxium analog. All of these observations are consistent with the theoretical predictions of Table 6.4. Because the n-hookiums are smaller-gap systems, MP2 and MP3 are less successful than for n-boxium, recovering roughly 85% and 96% of the correlation energy in 5-hookium.

	n-boxium ($L = \pi$)				*n*-hookium ($k = 1$)			
	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
E_{HF}	3.48451	10.37969	22.42489	40.79205	2.74367	6.63671	12.12335	19.16428
H-L gap	4.01	5.28	6.47	7.61	1.75	1.72	1.69	1.67
$-E_c^{\mathrm{LDA1}}$	46.1	72.5	99.4	126.5	42.2	65.9	90.1	114.7
$-E_c^{\mathrm{gLDA1}}$	11.0	26.3	44.0	63.0	12.7	28.0	44.9	62.9
$-E_c^{\mathrm{MP2}}$	8.3	23.1	41.8	62.8	10.8	26.0	43.7	63.0
$-E_c^{\mathrm{MP3}}$	9.5	25.6	45.4	67.3	12.7	30.0	49.8	71.1
$-E_c^{\mathrm{FCI}}$	9.8	26.2	46.1	68.0	13.5	31.8	52.4	74.3

Table 6.7: E_{HF} and HOMO–LUMO gap (in E_{h}) and E_c (in mE_{h}) in *n*-boxium and *n*-hookium

Our best estimates of the CBS limit HF and correlation energies are summarised in the right half of Table 6.7. As before, whereas LDA1 seriously overestimates the correlation energies, gLDA1 is within 15% of the true correlation energy in all cases. It is interesting to note that $|E_c(\textit{n}\text{-hookium})| > |E_c(\textit{n}\text{-boxium})|$ in all cases but that, whereas gLDA1 correctly predicts this trend, LDA1 reverses it.

6.6 Conclusion

The traditional Local Density Approximation (LDA) is exact by construction for an infinite uniform electron gas with Seitz radius r_s. However, it significantly overestimates the magnitudes of correlation energies in finite uniform electron gases, such as those created when n electrons are placed on the surface of a $\mathcal{D}$-dimensional sphere. This overestimation, which becomes even more pronounced in non-uniform gases, led us to seek generalizations of the LDA which are exact for both infinite and finite uniform gases and here we have proposed that the local hole curvature η provides the necessary information to achieve this goal. For present purposes, we have extracted η from the HF wave function: this requires only the occupied HF orbitals.

By fitting accurately calculated correlation energies for systems of n electrons on a ring, we have constructed the Generalised Local Density Approximation for one-dimensional systems and this has yielded a correlation kernel $\varepsilon_c(r_s, \eta)$ and a corresponding functional which we call GLDA1. To this point, we have considered only gases in which $\eta \leq 1$ and, consequently, the GLDA1 functional is not yet defined for gases with higher curvature. However, if we assume that the the correlation kernel becomes flat, *i.e.* that $\varepsilon_c(r_s, \eta) = \varepsilon_c(r_s, 1)$ when $\eta > 1$, we obtain an approximation to GLDA1 which we call gLDA1.

We have applied the traditional LDA1 functional and the curvature-corrected gLDA1 functional to electrons trapped in 1D boxes or in 1D harmonic wells and, by comparing the predicted correlation energies with those obtained from MP2, MP3 and Full CI calculations, we have discovered that gLDA1 is much more accurate than LDA1 in all cases.

We have also observed that gLDA1 tends to underestimate the magnitudes of correlation energies. This suggests that the true GLDA1 kernel continues to rise, *i.e.* that $|\varepsilon_c(r_s, \eta)| > |\varepsilon_c(r_s, 1)|$ but systematic examination of high-curvature $(\eta > 1)$ gases is required to test this. Such exploration is an important topic for future research and will allow the GLDA1 functional to be completely defined and tested.

Although we have presented relatively few calculations here, and much more investigation is warranted, these preliminary results suggest that "Curvature-Corrected Density Functional Theory (CC-DFT)" may offer an efficient pathway to improvements over existing functionals.

6.A   Calculation of matrix elements

The matrix elements in Eq. (6.22) are expressed as expectation values of operators over the HF wave function. Therefore, because Φ^2 and its reduced density matrices, *e.g.*

$$\rho_2(\theta_1, \theta_2) = \rho(\mathbf{r})^2 \left(1 - \left[\frac{\sin n(\theta_1 - \theta_2)/2}{n \sin(\theta_1 - \theta_2)/2} \right]^2 \right) \tag{6.50}$$

have *finite* Fourier expansions, integrals of their products with Fourier expansions of operators reduce to *finite* sums.

The Fourier expansions of bounded operators on a unit ring are straightforward, *e.g.*

$$r_{12}^2 = 2 - 2\cos(\theta_1 - \theta_2) \tag{6.51}$$

$$\nabla r_{12} \cdot \nabla r_{12} = 1 + \cos(\theta_1 - \theta_2) \tag{6.52}$$

$$r_{12} = -\frac{4}{\pi} \sum_{a=-\infty}^{\infty} \frac{e^{ia(\theta_1 - \theta_2)}}{4a^2 - 1} \tag{6.53}$$

$$\nabla r_{12} \cdot \nabla r_{13} = \left[\frac{4i}{\pi} \sum_{a=-\infty}^{\infty} \frac{a e^{ia(\theta_1 - \theta_2)}}{4a^2 - 1} \right] \left[\frac{4i}{\pi} \sum_{b=-\infty}^{\infty} \frac{b e^{ib(\theta_1 - \theta_3)}}{4b^2 - 1} \right] \tag{6.54}$$

The expansions of unbounded operators, *e.g.*

$$r_{12}^{-1} = -\frac{2}{\pi} \sum_{a=-\infty}^{\infty} \left(\sum_{p=1}^{|a|} \frac{1}{2p-1} \right) e^{ia(\theta_1 - \theta_2)} \tag{6.55}$$

are delicate (they converge only in the Cesàro mean [100]) but this is sufficient for our purposes because we require only a few of the low-order Fourier coefficients. The expansions of "cyclic" operators (*e.g.* $r_{12}r_{23}r_{31}$) are not simple products and must be derived separately.

Thus, for example, to find the $\langle \Phi | r_{12} r_{13} | \Phi \rangle$ integral in 3-ringium, the Fourier expansion

$$\Phi^2 = \frac{[2 - 2\cos(\theta_1 - \theta_2)]\,[2 - 2\cos(\theta_1 - \theta_3)]\,[2 - 2\cos(\theta_2 - \theta_3)]}{3!(2\pi)^3} \tag{6.56}$$

is combined with Eq. (6.53) to yield

$$\langle \Phi | r_{12} r_{13} | \Phi \rangle = \frac{16}{\pi^2} \sum_{a=-2}^{2} \sum_{b=-2}^{2} \iiint \frac{e^{ia(\theta_1 - \theta_2)}}{4a^2 - 1} \frac{e^{ib(\theta_1 - \theta_3)}}{4b^2 - 1} \Phi^2 d\theta_1 d\theta_2 d\theta_3$$
$$= \frac{16384}{675\pi^2} \tag{6.57}$$

6.B Extrapolation of perturbation energies

It is common these days to estimate the CBS limit of post-HF correlation energies by extrapolation [268]. Pioneering work by Schwartz [269], Hill [270] and Kutzelnigg and Morgan [271] showed that, for atoms in 3D, the second-order energy contributions from basis functions with angular momentum ℓ converge asymptotically as $(\ell + 1/2)^{-4}$.

While generating the data in Section 6.5, we found that the MP2 and MP3 energies converge so slowly (Tables 6.5 and 6.6) that the CBS limit is not reached (within our 0.1 mE_h target accuracy), even with our largest ($M = 30$) basis set. This is particularly noticeable for n-hookium. We therefore needed to develop and apply appropriate extrapolation procedures.

To this end, we analyzed the convergence of the second-order energy

$$E^{(2)} = \sum_{r=3}^{\infty} \sum_{s=r+1}^{\infty} \frac{\langle 12 || rs \rangle^2}{\epsilon_1 + \epsilon_2 - \epsilon_r - \epsilon_s} \tag{6.58}$$

obtained from the non-interacting orbitals and orbital energies in 2-boxium and
2-hookium. In n-hookium, the double-bar integral is

$$\langle 12||rs \rangle = \frac{(-1)^{(r-s+1)/2}\sqrt{2}}{\pi} \frac{\Gamma((r+s-2)/2)}{\sqrt{\Gamma(r)\Gamma(s)}} \tag{6.59}$$

if $r+s$ is odd but it vanishes if $r+s$ is even. The orbital energies are given by
$\epsilon_k = k - 1/2$. By substituting these expressions into (6.58) and making use of
Stirling's approximation [100], one can show that the error introduced by truncating
the basis after M functions is

$$\Delta E_M^{(2)} = \sum_{r=3}^{M} \sum_{s=r+1}^{M} \frac{\langle 12||rs \rangle^2}{\epsilon_1 + \epsilon_2 - \epsilon_r - \epsilon_s} - E^{(2)}$$
$$\sim \frac{1}{3(\pi M)^{3/2}} + O(M^{-2}) \tag{6.60}$$

The closed-form expression for the $\langle 12||rs \rangle$ integral in n-boxium is cumbersome
but a similar analysis reveals that the analogous truncation error is $O(M^{-3})$. The
truncation errors in the third-order energy can be found in the same way and all of
our results are summarised in Table 6.4.

The MP2, MP3 and FCI energies obtained with our largest basis sets conform
to these analytical predictions and allowed us to extrapolate reliably to the CBS
energies given in Table 6.7. The good agreement between our extrapolated FCI
energies and QMC energies further increases our confidence in these results.

Chapter 7

DFT benchmarks for one dimensional chemistry

7.1 Introduction

In Chapters 2–5 we have described a new paradigm for theoretical studies of atoms and molecules strictly confined to one dimension (1D). In this chapter we will apply the newly constructed gLDA1 functional to these molecules, comparing it to the LDA and the symmetry broken LDA recently constructed by Rogers, Loos and the present author [144]. This serves as both a comparative study of the functionals and test of DFT's applicability in our model of 1D chemistry.

One major advantage of the LDA is its large degree of fortuitous error cancellation when used to evaluate reaction energies [272], where despite the poor performance of the LDA functional for evaluating the components of the reaction the overall result can be relatively useful. Both the SBLDA and GLDA extensions to the LDA have been shown to offer improvements, sometimes quite substantial, when used to evaluate total correlation energies. Their ability to describe reactions has not yet been tested however, and it is unclear how the errors of each component involved will interact. Reaction descriptions are the ultimate goal of theoretical methods, and so one major aim of this chapter is to explore the new functionals performance at such a task.

We begin this chapter by describing the previously defined methods that are yet to appear, along with some of the necessary framework for performing the benchmarks. Of particular note, Sec. 7.2.3 discusses the unique feature that correlation in 1D is divided into two separate characters, which we have termed inter- and intra-domain correlation. In Sec. 7.3 we present the results of applying DFT to estimating intra-

domain correlation. The next chapter contains a section on our preliminary attempts to model inter-domain correlation.

7.2   Methods

In Chapter 6 we gave definitions of some of the functionals that we use here. In particular we have defined an LDA functional for 1D (See Eq. (6.39)), and the correlation kernel of the GLDA1 functional for $\eta \in [0,1]$ (See Eq. (6.40)). We continue to evaluate the functionals for the Hartree-Fock density, rather than self-consistently using the Kohn-Sham formalism.

7.2.1   Generalised LDA

In this chapter we generalise our definition of the gLDA1 functional (Eq. (6.41)) slightly. Here we define the αLDA functional

$$\varepsilon_c^{\alpha\text{LDA}}[\rho, \eta] = \begin{cases} \varepsilon_c^{\text{GLDA}}[\rho, \eta], & \eta < 1 \\ \alpha\varepsilon_c^{\text{LDA}}[\rho], & \eta \geq 1 \end{cases} \tag{7.1}$$

where the functional now uses a scaled LDA energy as a fallback when the correct GLDA functional is undefined. We will refer to three choices of α below:

- The αLDA functional will refer to an α chosen by minimising the errors in atomic energies.

- The 0LDA functional takes $\alpha = 0$, removing the LDA component completely and allowing the contribution of the correctly parameterised GLDA to be isolated.

- Finally, we continue to refer to the case where $\alpha = 1$ as gLDA.

7.2.2   Symmetry-Broken functionals

The symmetry broken (SB) correction is a much simpler approach to improving the performance of the raw LDA. The reference correlation values for the UEG that are used to construct the LDA functional are typically based upon the difference between a near-exact energy and the Hartree-Fock (HF) energy in the Fermi fluid state, where the electrons are completely delocalised. It can be shown however that the Fermi fluid state is never the true HF ground state however [273, 274].

Rather it is possible to show that in 1D the lowest energy state is always given by the Wigner crystal, where the electrons are allowed to localise at lattice sites to

	$\Delta\varepsilon_c^{\text{SBLDA}}$	a_{SB}	b_{SB}	c_{SB}
a_0	-0.0646228	-0.0348811	0.185088	1.53424
a_1	0.535062	0.0348811	0.00794842	-0.0263157
a_2	-0.490719	0.0438699	-0.174901	0.436547
a_3	0.168939			
b_0	53.1171	1	1	1.61107
b_1	1.53114	1.34815		0.361882
b_2	2.19606	-0.672131	-0.571199	-0.718872
b_3	1			0.245916

Table 7.1: Parameters of numerical approximants used in the definition of symmetry broken functionals.

minimise the energy [144]. The SBLDA correction measures the difference between the Fermi fluid and Wigner crystal HF energies and hence the ensuing change in correlation energy. This leads to the SBLDA functional giving a lower correlation estimate than the LDA, which is well known to overestimate the total correlation in both 3D [21] and, as we have determined earlier, 1D.

This functional was recently described by Rogers et al. [144], who give the following form for the correction

$$\Delta\varepsilon_c^{\text{SBLDA}}(r_s) = r_s^2 \frac{a_0 + a_1 r_s + a_2 r_s^2 + a_3 r_s^3}{b_0 + b_1 r_s^5 + b_2 r_s^{11/2} + b_3 r_s^6} \tag{7.2}$$

where $r_s = 1/2\rho$ is the Seitz radius and the a_i and b_i coefficients can be found in Table 7.1. This is added to the LDA energy to obtain the SBLDA energy.

The SB correction is not incompatible with the GLDA methodology, and we have also constructed a new SBgLDA functional. We have evaluated the Wigner crystal Hartree-Fock energies for the finite UEGs originally used to fit GLDA to obtain the new set of correlation energies given in Tab. 7.2. We then follow the same methodology as in Chapter 6, substituting the new reference correlation energies. Care has also been taken to accomodate the new low-density behaviour that is a consequence of taking the Wigner crystal state [144].

		r_s										
n	η	0	1/10	1/5	1/2	1	2	5	10	20	50	100
1	0	0	0	0	0	0	0	0	0	0	0	0
2	3/4	13.212	12.985	12.766	12.152	11.250	9.802	7.111	4.002	1.429	0.334	0.113
3	8/9	18.484	18.107	17.747	16.755	15.346	13.179	7.956	2.984	0.982	0.225	0.076
4	15/16	21.174	20.700	20.250	19.027	17.324	14.762	6.826	2.425	0.787	0.177	0.060
5	24/25	22.756	22.216	21.706	20.332	18.444	15.648	6.061	2.135	0.688	0.153	0.051
6	35/36	23.775	23.190	22.638	21.161	19.148	15.701	5.585	1.964	0.629	0.139	0.046
7	48/49	24.476	23.855	23.273	21.723	19.618	15.247	5.279	1.859	0.592	0.130	0.043
8	63/64	24.981	24.328	23.729	22.122	19.951	14.751	5.071	1.785	0.568	0.123	0.040
9	80/81	25.36	24.686	24.067	22.415	20.199	14.316	4.925	1.734	0.549	0.119	0.039
10	99/100	25.651	24.960	24.327	22.644	20.386	13.957	4.812	1.698	0.538	0.117	0.038
∞	1	27.416	26.597	25.91	23.430	18.882	11.979	4.314	1.524	0.477	0.086	0.032

Table 7.2: η and $-\varepsilon_c(r_s, \eta)$ (mE_{h} per electron) for the ground state of n electrons on a ring using the HF Wigner crystal reference

This process has resulted in the following functional form

$$\varepsilon_c^{\mathrm{SBGLDA}}[\rho, \eta] = -a_{SB}(\eta)\,_2F_1\left(\frac{3}{2}; 2, c_{SB}(\eta); -\left[\sqrt{\pi}\frac{\Gamma(c_{SB}(\eta))}{\Gamma(c_{SB}(\eta) - 3/2)}\frac{a_{SB}}{b_{SB}}\right]^{2/3}\right)$$
(7.3)

$$a_{SB}(\eta) = \frac{a_0 + a_1\sqrt{1-\eta} + a_2\eta}{b_0 + b_1\sqrt{1-\eta} + b_2\eta}$$
(7.4)

$$b_{SB}(\eta) = \frac{a_0 + a_1\sqrt{1-\eta} + a_2\eta}{b_0 + b_2\eta}$$
(7.5)

$$c_{SB}(\eta) = e^{-\frac{1}{20\rho}}\left(a_0 + a_1\sqrt{1-\eta} + a_2(1-\eta)\right)$$
$$+ \left(1 - e^{-\frac{1}{20\rho}}\right)\left(b_0 + b_1\sqrt{1-\eta} + b_2(1-\eta) + b_3(1-\eta)^{3/2}\right)$$
(7.6)

where the a_i and b_i coefficients can once again be found in Table 7.1.

7.2.3   Inter- and Intra-domain correlation

The separation of electrons into different domains causes a unique partitioning of the correlation energy within the molecule. One can conceive of two kinds of correlation present: interactions between electrons in the same domain (intra-domain correlation) and interactions between electrons in two different domains (inter-domain correlation). Intra-domain correlation is the same as observed in 3D molecules. Inter-domain correlation, however, is purely dispersive in nature, since there is rigorously no overlap in density of the two electrons.

While the similarity of intra-domain correlation to 3D correlation means any standard method can be used to evaluate it, dispersion is a traditional weak point of

DFT, which causes issues for evaluating the inter-domain correlation. The GLDA-type functionals exemplify this weakness. If we consider the ground-state helium atom, which has one electron on either side of the nucleus, then one can show that η must be equal to 0 everywhere. The electron to the left of the nucleus can never be found at the same point as the electron on the right. Due to the Dirichlet condition the quantity $P(u|r)$ (see Eq. (6.3)) must be 0 in any neighbourhood of $u = 0$ because $r \neq 0$ in the entire neighbourhood. Hence it follows that $d^2 P(u|r)/du^2 = 0$ when $u = 0$, and from that $\eta = 0$.

All GLDA-type functionals are equal to 0 when $\eta = 0$. This is an advantage, since it removes the self-interaction error that is inherent to the LDA. In this case, however, it also forces the functional to remove the dispersive correlation between the two electrons. LDA does give a non-zero correlation energy in this case, but this is precisely due to the self-interaction error, and not an indication that the LDA is correctly representing the physics involved.

As a result the density functionals used in this study can only evaluate the intra-domain correlation within a molecule.

7.2.4 The G1D test set

For testing the methods under consideration over a range of situations we have defined a set of molecules and their reactions which we call the G1D set. As the name suggests, we take inspiration from what is commonly known as the Pople G1 test set [275]. This set contains molecules with a variety bonding characters and four types of reactions for each: atomisation energies, ionisation energies, electron affinities and proton affinities.

We have selected a set of 42 stable molecules that were identified in Chapter 5. These include diatomics, triatomics and tetra-atomics and contains all three kinds of diatomic bonds identified as well as a group of weakly bound molecules. A full listing of the molecules chosen for the test set can be found in Appendix A.

The set also contains the same set of reactions as the G1 set. Both a "left" and "right" sided reaction is included where it is appropriate. For example, an electron can be removed from either the left-most or right-most domain of the molecule $_1H_2Li_2$ during an ionisation process.

7.2.5 Benchmark methods

For benchmark values to compare the functionals against we have chosen to use the MP3 level of theory, which was observed to be exceptionally accurate in Chapter 2.

As an additional benchmark, we also compare the MP2 level of theory along with functionals.

MPn energies are able to capture both intra- and inter-domain correlation. Slight modifications to the MPn paradigm must be taken in order to consider these separately. We define the intra-MPn method to this end. Here one simply performs an MPn calculation on each domain and sums the results, giving the intra-domain correlation exclusively. In each of these sub-calculations only the occupied and virtual domains are used in the summation, preventing any excitation coupling between electrons in different domains. Inter-domain correlation can be found easily by subtracting the intra-MPn energy from the total MPn energy.

7.3   Intra-domain correlation

To compute the intra-domain correlation energies we have added density functional capabilities to LegLag, which has been described in Chapter 4. We use the same methodologies, such as basis set choice, as we did in Chapter 5. In Sec. 7.3.1 we consider the ability of the functionals to reproduce the total correlation energy of atoms and molecules. In Sec. 7.3.2 we present a possible cause for the errors in the gLDA functional when applied to molecular systems and in Sec. 7.3.3 we compare the accuracy of the functionals for determining reaction behaviour.

We present statistical summaries of the correlation errors throughout this section. See Appendix A for a listing of the absolute correlation energies for each species of the G1D test set.

7.3.1   Absolute correlation energies

Tables 7.3 and 7.4 contain a summary of the correlation performance of the functionals described in Sec. 7.2. Table 7.3 contains total intra-domain correlation energies for the atoms lithium through neon, while Table 7.4 contains a summary of the same energies for the molecules within the G1D test set.

The most striking result is how incredibly poor the LDA becomes in 1D, which we have seen in Chapter 6. Unfortunately, while the SBLDA correction makes a significant effort to improve the LDA's results it is unable to overcome the failure of the parent LDA, leaving its performance at roughly the same level of quality. Note that the LDA does improve significantly as the electron density becomes more diffuse, much like in 3D, as evidenced by the increasing relative accuracy for atoms as the atomic charge increases. The more diffuse electron distribution better mimics the situation within a UEG, especially in 1D where the combination of the Super-Pauli

| | Correlation error (mE_h) | | | | | |
Atom	intra-MP2	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
Li	−0.158	59.928	50.925	−0.576	0.718	1.986	1.059
Be	−0.799	93.750	78.921	−4.434	1.792	7.894	5.890
B	−0.480	79.547	67.799	−2.508	1.637	5.701	4.194
C	−1.462	110.121	91.646	−7.918	1.593	10.915	7.754
N	−2.005	122.891	101.387	−11.197	0.894	12.745	9.504
O	−3.615	150.738	121.954	−19.840	−1.418	16.640	11.789
F	−2.787	136.701	111.268	−15.255	0.003	14.958	10.622
Ne	−4.402	161.458	129.518	−24.111	−2.750	18.188	12.882
	Correlation percentage					
Atom	intra-MP2	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
Li	88	4842	4130	54	157	257	184
Be	90	1325	1132	42	123	203	177
B	89	1844	1586	45	136	225	192
C	89	952	809	39	112	184	160
N	89	792	671	37	105	172	154
O	88	599	504	34	95	155	139
F	88	674	567	36	100	163	145
Ne	88	542	455	34	92	150	135

Table 7.3: Errors in intra-domain correlation energy estimates for 1D atoms. Intra-MP3 energies have been used as reference values, errors are given in milli-hartrees (mE_h) and as percentages of the intra-MP3 energy.

exclusion principle and the Dirichlet condition pushes electrons to spectacularly large distances from the nuclei.

In contrast to the LDA, the GLDA family of functionals are outstanding. Unfortunately they do not reach the quality of perturbation theory, particularly the consistency of its description. If we assume that the 0LDA is an accurate estimate of the correct correlation in the regions of space that it is able to describe then it appears that areas with high η values are crucial to correctly describing correlation. The atomic results show that this is particularly true as the overall density of electrons decreases.

This, combined with the poor LDA results, explains why the gLDA is overall the worst of the GLDA functionals. Despite clearly removing a large portion of the worst offending parts of the LDA it still results in a relatively large overestimation of the total correlation. In this case, however, the symmetry broken correction of SBgLDA gives a substantial overall improvement to the results of the gLDA functional, owing to the much improved starting point.

Despite this it is the αLDA functional which performs the best overall. Here the coefficient is fitted to minimise the error over the set of atomic correlations

	Correlation error (mE_h)					
Atom	intra-MP2	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
Mean	-1.051	170.547	143.360	-5.518	2.842	11.036	7.585
σ	0.568	45.999	38.482	3.400	1.089	4.696	3.604
Max	-0.146	267.896	224.836	-0.455	5.437	21.861	16.072
Min	-2.230	78.878	65.227	-12.506	0.783	2.024	1.128
	Correlation percentage					
Atom	intra-MP2	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
Mean	89	2550	2173	48	138	226	182
σ	1	1907	1656	9	22	35	12
Max	90	10313	9032	68	186	301	202
Min	87	1041	878	38	111	181	156

Table 7.4: A summary of errors in intra-domain correlation energy estimates for the molecules in the G1D test set. Intra-MP3 energies have been used as reference values, errors are given in milli-hartrees (mE_h) and as percentages of the intra-MP3 energy.

($\alpha = 0.505$). It is fortunate that this also works relatively well for the molecules of the G1D test set, suggesting that nearby nuclei do not significantly alter the behaviour of the electrons. This is easily explained by the Super-Pauli principle preventing the pairing of electrons into a single orbital. Therefore there are no truly significant shifts in environment for any electrons during the bonding process.

It must be noted that the parameter of the αLDA functional is ultimately compromised by the shifting importance of the regions with large η values as atomic charge varies. It would, of course, be possible to choose α such that the molecular estimates have an average error of 0 overall. The G1D set is comprised of molecules built primarily from lighter atoms, and so the compromise of fitting the αLDA to the full set of atoms does not favour its treatment of this set. Ultimately, this suggests a lack of consistency in its treatment of electron correlation.

7.3.2   High density limit of helium-like ions

When the gLDA functional was first tested on electrons confined to boxes and harmonic wells it was found to have exceptional, MP2 quality accuracy (See Sec. 6.5). The results in Sec. 7.3.1 show that this performance is not retained across other systems. It is not immediately clear why this is the case. To attempt to gain some clarity we have evaluated the high-density correlation limit of the one-sided helium-like ions, i.e. the limit as $Z \to \infty$ of an atom with two electrons found on one side of a nucleus with charge Z. This is the same limit as was found for the two-sided helium-like ions in Chapter 2 (Sec. 2.3.2).

This requires determining both the exact energy and the Hartree-Fock energy in this high-density limit. For the exact energy we have used Hylleraas type expansions [96–98] evaluated numerically, finding the following power series for the energy.

$$E_{\text{exact}} = \sum_{n=0}^{\infty} \varepsilon_{\text{exact},n} Z^{2-n} \tag{7.7}$$

$$\varepsilon_{\text{exact},0} = 0.625000 \tag{7.8}$$

$$\varepsilon_{\text{exact},1} = 0.282341 \tag{7.9}$$

$$\varepsilon_{\text{exact},2} = 0.185226 \tag{7.10}$$

To evaluate the corresponding expansion of the Hartree-Fock energy we have followed the method of Linderberg[105] where a perturbative expansion is applied to the problem. One of the key insights of Linderberg is to introduce a scaled unit charge $e\sqrt{Z}$, and the modified Hartree energy $E_{\text{HM}} = E_{\text{h}}$. The energy E_{HF} and the two occupied orbitals $u(r)$ and $v(r)$ are then expanded as the power series

$$E_{\text{HF}} = \sum_{n=0}^{\infty} \varepsilon_{\text{HF},n} Z^{-n} \tag{7.11}$$

$$u(r) = \sum_{n=0}^{\infty} u_n(r) Z^{-n} \tag{7.12}$$

$$v(r) = \sum_{n=0}^{\infty} v_n(r) Z^{-n} \tag{7.13}$$

such that the zero-th order gives the non-interacting orbitals and energy.

Assuming that the electrons are on the right side of the nucleus (i.e. $r > 0$) then u_0 and v_0 are the first two hydrogenic orbitals from Eq. (2.2)

$$u_0(r) = 2r \exp(-r) \tag{7.14}$$

$$v_0(r) = \frac{r(2-r)}{\sqrt{8}} \exp(-r/2) \tag{7.15}$$

Both $\varepsilon_{\text{HF},0}$ and $\varepsilon_{\text{HF},1}$ can be expressed in terms of the zero-th order orbitals. The former is simply the sum of the non-interacting orbital energies while the latter is the remainder of the Fock operator

$$\varepsilon_{\text{HF},0} = \langle u(r)|\hat{h}_0|u(r)\rangle + \langle v(r)|\hat{h}_0|v(r)\rangle = -\frac{5}{8} \tag{7.16}$$

$$\varepsilon_{\text{HF},1} = (uu||vv) = \{uu|vv\} - \{uv|uv\} = \frac{1}{729}\left(203 + 24\ln\frac{9}{8}\right) \tag{7.17}$$

where $\hat{h}_0 = -1/2\, d^2/dr^2 + 1/r$ and $\{uu|vv\}$ is a quasi-integral (see Sec. 3.2.4).

These two coefficients match those obtained from the Hylleraas expansion, and so these terms cancel when evaluating the correlation energy. The next coefficient can be obtained from u_1 and v_1, which are given by the differential equations

$$\left(\hat{h}_0 + \frac{1}{2}\right) u_1 = \varepsilon_{\text{HF},1} u_0 - u_0 |v_0 v_0) + v_0 |u_0 v_0) \tag{7.18}$$

$$\left(\hat{h}_0 + \frac{1}{2}\right) v_1 = \varepsilon_{\text{HF},1} v_0 - v_0 |u_0 u_0) + u_0 |u_0 v_0) \tag{7.19}$$

$$|u_0 u_0) = (8r + 2)e^{-2r} - (2r + 2\omega + 1) - (2r - 2\omega)^2 e^{-2(r-\omega)} \, \text{Ei}(-2\omega)$$
$$- (2r + 2\omega)^2 e^{-2(r+\omega)} \big(\text{Ei}(2\omega) - \text{Ei}(2r + 2\omega) \big) \tag{7.20}$$

$$|u_0 v_0) = \frac{1}{\sqrt{2}} \Bigg[\frac{8}{9}\left(r + \omega + \frac{1}{3}\right) - \frac{2}{3}(r + \omega)^2 + \frac{4}{3}\left(\omega^2 + 3t^2 - 2t - \frac{4}{9}\right) e^{-3/2t}$$
$$- (r - \omega - 2)(r - \omega)^2 e^{-3/2(r-\omega)} \, \text{Ei}\left(-\frac{3}{2}\omega\right)$$
$$- (r + \omega - 2)(r + \omega)^2 e^{-3/2(r+\omega)} \, \text{Ei}\left(\frac{3}{2}\omega\right)$$
$$+ (r + \omega - 2)(r + \omega)^2 e^{-3/2(r+\omega)} \, \text{Ei}\left(\frac{3}{2}r + \frac{3}{2}\omega\right) \Bigg] \tag{7.21}$$

$$|v_0 v_0) = \frac{e^{-r}}{8} \Bigg[4(2r^3 - 3r^2 + 2r + 1) + 2(4r - 3)\omega^2$$
$$- \big((r + \omega)^3 - 3(r + \omega)^2 + 2(r + \omega + 1)\big)e^r$$
$$- (r - \omega)^2 (r - \omega - 2)^2 e^{\omega} \, \text{Ei}(-\omega)$$
$$- (r + \omega)^2 (r + \omega - 2)^2 e^{-\omega} \big(\text{Ei}(\omega) - \text{Ei}(r + \omega) \big) \Bigg] \tag{7.22}$$

where ω is a softening parameter with the same function as in Sec. 3.2.4. The limit as $\omega \to 0$ must be taken once the solution is obtained.

Unfortunately solving these equations is extremely difficult due to the Ei functions in the inhomogoneities of the equations. Instead we use LegLag to determine the $\varepsilon_{\text{HF},2}$ coefficient numerically, from which we eventually obtain

$$\varepsilon_{\text{HF},2} = 0.178102 \tag{7.23}$$

The correlation can then be written as

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} = \sum_{n=0}^{\infty} \frac{\varepsilon_{\text{exact},n} - \varepsilon_{\text{HF},n}}{Z^n} \tag{7.24}$$

where the leading term of the expansion becomes 0.007124.

This is at least one of the sources of gLDA's worse than expected performance. The functional was constructed to obtain the correct correlation energy of the finite and infinite UEGs, whose high density correlation ranges from 13 to 27 mE_{h} (See Sec. 6.3), which are all larger than the 7 mE_{h} limit of the one-sided helium-like ions. Note, however, that the 0LDA functional underestimates the correct correlation energy of all the atoms considered. This leaves the possibility that the fully parameterised GLDA functional, which does not rely upon the LDA when η is above 1, could correctly reproduce the correlation of 1D atoms.

Aside from a possible explanation for the worse than expected gLDA functional performance, this discrepancy between the high-density limit of atoms and UEGs is curious on a more fundamental level. It is known that in the infinite-dimensional limit that high-density correlation energies are invariant under change in the external potential that confines the electrons [106]. While this does not remain rigorously true in finite dimensions, it has been shown numerically that assuming this fact is an excellent approximation down to two-dimensions. This assumption is at the heart of the motivation for the construction of the LDA, where the functional is constructed using UEGs before being applied to a wide variety of different external potentials. Clearly this assumption does not hold in 1D, and the LDA is a far weaker approximation than it is in other dimensionalities.

7.3.3 Reaction correlation energies

The primary goal of this chapter is to model reaction energies. A summary of applying DFT to find the intra-domain correlation contributions to the G1D reaction set described in Sec. 7.2.4 can be found in Table 7.5.

Immediately obvious is a wide degree of error cancellation for all the methods tested. The average error over the full set of reactions is typically 1–1.5 orders of magnitude less than the average error over the set of molecules. This is even greater for the LDA (approximately 2 orders of magnitude) and especially the SBLDA and SBgLDA (approximately 2.5 orders of magnitude).

Although the reaction errors are greatly improved over those for the molecular correlation energies the total correlation contribution in the reactions ranges between -3 mE_{h} and $+3$ mE_{h}, typically falling in the range of ± 1 mE_{h}. This completely

	Correlation error (mE_h)						
	intra-MP2	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
Atomisation energies							
MAE	0.042	2.264	2.167	0.095	0.445	0.824	0.333
MSE	0.000	−2.259	−1.664	−0.062	−0.442	−0.815	−0.320
Median (AE)	0.023	1.541	1.034	0.062	0.317	0.551	0.130
σ (AE)	0.048	2.247	2.503	0.114	0.403	0.824	0.370
Maximum (AE)	0.244	7.784	9.327	0.623	1.418	2.820	1.274
Ionisation energies							
MAE	0.146	14.866	9.119	0.607	0.429	1.339	0.743
MSE	0.130	−14.866	−9.119	0.491	−0.412	−1.297	−0.660
Median (AE)	0.139	13.892	7.591	0.535	0.286	1.121	0.409
σ (AE)	0.091	3.567	5.140	0.409	0.355	0.981	0.847
Maximum (AE)	0.530	23.226	21.894	2.408	1.193	3.320	2.895
Electron affinities							
MAE	0.077	11.274	4.277	0.195	0.149	0.441	0.106
MSE	−0.073	11.274	4.277	−0.160	0.136	0.426	0.065
Median (AE)	0.051	10.694	3.549	0.106	0.111	0.327	0.059
σ (AE)	0.078	2.220	2.383	0.225	0.165	0.357	0.165
Maximum (AE)	0.433	16.750	10.849	1.274	0.719	1.638	0.775
Proton affinities							
MAE	0.068	4.177	6.045	0.262	0.059	0.193	0.082
MSE	0.055	4.177	6.045	0.218	0.023	−0.168	−0.059
Median (AE)	0.047	3.762	6.051	0.240	0.038	0.159	0.044
σ (AE)	0.067	2.926	4.080	0.228	0.076	0.167	0.086
Maximum (AE)	0.264	9.793	12.739	0.972	0.496	0.880	0.413
All reactions							
MAE	0.087	8.779	5.750	0.307	0.239	0.658	0.298
MSE	0.028	0.363	0.431	0.136	−0.126	−0.382	−0.218
Median (AE)	0.061	9.321	4.868	0.196	0.110	0.346	0.081
σ (AE)	0.083	5.665	4.451	0.331	0.306	0.772	0.540
Maximum (AE)	0.530	23.226	21.894	2.408	1.418	3.320	2.895

Table 7.5: Statistical summary of errors in the intra-domain correlation energy estimates for the G1D reaction set. Intra-MP3 energies have been used as reference values, absolute errors are given in milli-hartrees (mE_h).

eliminates the LDA from being a useful approximation in this context, and although the SBLDA receives greater error cancellation and makes a significant improvement to the error it too remains unusable.

The four GLDA-type functionals do give errors which fall within a useful range. Interestingly, the 0LDA gives a relatively high quality estimate for the atomisation energy, suggesting that the correlation in the $\eta > 1$ regions should largely cancel across the reaction. All of the functionals extended to treat this range (αLDA, gLDA and SBgLDA) have a much more significant error for the atomisations. Clearly there is an aspect of electron correlation in this region that an LDA functional is unable to describe.

The accuracy of the 0LDA does not hold up over the other classes of reactions, where the $\eta > 1$ extension plays an important role in describing the energetics. The gLDA functional is not a useful extension either, and is almost strictly inferior to the 0LDA. Both the αLDAand SBgLDA do improve upon the 0LDA, although when comparing their average overall results they are only capable of a modest improvement.

The functional results suffer from some larger outliers however, and this appears to skew the average measures. The median error, which is much more resilient to such outliers, gives a more promising outlook for the GLDA-type functionals. In particular, both the αLDAand SBgLDA appear to approach the overall performance of intra-MP2. Recalling that the parametrisation of αLDAis somewhat compromised (as discussed in Sec. 7.3.1), the SBgLDA appears to be the best performing functional in this group. Of course, the large maximum errors means neither of these methods can be considered as reliable as intra-MP2.

7.4 Conclusion

It appears that DFT has a difficult relationship with 1D chemistry. The high accuracy of the Hartree-Fock approximation gives perturbation theory an exceptional starting point, resulting in very high accuracy that the LDA and the derivatives examined here cannot match. In addition, the separation of correlation into inter- and intra-domain regions, combined with the locality of the LDA and its derivatives, prevent DFT from being applied to the full correlation problem in a 1D molecule.

Unfortunately, over the G1D test set inter-domain correlation accounts for, on average, roughly half of the total correlation in a molecule. In order to model correlation in 1D chemistry with density functionals a dispersion correction that can estimate the inter-domain correlation is required.

Despite this 1D molecules have proven to be an interesting test for the GLDA type functionals described in Chapter 6. Our results here suggest that the GLDA in 1D does not benefit from the kind of error cancellation for reaction energies that the LDA does in 3D. The 1D LDA, however, does not appear to achieve the same feat as it does in 3D either. This makes it unclear how the GLDA will perform in 3D, leaving the original intent of this chapter unanswered. We do find, however, that such functionals are significant improvements over the standard LDA for reaction estimates. This should bode well for future development of GLDA functionals.

Chapter 8

Unpublished work

8.1 Self consistent GLDA theory

In the preceding chapters we have applied density functionals as a Post-Hartree-Fock level of theory. That is, instead of using the Kohn-Sham (KS) formalism we have performed a HF calculation and then applied the correlation functional to the HF density. Employing the KS equations would yield a superior result as it would allow the density to relax according to the correlation effects. Unfortunately a self-consistent formulation using the GLDA functional is not as simple as the KS equations due to the added complexity of including the η parameter.

At this point it is important to note where the GLDA functional sits in the context of other functional development. The key component of the η parameter is the Laplacian of the positional intracule that can be found in Eq. (6.7). This expression can be partially rewritten using the "kinetic energy density"

$$\nabla_{\mathbf{u}}^2 P(\mathbf{0}|\mathbf{r}) = 2\tau - \frac{|\nabla\rho|^2}{2\rho} \tag{8.1}$$

$$\tau = \sum_i^{\text{occ}} |\nabla\psi_i|^2 \tag{8.2}$$

where ψ_i are the HF orbitals and τ is the kinetic energy density. The use of the term τ puts the GLDA functional into the class of meta-GGA functionals. This class has gained some traction in the literature, and so there already exists some discussion of the self-consistent implementation of functionals using these kinds of terms.

In this section we develop the theory required to perform self-consistent GLDA calculations by tailoring previous work to the η parameter. We will work within the finite-basis formulation for GGA functionals developed by Pople, Gill and Johnson [276]. This necessitates determining the behaviour of the functional as the density

changes. Since τ is not expressible in terms of the density this proves to be somewhat difficult. This problem was first solved by Neumann *et. al.* [277], however we follow the slightly different notations used by Adamo *et. al.* during their implementation of a meta-GGA functional [278].

8.1.1 Theory

The GLDA functional that was defined in Chapter 6 is a 1D functional only. The method can scale to higher dimensions and here we will construct the self-consistent theory in a general manner so that it may be applied to any dimensionality. This will require the reintroduction of spin-coordinates. For the remainder of this section we will assume we are considering an α spin electron, the modifications to consider a β electron are trivial.

A key component of the Kohn-Sham equations is the potential of the exchange/correlation functional. This potential can be found by applying calculus of variations to the functional with respect to the density. When the functional includes τ, however, it is common to instead find the variation with respect to the KS orbitals. Adamo *et. al.* give the following expression for this, except we have added spin coordinates and a Laplacian of the density

$$\frac{\delta E_{\text{XC}}}{\delta \psi_i^\alpha} = \left(\nabla^2 \frac{\delta f}{\delta(\nabla^2 \rho_\alpha)} \right) \psi_i^\alpha - \left(\nabla \frac{\delta f}{\delta(\nabla \rho_\alpha)} \right) \psi_i^\alpha + \left(\frac{\delta f}{\delta \rho_\alpha} \right) \psi_i^\alpha$$
$$- \left(\nabla \frac{\delta f}{\delta \tau_\alpha} \right) \nabla \psi_i^\alpha - \left(\frac{\delta f}{\delta \tau_\alpha} \right) \nabla^2 \psi_i^\alpha \quad (8.3)$$

where f is the exchange/correlation functional. The exchange/correlation energy in the one-particle KS equations is then given by

$$E_{\text{XC}} = \sum_{ij} \int \frac{\delta E_{\text{XC}}}{\delta \psi_i^\alpha} \psi_j \qquad (8.4)$$

Using this potential will optimise the KS orbitals rather than the electronic density, which is a subtle but important distinction [279]. While there are methods which can optimise the density, e.g. the Optimised Effective Potential [279–282] or more recent work by Ryabinkin and Staroverov [283], the computational effort required is significantly greater. As a result working with the potential derived from varying the orbitals is generally preferred, and we will follow that preference here.

The finite-basis method of Pople *et. al.* defines a Fock-like matrix (in reference to the Fock matrix of the Roothaan-Hall equations)

$$\mathbf{F}^\alpha = \mathbf{H}^{\text{core}} + \mathbf{J} + \mathbf{F}^{\text{XC}\alpha} \tag{8.5}$$

where $\mathbf{H}^{\text{core}}$ contains the kinetic energies and interaction with the external potential, $\mathbf{J}$ contains the Coulomb repulsion and $\mathbf{F}^{\text{XC}\alpha}$ contains the effects of the exchange/correlation functional. We can expand the KS orbitals in a basis set

$$\psi_i^\alpha = \sum_\mu c_{i\mu}\phi_\mu \tag{8.6}$$

$$\rho_\alpha = \sum_{\mu\nu}\sum_i (c_{i\mu}^\alpha)^* c_{i\nu}^\alpha \phi_\mu \phi_\nu = \sum_{\mu\nu} \mathbf{P}_{\mu\nu}^\alpha \phi_\mu \phi_\nu \tag{8.7}$$

which we can then substitute into Eq. (8.3) and (8.4) to obtain an expression for the elements of the $\mathbf{F}^{\text{XC}\alpha}$ matrix

$$\mathbf{F}_{\mu\nu}^{\text{XC}\alpha} = \mathbf{P}_{\mu\nu}^\alpha \int \left[\left(\nabla^2 \frac{\delta f}{\delta(\nabla^2 \rho_\alpha)} \right) \phi_\mu \phi_\nu - \left(\nabla \frac{\delta f}{\delta(\nabla \rho_\alpha)} \right) \phi_\mu \phi_\nu + \left(\frac{\delta f}{\delta \rho_\alpha} \right) \phi_\mu \phi_\nu \right.$$
$$\left. - \left(\nabla \frac{\delta f}{\delta \tau_\alpha} \right) (\nabla \phi_\mu)\phi_\nu - \left(\frac{\delta f}{\delta \tau_\alpha} \right) (\nabla^2 \phi_\mu)\phi_\nu \right] \tag{8.8}$$

Judiciously applying integration by parts to the above simplifies this to

$$\mathbf{F}_{\mu\nu}^{\text{XC}\alpha} = \mathbf{P}_{\mu\nu}^\alpha \int \left(\frac{\delta f}{\delta \rho_\alpha} \phi_\mu \phi_\nu + \frac{\delta f}{\delta(\nabla \rho_\alpha)} \cdot \nabla(\phi_\mu \phi_\nu) \right.$$
$$\left. + \frac{\delta f}{\delta(\nabla^2 \rho_\alpha)} \nabla^2(\phi_\mu \phi_\nu) + \frac{\delta f}{\delta \tau_\alpha} (\nabla \phi_\mu . \nabla \phi_\nu) \right) \tag{8.9}$$

So far the theory developed is general and can be applied to any functional. Applying this to the GLDA functional causes difficulty because the definition of η from Chapter 6 can only be applied to same spin electrons. Higher dimensional GLDA functionals will require us to define two parameters $\eta_{\alpha\alpha}$ and $\eta_{\alpha\beta}$ as measures of the proximity of same spin electrons and opposite spin electrons respectively. The functional can then be written as

$$f(\mathbf{r}) = \rho_\alpha(\mathbf{r})\varepsilon_c^{\text{GLDA}}[\rho_\alpha(\mathbf{r}), \eta_{\alpha\alpha}(\mathbf{r}), \eta_{\alpha\beta}(\mathbf{r})] \tag{8.10}$$

The $\eta_{\alpha\alpha}$ parameter retains the same definition as η previously did. Unfortunately $\eta_{\alpha\beta}$ has yet to be defined.

The following quantities are required when Eq. (8.10) is substituted into Eq. (8.9)

$$\frac{\delta f}{\delta \rho_\alpha} = \varepsilon_c^{\text{GLDA}} + \rho_\alpha \left(\frac{\partial}{\partial \rho_\alpha} \varepsilon_c^{\text{GLDA}} + \frac{\partial \eta_{\alpha\alpha}}{\partial \rho_\alpha} \frac{\partial}{\partial \eta_{\alpha\alpha}} \varepsilon_c^{\text{GLDA}} + \frac{\partial \eta_{\alpha\beta}}{\partial \rho_\alpha} \frac{\partial}{\partial \eta_{\alpha\beta}} \varepsilon_c^{\text{GLDA}} \right) \quad (8.11)$$

$$\frac{\delta f}{\delta (\nabla \rho_\alpha)} = \rho_\alpha \left(\frac{\partial \eta_{\alpha\alpha}}{\partial (\nabla \rho_\alpha)} \frac{\partial}{\partial \eta_{\alpha\alpha}} \varepsilon_c^{\text{GLDA}} + \frac{\partial \eta_{\alpha\beta}}{\partial (\nabla \rho_\alpha)} \frac{\partial}{\partial \eta_{\alpha\beta}} \varepsilon_c^{\text{GLDA}} \right) \quad (8.12)$$

$$\frac{\delta f}{\delta (\nabla^2 \rho_\alpha)} = \rho_\alpha \left(\frac{\partial \eta_{\alpha\alpha}}{\partial (\nabla^2 \rho_\alpha)} \frac{\partial}{\partial \eta_{\alpha\alpha}} \varepsilon_c^{\text{GLDA}} + \frac{\partial \eta_{\alpha\beta}}{\partial (\nabla^2 \rho_\alpha)} \frac{\partial}{\partial \eta_{\alpha\beta}} \varepsilon_c^{\text{GLDA}} \right) \quad (8.13)$$

$$\frac{\delta f}{\delta \tau_\alpha} = \rho_\alpha \left(\frac{\partial \eta_{\alpha\alpha}}{\partial \tau_\alpha} \frac{\partial}{\partial \eta_{\alpha\alpha}} \varepsilon_c^{\text{GLDA}} + \frac{\partial \eta_{\alpha\beta}}{\partial \tau_\alpha} \frac{\partial}{\partial \eta_{\alpha\beta}} \varepsilon_c^{\text{GLDA}} \right) \quad (8.14)$$

Without a definition of $\eta_{\alpha\beta}$ the terms involving it cannot yet be computed. The remaining terms are given by

$$\frac{\partial \eta_{\alpha\alpha}}{\partial \rho_\alpha} = \mathcal{C}_d \left(\frac{\Gamma(d/2+1)}{\pi^{d/2}} \right)^{(d+2)/d}$$
$$\left(\frac{d+1}{d} |\nabla \rho_\alpha|^2 \rho_\alpha^{-(3d+2)/d} - 2 \frac{d+2}{d} \rho_\alpha^{-2(d+1)/d} \tau_\alpha \right) \quad (8.15)$$

$$\frac{\partial \eta_{\alpha\alpha}}{\partial (\nabla \rho_\alpha)} = -\mathcal{C}_d \left(\frac{\Gamma(d/2+1)}{\pi^{d/2}} \right)^{(d+2)/d} (\nabla \rho_\alpha) \rho_\alpha^{-2(d+1)/d} \quad (8.16)$$

$$\frac{\partial \eta_{\alpha\alpha}}{\partial (\nabla^2 \rho_\alpha)} = 0 \quad (8.17)$$

$$\frac{\partial \eta_{\alpha\alpha}}{\partial \tau_\alpha} = 2\mathcal{C}_d \left(\frac{\Gamma(d/2+1)}{\pi^{d/2}} \right)^{(d+2)/d} \rho_\alpha^{-(d+2)/d} \quad (8.18)$$

This is sufficiently developed for the self-consistent evaluation of the 1D GLDA functionals, although it has not been implemented.

Higher dimensional calculations will require the definition of an opposite spin η parameter, i.e. a measure of the distance between an α electron at a given point and a β electron. An unrestricted Hartree-Fock wavefunction contains two independent determinants for the α and β electrons. It appears that this results in an opposite spin parameter being impossible to obtain from a HF calculation, unlike the same spin parameter, and is the reason why this has not been done here.

8.2   The full-range GLDA functional

It is clear that one of the main drawbacks of the GLDA functional described in Chapter 6 is its lack of coverage for the full range of the η parameter. This is especially true given some of the results found in Chapter 7. We have made a number

of attempts to extend the parametrisation of the functional, but unfortunately none of them have borne fruit thus far. In this section we will give brief attention to the two most promising strategies that have been used.

8.2.1 Excited states of n-ringium

A central assumption in deriving the Hohenberg-Kohn theorems is that the ground-state wavefunction and density is used. As a result, traditional DFT cannot be used with systems in excited states. However it is not difficult to show that excited states of n-ringium have arbitrarily high values of η while maintaining their uniformity.

In Chapter 6 we found the HF orbitals of 2-ringium to be

$$\psi(\theta) = \frac{1}{\sqrt{2\pi R}} \exp\left(\frac{im\theta}{2}\right) \tag{8.19}$$

where, for the ground-state, $m = \pm 1$. In the ground-state we also find $\eta = 1/2$. If we consider "symmetrically" excited-states, where $m = \pm 3, \pm 5, \ldots$, then using a combination of Eqs. (6.6), (6.7) and (6.14) it is reasonably trivial to determine

$$\eta = \frac{3}{4}m^2 \tag{8.20}$$

given that ρ must remain the same for all states due to the finite spatial extent and the symmetry of the ring. Clearly η grows quickly as excitation levels increase, and is not difficult to determine that the same also happens for other n-ringium systems.

The hypothesis we proceeded upon is simple: by expanding Tab. 6.2 to also include an appropriately selected collection of excited states it should be possible to obtain data for $\eta > 1$ that the GLDA functional can be fitted to. This proved promising initially. In Fig. 8.1 we demonstrate that the HF energies for a wide range of excitations of 3-ringium have an all but perfectly linear relationship with their η values. This suggested that the excited state energies would paint a consistent picture of the GLDA functional.

Unfortunately, as this project continued we eventually discovered that although the excited states of a particular n-ringium are able to suggest an attractive, self-consistent picture of the GLDA functional past $\eta = 1$, this does not hold true when comparing different electron counts. That is to say one can eventually find a state of n-ringium and a state of m-ringium with equal ρ and η (or at least very similar) but wildly differing correlation energies.

As a result, this approach ultimately leads to a situation where the GLDA functional is not uniquely determined. This clearly does not lead to the correct
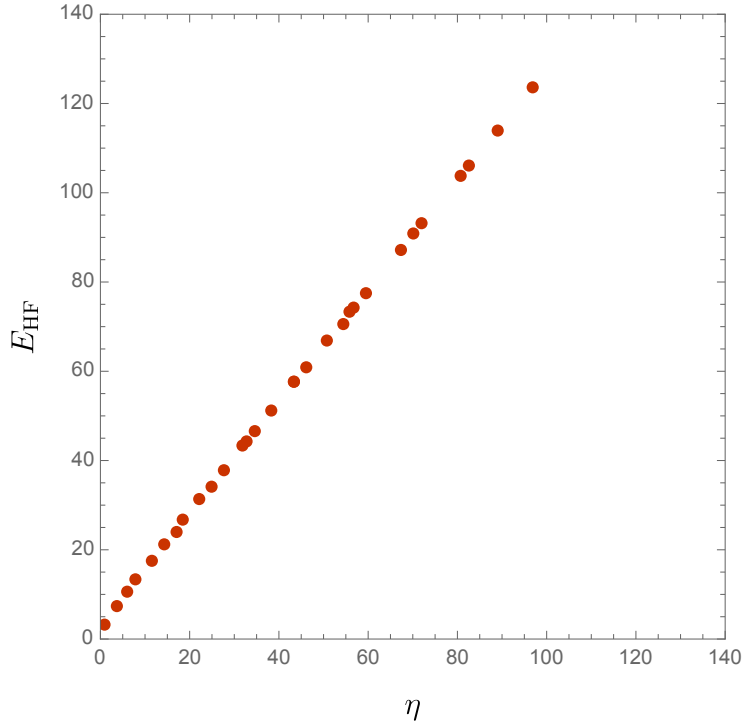
Figure 8.1:   Hartree-Fock energies (in E_h) compared to η parameters of excited states of 3-ringium with $r_s = 1$.

parametrisation, which is unsurprising given the nature of the Hohenberg-Kohn theorems and the wider understanding of mappings between external potentials and excited state densities [284].

8.2.2   Broken symmetry states of n-ringium

In Sec. 7.2.2 we briefly described the SBLDA correction of Rogers *et. al.* [144]. This correction revolves around the difference between the Fermi fluid (FF) state and the Wigner crystal (WC) state of the HF solution for the UEG. The Wigner crystal state also has connections to high values of η, which we will discuss here.

The FF state allows the electrons to delocalise to minimise their kinetic energy. In contrast, the WC state localises the electrons onto lattice sites to minimise their potential energy. This state can be found by breaking the spherical symmetry of a HF solution for the finite UEGs. This yields a solution with a non-uniform density due to the localisation of the electrons, resulting in a non-uniform η parameter which can take values greater than 1.

One could in principle use these solutions (or any non-uniform η system) to fit the full GLDA functional by solving a set of integral equations. However the non-uniform density of the WC solution is something of a mathematical oddity. The

WC should only demand that electrons localise with respect to each other, not at absolute positions on the ring. The lattice should rotate freely around the ring. If this were taken into account the spherical symmetry should be restored along with the uniformity of ρ and η.

Naïvely, one can attempt to do this by averaging the density

$$\rho = \int_0^{2\pi} \frac{\rho_{\mathrm{SB}}(\theta)}{2\pi} d\theta \tag{8.21}$$

where ρ is the re-symmetrised density and ρ_{SB} is the symmetry-broken WC density. This does indeed return the density to the expected uniform value ($\rho = n/2\pi R$), however complications arise from the subtle dependency of the η parameter on the density through Eq. (6.7). The averaging process does not commute with the evaluation of η, and it is unclear what the correct value should be. All other attempts that have been made to re-symmetrise the solution have met with the same difficulty.

If one were able to overcome this problem then the WC solutions show some promise for constructing the GLDA functional past $\eta = 1$. The methods that we have attempted thus far suggest that the WC solutions do indeed take uniform η values larger than 1. Unfortunately they cannot form a complete data set for fitting the GLDA functional since the WC can only form at sufficiently low densities (the point at which they do form depends upon the number of electrons). This is most clear in the extreme case of the high-density limit where the WC and FF states are always degenerate.

Chapter 9

Concluding remarks

In this thesis we have explored two very distinct lines of research. In Chapters 2 through 5 we explored a new model of one-dimensional (1D) chemistry and ultimately described the basics of molecular structure within it. In Chapter 6 we constructed the Generalised Local Density Approximation (GLDA) and laid some groundwork for its further development in Chapter 8. We were also able to explore the intersection of these ideas in Chapter 7, where we tested the efficacy of the GLDA in modelling reactions.

Until this point studying 1D molecules has been the domain of physicists, being completely ignored by chemists since the early work of Loudon [69]. This is understandable given the conclusions he reached, where under his assumptions 1D chemistry is almost certainly an exceedingly dull topic. More recent developments show that the situation is far more complicated than earlier researchers realised, and a different approach reveals a far more curious set of results. Dealing with the confinement using the Dirichlet conditions produces a surprisingly nuanced model with a number of very unique features.

While we do freely admit that this may not be the most applicable model to real-life situations (there are exceptions e.g. the unique bonds formed in extremely large magnetic fields [155, 156]), the construction presented here offers possibilities for some unique theoretical insight. A specific example of this, and one which has been encountered in this thesis (See Sec. 7.2.3), is the possibility for new insight into dispersion models. This arises due to the ability to isolate electrons from one another, something which cannot be done in 3D molecules.

In contrast, the GLDA is a new approach to a long-standing body of literature which is only increasing in popularity. Most modern functional development builds upon its immediate predecessors. While this idea is at the heart of all scientific study, contemporary work has also turned to extensive parameterisation against reference

(experimental or theoretical) data. This is a pragmatic approach, but it slows the development of our understanding of the fundamental physics at play.

As we pointed out in Sec. 8.1, one can consider the GLDA to be a high level meta-GGA functional due to its use of the kinetic energy density. But the method of its construction builds directly upon the LDA. This makes its development in some sense orthogonal to the directions that density functional development has taken so far. This gives us some unique insights into the behaviour of electrons, and it may form the basis for other new developments.

Apart from that we observe the GLDA to be a promisingly successful improvement over the LDA in 1D. Although our work stops there it lays most of the necessary foundation for constructing the same functional in 3D. Furthermore, our observations in Sec. 7.3.2 suggest that it might be possible that the GLDA functional performs even better in 3D than it does in 1D.

Appendix A

G1D correlation energies

The G1D test set is subset of the molecules studied in Chapter 5. It is a collection of 42 molecules which are intended to contain a range of the different bond characters observed in 1D. The set includes diatomic, triatomic and tetra-atomic molecules with a variety of aligned and opposed dipole bonds, noble atom bonds. It also includes a set of weakly bound molecules. The full list of molecules is found in Tab. A.1 and correlation estimates in Tab. A.3.

The set also includes seven different reactions for each of the molecules (where the reaction is possible). Using $_1H_1He_1$ as an example, the following reactions are included (the table in which the energies are listed is given in brackets):

- Atomisation energy $\qquad\qquad _1H_1He_1 \quad \rightarrow \quad H_1 + {}_1He_1 \qquad$ (Tab. A.4)
- Left ionisation energy $\qquad\quad _1H_1He_1 \quad \rightarrow \quad H_1He_1 + e^- \quad$ (Tab. A.5)
- Right ionisation energy $\qquad _1H_1He_1 \quad \rightarrow \quad {}_1H_1He + e^- \quad$ (Tab. A.6)
- Left electron affinity $\qquad _1H_1He_1 + e^- \quad \rightarrow \quad {}_2H_1He_1 \qquad$ (Tab. A.7)
- Right electron affinity $\qquad _1H_1He_1 + e^- \quad \rightarrow \quad {}_1H_1He_2 \qquad$ (Tab. A.8)
- Left proton affinity $\qquad _1H_1He_1 + H^+ \quad \rightarrow \quad H_1H_1He_1{}^+ \qquad$ (Tab. A.9)
- Right proton affinity $\qquad _1H_1He_1 + H^+ \quad \rightarrow \quad {}_1H_1He_1H^+ \quad$ (Tab. A.10)

Geometries for each of the required species have been optimised at the Hartree-Fock level. We also report correlation energy estimates for the first ten atoms in Tab. A.2.

Diatomics	Triatomics	Tetra-atomics
H_1H_1	$H_2Li_3Li_2$	$_2B_3H_1H_3B_3$
$_1H_1He_1$	$H_3B_5B_3$	$_1H_3B_5B_3H$
H_2Li_2	$H_3B_4Li_2$	$_1H_3B_2H_3B_3$
$_1H_2Li_1$	$_1H_3B_4Li_1$	
$_1H_2Be_2$	$_1Li_2H_3B_3$	Weakly bound species
H_3B_3	$_1Li_3Li_3Li_2$	$H_2H_2B_3$
$_1H_3B_2$	$_1Li_3Li_4B_3$	$H_2He_3B_3$
$_1H_2B_3$	$_2Li_2H_2B_3$	$_1H_1He_4C_3$
$_1H_3C_3$	$_1H_2Li_3Be_2$	$_1He_2He_2Li_2$
$_1He_2Li_2$	$_2Li_3Li_3Be_2$	$_2Li_2He_3Be_2$
$_1Li_3Li_2$	$_2Be_2H_3B_3$	$_2Li_2He_4B_2$
$_2Li_3Be_2$	$_1H_3C_3H_1$	$_2B_4H_2B_3$
$_2Li_4B_2$	$_2Li_3Be_3Li_2$	$_3B_3He_4C_3$
$_2Li_5N_3$		$_2B_5Be_4B_3$
$_2Li_4N_4$		$_3B_2H_4B_3H$
$_2B_5B_3$		

Table A.1: Molecules contained in the G1D test set.

	Total correlation		Intra-domain correlation							
Atom	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1	0	0	0	0	-19.054	-15.024	0	0	0	0
$_1He_1$	-2.063	-2.688	0	0	-43.024	-38.826	0	0	0	0
$_1Li_2$	-3.364	-4.032	-1.106	-1.264	-61.192	-52.189	-0.687	-1.982	-3.250	-2.323
$_2B_3$	-9.576	-10.853	-6.852	-7.650	-101.400	-86.571	-3.216	-9.442	-15.545	-13.540
$_2Be_2$	-6.666	-7.682	-4.081	-4.561	-84.109	-72.360	-2.053	-6.199	-10.263	-8.756
$_3C_3$	-14.337	-16.261	-11.464	-12.926	-123.047	-104.571	-5.008	-14.518	-23.840	-20.680
$_3N_4$	-18.718	-21.152	-15.747	-17.751	-140.643	-119.138	-6.554	-18.645	-30.496	-27.255
$_4F_5$	-29.740	-33.740	-26.608	-30.224	-180.962	-152.178	-10.384	-28.806	-46.864	-42.013
$_4O_4$	-24.082	-27.278	-21.020	-23.807	-160.508	-135.075	-8.553	-23.810	-38.766	-34.429
$_5Ne_5$	-35.289	-40.056	-32.095	-36.497	-197.955	-166.015	-12.386	-33.747	-54.685	-49.378

Table A.2: Correlation energy estimates (in mE_h) for 1D atoms.

Atom	Total correlation		Intra-domain correlation							
	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	−0.846	−1.156	0	0	−39.380	−31.699	0	0	0	0
$H_2H_2B_3$	−10.316	−11.869	−6.815	−7.604	−141.309	−118.045	−3.190	−9.409	−15.506	−13.416
$H_2He_3B_3$	−11.711	−13.597	−6.924	−7.707	−163.904	−139.914	−3.207	−9.480	−15.628	−13.566
H_2Li_2	−3.751	−4.489	−1.475	−1.688	−80.566	−66.915	−1.106	−3.090	−5.034	−3.383
$H_2Li_3Li_2$	−7.595	−9.102	−2.998	−3.424	−143.048	−120.009	−2.250	−6.254	−10.179	−6.816
H_3B_3	−9.670	−10.937	−6.947	−7.736	−120.654	−101.125	−3.360	−9.668	−15.852	−13.649
$H_3B_4Li_2$	−13.508	−15.532	−8.423	−9.385	−184.421	−155.376	−4.386	−12.696	−20.842	−16.988
$H_3B_5B_3$	−19.493	−22.079	−14.035	−15.653	−222.926	−187.864	−6.804	−19.671	−32.282	−27.546
$_1H_1He_1$	−2.305	−2.987	0	0	−62.754	−54.300	0	0	0	0
$_1H_1He_4C_3$	−16.751	−19.343	−11.574	−13.021	−186.242	−158.576	−4.996	−14.638	−24.090	−20.778
$_1H_2B_3$	−10.329	−11.885	−6.826	−7.617	−122.232	−103.093	−3.194	−9.421	−15.525	−13.430
$_1H_2Be_2$	−7.522	−8.898	−3.951	−4.407	−105.164	−89.677	−2.032	−6.103	−10.093	−8.596
$_1H_2Li_1$	−4.583	−5.805	−1.026	−1.172	−83.661	−71.386	−0.689	−1.955	−3.196	−2.300
$_1H_2Li_3Be_2$	−11.620	−13.911	−5.463	−6.130	−168.516	−143.917	−3.148	−9.205	−15.142	−12.014
$_1H_3B_2$	−10.958	−12.874	−6.678	−7.397	−126.145	−107.826	−3.165	−9.270	−15.254	−13.385
$_1H_3B_4Li_1$	−14.557	−17.179	−8.004	−8.909	−187.785	−160.033	−4.049	−11.715	−19.230	−15.985
$_1H_3B_5B_3H_0$	−21.133	−24.416	−14.102	−15.692	−248.510	−210.484	−6.859	−19.745	−32.375	−27.650
$_1H_3C_3H_1$	−16.549	−19.434	−10.935	−12.153	−168.166	−143.175	−4.858	−14.010	−22.980	−20.005
$_1H_3C_3$	−15.442	−17.843	−11.203	−12.544	−145.806	−123.722	−4.910	−14.255	−23.414	−20.347
$_1He_2He_2Li_2$	−7.629	−9.573	−1.250	−1.432	−147.679	−129.332	−0.977	−2.659	−4.309	−2.894
$_1Li_2H_3B_3$	−14.293	−16.790	−8.028	−8.977	−185.070	−157.466	−4.069	−11.793	−19.364	−16.059
$_1Li_3Li_2$	−7.178	−8.611	−2.598	−2.964	−123.689	−104.904	−1.810	−5.103	−8.330	−5.699
$_1Li_3Li_3Li_2$	−11.016	−13.210	−4.127	−4.707	−186.314	−158.166	−2.964	−8.278	−13.486	−9.141
$_1Li_3Li_4B_3$	−16.989	−19.730	−9.687	−10.881	−225.395	−191.226	−5.292	−15.099	−24.711	−19.563
$_2B_4H_2B_3$	−19.869	−22.701	−13.647	−15.238	−224.163	−189.595	−6.428	−18.941	−31.205	−26.998
$_2B_5B_3$	−19.385	−21.976	−13.926	−15.550	−203.894	−172.965	−6.613	−19.400	−31.934	−27.411
$_2B_5Be_4B_3$	−25.863	−29.443	−17.832	−19.918	−287.814	−244.754	−8.598	−25.355	−41.779	−35.990
$_2Be_2H_3B_3$	−17.236	−19.891	−10.943	−12.199	−206.724	−175.862	−5.391	−15.875	−26.151	−22.330
$_2Li_2H_2B_3$	−14.150	−16.462	−8.392	−9.419	−184.177	−155.505	−4.315	−12.579	−20.679	−16.905
$_2Li_2He_1$	−5.565	−6.884	−1.249	−1.431	−104.530	−90.637	−0.967	−2.652	−4.303	−2.890
$_2Li_2He_3Be_2$	−12.262	−14.590	−5.364	−6.021	−188.870	−162.731	−3.073	−8.938	−14.688	−11.695
$_2Li_2He_4B_2$	−15.176	−17.771	−8.141	−9.120	−206.601	−177.113	−4.252	−12.337	−20.261	−16.600
$_2Li_3Be_2$	−10.378	−12.114	−5.518	−6.193	−146.013	−124.280	−3.139	−9.204	−15.150	−11.996
$_2Li_3Be_3Li_2$	−14.072	−16.527	−6.933	−7.800	−207.730	−176.501	−4.173	−12.145	−19.960	−15.186
$_2Li_3Li_3Be_2$	−14.209	−16.710	−7.035	−7.922	−208.472	−177.351	−4.315	−12.382	−20.289	−15.419
$_2Li_4B_2$	−13.393	−15.421	−8.308	−9.274	−165.332	−140.427	−4.289	−12.471	−20.490	−16.849
$_2Li_4N_4$	−22.621	−25.746	−17.348	−19.503	−203.015	−171.141	−7.569	−21.585	−35.323	−30.375
$_2Li_5N_3$	−22.555	−25.738	−17.198	−19.352	−205.572	−173.433	−7.577	−21.613	−35.372	−30.561
$_3B_2H_4B_3H_0$	−20.045	−22.880	−13.827	−15.422	−243.544	−204.850	−6.546	−19.257	−31.717	−27.211
$_3B_3He_4C_3$	−26.115	−29.911	−18.456	−20.686	−268.231	−229.193	−8.180	−24.025	−39.556	−34.273

Table A.3: Correlation energy estimates (in mE_h) for 1D molecules in the G1D test set.

	Total correlation		Intra-domain correlation							
Atom	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	0.846	1.156	0	0	1.272	1.651	0	0	0	0
$H_2H_2B_3$	0.741	1.016	−0.037	−0.047	1.802	1.425	−0.026	−0.033	−0.039	−0.124
$H_2He_3B_3$	0.072	0.056	0.073	0.057	0.427	−0.507	−0.008	0.038	0.083	0.026
H_2Li_2	0.387	0.457	0.369	0.425	0.321	−0.298	0.418	1.108	1.784	1.060
$H_2Li_3Li_2$	0.867	1.037	0.786	0.897	1.610	0.606	0.875	2.291	3.679	2.171
H_3B_3	0.094	0.084	0.096	0.086	0.201	−0.470	0.144	0.226	0.307	0.109
$H_3B_4Li_2$	0.569	0.647	0.466	0.472	2.775	1.592	0.483	1.273	2.047	1.125
$H_3B_5B_3$	0.342	0.372	0.332	0.353	1.073	−0.303	0.372	0.787	1.193	0.465
$_1H_1He_1$	0.242	0.299	0	0	0.676	0.450	0	0	0	0
$_1H_1He_4C_3$	0.351	0.393	0.110	0.096	1.118	0.154	−0.012	0.120	0.250	0.098
$_1H_2B_3$	0.753	1.032	−0.025	−0.033	1.778	1.498	−0.022	−0.021	−0.020	−0.111
$_1H_2Be_2$	0.856	1.216	−0.130	−0.155	2.002	2.292	−0.021	−0.096	−0.170	−0.160
$_1H_2Li_1$	1.219	1.773	−0.080	−0.092	3.415	4.173	0.001	−0.027	−0.054	−0.023
$_1H_2Li_3Be_2$	1.590	2.197	0.275	0.305	4.161	4.344	0.407	1.024	1.629	0.935
$_1H_3B_2$	1.383	2.021	−0.173	−0.253	5.691	6.231	−0.051	−0.172	−0.291	−0.155
$_1H_3B_4Li_1$	1.618	2.294	0.047	−0.005	6.139	6.249	0.146	0.292	0.435	0.122
$_1H_3B_5B_3H_0$	1.982	2.710	0.399	0.392	7.603	7.294	0.427	0.861	1.286	0.569
$_1H_3C_3H_1$	2.212	3.173	−0.528	−0.772	7.011	8.555	−0.149	−0.508	−0.860	−0.674
$_1H_3C_3$	1.105	1.582	−0.261	−0.382	3.706	4.127	−0.097	−0.263	−0.426	−0.333
$_1He_2He_2Li_2$	0.139	0.164	0.144	0.168	0.440	−0.509	0.289	0.678	1.059	0.571
$_1Li_2H_3B_3$	1.354	1.904	0.071	0.063	3.424	3.682	0.165	0.369	0.569	0.196
$_1Li_3Li_2$	0.450	0.546	0.387	0.437	1.305	0.526	0.435	1.139	1.830	1.054
$_1Li_3Li_3Li_2$	0.925	1.113	0.809	0.916	2.738	1.599	0.902	2.333	3.736	2.173
$_1Li_3Li_4B_3$	0.686	0.812	0.623	0.703	1.611	0.277	0.701	1.694	2.666	1.378
$_2B_4H_2B_3$	0.718	0.995	−0.056	−0.063	2.309	1.428	−0.003	0.057	0.116	−0.083
$_2B_5B_3$	0.233	0.270	0.223	0.249	1.094	−0.178	0.181	0.516	0.845	0.331
$_2B_5Be_4B_3$	0.046	0.055	0.048	0.056	0.905	−0.749	0.113	0.272	0.428	0.154
$_2Be_2H_3B_3$	0.995	1.356	0.010	−0.013	2.162	1.906	0.122	0.234	0.344	0.034
$_2Li_2H_2B_3$	1.211	1.577	0.434	0.505	2.531	1.721	0.412	1.155	1.884	1.042
$_2Li_2He_1$	0.138	0.163	0.143	0.167	0.314	−0.378	0.280	0.670	1.053	0.568
$_2Li_2He_3Be_2$	0.169	0.187	0.177	0.195	0.546	−0.644	0.332	0.758	1.175	0.617
$_2Li_2He_4B_2$	0.173	0.197	0.183	0.206	0.986	−0.474	0.348	0.913	1.467	0.737
$_2Li_3Be_2$	0.348	0.400	0.331	0.368	0.713	−0.269	0.398	1.024	1.637	0.917
$_2Li_3Be_3Li_2$	0.679	0.780	0.641	0.711	1.238	−0.238	0.745	1.983	3.197	1.785
$_2Li_3Li_3Be_2$	0.816	0.963	0.742	0.833	1.980	0.612	0.886	2.220	3.526	2.018
$_2Li_4B_2$	0.453	0.536	0.350	0.360	2.740	1.666	0.386	1.047	1.695	0.986
$_2Li_4N_4$	0.539	0.561	0.496	0.488	1.181	−0.186	0.327	0.958	1.576	0.797
$_2Li_5N_3$	0.473	0.554	0.346	0.337	3.738	2.106	0.335	0.987	1.626	0.983
$_3B_2H_4B_3H_0$	0.894	1.174	0.124	0.121	2.636	1.659	0.114	0.373	0.627	0.131
$_3B_3He_4C_3$	0.140	0.109	0.140	0.110	0.760	−0.775	−0.044	0.065	0.171	0.053

Table A.4: Correlation contribution to atomisation energies (in mE_h) for 1D molecules in the G1D test set.

Atom	Total correlation		Intra-domain correlation							
	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1										
$H_2H_2B_3$										
$H_2He_3B_3$										
H_2Li_2										
$H_2Li_3Li_2$										
H_3B_3										
$H_3B_4Li_2$										
$H_3B_5B_3$										
$_1H_1He_1$	0.592	0.785	0	0	18.673	14.194	0	0	0	0
$_1H_1He_4C_3$	0.780	1.010	0.183	0.218	19.302	15.026	−0.101	−0.002	0.095	0.194
$_1H_2B_3$	0.646	0.878	−0.200	−0.285	17.587	13.310	−0.101	−0.252	−0.401	−0.201
$_1H_2Be_2$	1.388	1.863	0.296	0.327	18.063	13.079	0.096	0.298	0.497	0.510
$_1H_2Li_1$	1.703	2.331	0.380	0.432	17.618	12.091	0.160	0.478	0.790	0.698
$_1H_2Li_3Be_2$	1.699	2.337	0.353	0.401	17.547	12.001	0.164	0.443	0.716	0.642
$_1H_3B_2$	1.948	2.752	0.382	0.461	16.810	10.708	0.224	0.524	0.818	0.524
$_1H_3B_4Li_1$	1.938	2.742	0.356	0.426	16.928	10.885	0.120	0.450	0.773	0.523
$_1H_3B_5B_3H_0$	1.913	2.721	0.321	0.388	16.772	10.648	0.277	0.499	0.716	0.444
$_1H_3C_3H_1$	2.209	2.990	0.778	0.919	17.645	12.030	0.353	0.847	1.330	1.095
$_1H_3C_3$	2.111	2.860	0.707	0.830	17.329	11.663	0.259	0.760	1.250	1.003
$_1He_2He_2Li_2$	2.008	2.624	−0.052	−0.060	20.821	18.189	0.002	−0.047	−0.094	−0.092
$_1Li_2H_3B_3$	1.547	2.043	−0.857	−0.961	21.015	18.769	−0.158	−0.665	−1.162	−1.506
$_1Li_3Li_2$	1.331	1.745	−0.951	−1.065	20.662	18.073	−0.289	−0.923	−1.545	−1.686
$_1Li_3Li_3Li_2$	1.677	2.143	−0.605	−0.669	21.681	19.415	−0.076	−0.361	−0.639	−0.992
$_1Li_3Li_4B_3$	1.588	2.042	−0.697	−0.775	21.079	18.744	−0.138	−0.454	−0.763	−1.149
$_2B_4H_2B_3$	1.896	2.066	1.660	1.725	17.550	13.476	0.959	2.902	4.807	4.450
$_2B_5B_3$	2.166	2.401	1.922	2.045	18.177	14.168	0.983	3.177	5.327	4.798
$_2B_5Be_4B_3$	1.930	2.101	1.691	1.756	17.442	13.391	0.926	2.868	4.771	4.454
$_2Be_2H_3B_3$	2.393	2.679	2.198	2.387	17.503	11.986	1.126	3.208	5.249	4.169
$_2Li_2H_2B_3$	0.848	0.973	0.811	0.913	14.423	7.672	0.548	1.628	2.687	1.864
$_2Li_2He_1$	1.093	1.259	1.070	1.223	14.977	8.182	0.683	1.937	3.167	2.248
$_2Li_2He_3Be_2$	1.096	1.262	1.073	1.227	14.996	8.218	0.705	1.942	3.154	2.247
$_2Li_2He_4B_2$	1.225	1.427	1.198	1.384	15.359	8.597	0.668	1.973	3.252	2.329
$_2Li_3Be_2$	1.042	1.216	0.985	1.124	14.741	7.945	0.636	1.796	2.932	2.073
$_2Li_3Be_3Li_2$	1.051	1.226	0.995	1.136	14.720	7.922	0.592	1.792	2.968	2.095
$_2Li_3Li_3Be_2$	0.937	1.116	0.852	0.974	14.229	7.253	0.662	1.694	2.707	1.843
$_2Li_4B_2$	0.945	1.162	0.834	0.973	13.526	6.197	0.612	1.644	2.656	1.650
$_2Li_4N_4$	0.986	1.133	0.921	1.027	14.379	7.800	0.567	1.647	2.705	2.033
$_2Li_5N_3$	0.793	0.999	0.663	0.777	12.767	5.293	0.651	1.508	2.349	1.362
$_3B_2H_4B_3H_0$	1.126	1.379	1.108	1.350	12.298	5.052	0.489	1.483	2.457	1.447
$_3B_3He_4C_3$	0.931	1.139	0.925	1.128	11.727	4.485	0.458	1.320	2.166	1.216

Table A.5: Correlation contribution to left side ionisation energies (in mE_h) for 1D molecules in the G1D test set. Blank entries indicate that the molecule has no possible left ionisation.

Atom	Total correlation		Intra-domain correlation							
	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	0.846	1.156	0	0	18.212	13.367	0	0	0	0
$H_2H_2B_3$	1.092	1.337	1.077	1.312	12.160	4.886	0.597	1.529	2.443	1.411
$H_2He_3B_3$	0.911	1.116	0.905	1.105	11.683	4.427	0.538	1.354	2.153	1.193
H_2Li_2	0.898	1.045	0.855	0.978	14.570	7.694	0.577	1.644	2.691	1.835
$H_2Li_3Li_2$	0.924	1.101	0.839	0.959	14.220	7.245	0.574	1.645	2.694	1.823
H_3B_3	0.871	1.067	0.864	1.055	11.579	4.354	0.599	1.343	2.073	1.139
$H_3B_4Li_2$	0.928	1.143	0.816	0.952	13.493	6.164	0.595	1.623	2.631	1.625
$H_3B_5B_3$	0.734	0.907	0.718	0.877	11.078	3.907	0.439	1.119	1.786	0.958
$_1H_1He_1$	1.851	2.392	0	0	20.982	18.299	0	0	0	0
$_1H_1He_4C_3$	1.814	2.195	1.775	2.125	14.307	7.227	0.710	2.285	3.829	2.565
$_1H_2B_3$	1.106	1.353	1.090	1.328	12.187	4.901	0.601	1.543	2.466	1.428
$_1H_2Be_2$	2.008	2.322	1.821	2.039	17.757	12.224	0.954	2.893	4.794	4.018
$_1H_2Li_1$	1.856	2.349	−0.493	−0.544	22.682	21.350	−0.166	−0.549	−0.924	−0.924
$_1H_2Li_3Be_2$	2.070	2.401	1.879	2.107	17.855	12.319	1.027	3.060	5.053	4.174
$_1H_3B_2$	2.251	2.552	1.913	2.047	19.061	14.943	0.963	2.989	4.975	4.690
$_1H_3B_4Li_1$	1.394	1.822	−0.865	−0.949	20.563	17.892	−0.193	−0.681	−1.159	−1.591
$_1H_3B_5B_3H_0$										
$_1H_3C_3H_1$	2.209	2.991	0.779	0.919	17.645	12.031	0.354	0.847	1.331	1.096
$_1H_3C_3$	2.089	2.511	2.052	2.454	14.961	8.172	0.911	2.533	4.122	2.923
$_1He_2He_2Li_2$	1.091	1.257	1.069	1.222	14.975	8.179	0.691	1.940	3.164	2.246
$_1Li_2H_3B_3$	0.772	0.943	0.774	0.944	11.262	4.089	0.441	1.179	1.902	1.034
$_1Li_3Li_2$	0.965	1.145	0.883	1.010	14.318	7.363	0.600	1.698	2.773	1.903
$_1Li_3Li_3Li_2$	0.913	1.091	0.824	0.943	14.095	7.068	0.614	1.653	2.672	1.799
$_1Li_3Li_4B_3$	0.751	0.919	0.740	0.899	11.298	4.198	0.513	1.183	1.841	1.031
$_2B_4H_2B_3$	1.152	1.411	1.135	1.383	12.443	5.170	0.646	1.579	2.494	1.471
$_2B_5B_3$	0.752	0.929	0.736	0.899	11.129	3.951	0.551	1.185	1.807	0.974
$_2B_5Be_4B_3$	0.937	1.147	0.930	1.134	11.840	4.608	0.617	1.396	2.159	1.221
$_2Be_2H_3B_3$	0.834	1.020	0.833	1.017	11.494	4.277	0.498	1.260	2.007	1.096
$_2Li_2H_2B_3$	1.185	1.452	1.171	1.429	12.356	5.032	0.632	1.628	2.605	1.516
$_2Li_2He_1$										
$_2Li_2He_3Be_2$	1.865	2.169	1.675	1.874	17.295	11.553	0.913	2.785	4.621	3.800
$_2Li_2He_4B_2$	1.957	2.145	1.717	1.798	17.888	13.696	0.946	2.992	4.997	4.563
$_2Li_3Be_2$	1.961	2.274	1.777	1.989	17.374	11.689	0.950	2.924	4.859	3.989
$_2Li_3Be_3Li_2$	1.051	1.226	0.995	1.136	14.720	7.922	0.592	1.792	2.968	2.095
$_2Li_3Li_3Be_2$	2.028	2.352	1.843	2.066	17.617	12.020	1.045	3.040	4.995	4.110
$_2Li_4B_2$	2.506	2.805	2.254	2.438	19.810	16.161	1.132	3.468	5.758	5.333
$_2Li_4N_4$	0.666	0.844	0.662	0.838	10.046	3.049	0.349	0.876	1.392	0.707
$_2Li_5N_3$	3.163	3.726	3.108	3.638	17.446	11.607	1.230	3.500	5.725	4.726
$_3B_2H_4B_3H_0$										
$_3B_3He_4C_3$	1.748	2.113	1.709	2.044	14.035	6.987	0.771	2.239	3.678	2.466

Table A.6: Correlation contribution to right side ionisation energies (in mE_h) for 1D molecules in the G1D test set. Blank entries indicate that the molecule has no possible right ionisation.

Atom	Total correlation		Intra-domain correlation							
	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	−1.077	−1.519	0	0	−16.750	−10.849	0	0	0	0
$H_2H_2B_3$	−1.601	−2.389	0.106	0.130	−15.609	−9.266	0.006	0.088	0.167	0.129
$H_2He_3B_3$	−1.612	−2.402	0.097	0.119	−15.560	−9.216	0.013	0.067	0.119	0.100
H_2Li_2	−0.983	−1.508	0.348	0.404	−11.576	−6.164	0.208	0.616	1.016	0.665
$H_2Li_3Li_2$	−1.614	−2.259	−0.230	−0.262	−16.842	−11.014	−0.102	−0.323	−0.540	−0.441
H_3B_3	−1.562	−2.319	0.092	0.115	−15.045	−8.930	0.052	0.007	−0.036	0.050
$H_3B_4Li_2$	−1.781	−2.586	−0.128	−0.154	−16.245	−9.994	−0.112	−0.281	−0.446	−0.227
$H_3B_5B_3$	−1.725	−2.518	−0.074	−0.088	−15.892	−9.687	−0.141	−0.247	−0.351	−0.150
$_1H_1He_1$	−0.023	−0.031	−0.024	−0.032	−9.263	−2.314	−0.065	−0.143	−0.220	−0.072
$_1H_1He_4C_3$	−0.039	−0.049	−0.039	−0.048	−9.389	−2.434	−0.195	−0.203	−0.211	−0.091
$_1H_2B_3$	0.057	0.065	0.054	0.063	−8.287	−1.537	−0.080	−0.109	−0.137	0.036
$_1H_2Be_2$	−0.016	−0.020	−0.024	−0.032	−9.404	−2.397	−0.070	−0.113	−0.155	−0.077
$_1H_2Li_1$	−0.021	−0.023	−0.033	−0.042	−9.816	−2.726	−0.058	−0.110	−0.161	−0.075
$_1H_2Li_3Be_2$	−0.016	−0.017	−0.029	−0.037	−9.852	−2.752	−0.043	−0.098	−0.150	−0.066
$_1H_3B_2$	−0.023	−0.025	−0.039	−0.049	−10.043	−2.878	0.014	−0.087	−0.185	−0.082
$_1H_3B_4Li_1$	−0.025	−0.025	−0.040	−0.050	−10.161	−3.002	−0.026	−0.107	−0.185	−0.093
$_1H_3B_5B_3H_0$	−0.025	−0.027	−0.040	−0.051	−10.207	−3.028	0.038	−0.073	−0.182	−0.089
$_1H_3C_3H_1$	−0.031	−0.036	−0.043	−0.054	−9.620	−2.564	−0.066	−0.118	−0.169	−0.080
$_1H_3C_3$	0.009	0.013	−0.005	−0.009	−9.537	−2.445	−0.032	−0.073	−0.114	−0.025
$_1He_2He_2Li_2$	−0.006	−0.009	−0.007	−0.009	−8.225	−1.555	−0.070	−0.145	−0.218	−0.081
$_1Li_2H_3B_3$	−0.613	−0.743	−0.513	−0.580	−12.603	−5.371	−0.491	−1.299	−2.092	−1.258
$_1Li_3Li_2$	−0.725	−0.879	−0.630	−0.720	−13.004	−5.719	−0.520	−1.412	−2.287	−1.443
$_1Li_3Li_3Li_2$	−0.795	−0.957	−0.703	−0.804	−13.428	−6.236	−0.504	−1.482	−2.441	−1.579
$_1Li_3Li_4B_3$	−0.748	−0.902	−0.657	−0.749	−13.039	−5.882	−0.476	−1.429	−2.364	−1.506
$_2B_4H_2B_3$	−0.331	−0.419	−0.311	−0.380	−9.442	−2.327	−0.359	−0.747	−1.126	−0.462
$_2B_5B_3$	−0.503	−0.627	−0.484	−0.591	−9.930	−2.860	−0.419	−0.907	−1.385	−0.672
$_2B_5Be_4B_3$	−0.331	−0.420	−0.307	−0.375	−9.235	−2.166	−0.312	−0.711	−1.103	−0.460
$_2Be_2H_3B_3$	−0.077	−0.099	−0.073	−0.092	−8.832	−1.974	−0.093	−0.228	−0.361	−0.126
$_2Li_2H_2B_3$	−0.132	−0.168	−0.135	−0.171	−10.580	−3.453	−0.141	−0.286	−0.428	−0.203
$_2Li_2He_1$	−0.192	−0.240	−0.193	−0.241	−10.858	−3.712	−0.130	−0.328	−0.522	−0.287
$_2Li_2He_3Be_2$	−0.194	−0.242	−0.195	−0.243	−10.889	−3.745	−0.136	−0.333	−0.527	−0.290
$_2Li_2He_4B_2$	−0.243	−0.301	−0.243	−0.302	−11.187	−4.055	−0.171	−0.387	−0.599	−0.346
$_2Li_3Be_2$	−0.197	−0.245	−0.198	−0.248	−10.989	−3.844	−0.137	−0.348	−0.554	−0.293
$_2Li_3Be_3Li_2$	−0.189	−0.237	−0.190	−0.239	−10.813	−3.659	−0.188	−0.357	−0.523	−0.271
$_2Li_3Li_3Be_2$	−0.209	−0.258	−0.213	−0.266	−11.368	−4.218	−0.098	−0.376	−0.649	−0.346
$_2Li_4B_2$	−0.231	−0.285	−0.240	−0.301	−11.492	−4.292	−0.222	−0.457	−0.688	−0.383
$_2Li_4N_4$	−0.128	−0.158	−0.131	−0.162	−10.423	−3.353	−0.104	−0.243	−0.379	−0.214
$_2Li_5N_3$	−0.216	−0.268	−0.228	−0.289	−11.533	−4.296	−0.216	−0.434	−0.648	−0.362
$_3B_2H_4B_3H_0$	−0.735	−0.931	−0.732	−0.926	−13.591	−6.351	−0.412	−0.914	−1.405	−0.867
$_3B_3He_4C_3$	−0.717	−0.910	−0.717	−0.910	−13.644	−6.367	−0.441	−0.942	−1.433	−0.835

Table A.7: Correlation contribution to left side electron affinities (in mE_h) for 1D molecules in the G1D test set.

	Total correlation		Intra-domain correlation							
Atom	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	−0.017	−0.022	−0.021	−0.028	−9.539	−2.525	−0.062	−0.100	−0.136	−0.065
$H_2H_2B_3$	−0.667	−0.848	−0.664	−0.842	−13.197	−5.957	−0.339	−0.827	−1.304	−0.797
$H_2He_3B_3$	−0.704	−0.896	−0.705	−0.896	−13.610	−6.329	−0.387	−0.911	−1.424	−0.821
H_2Li_2	−0.195	−0.244	−0.196	−0.245	−11.082	−3.941	−0.144	−0.358	−0.568	−0.296
$H_2Li_3Li_2$	−0.208	−0.257	−0.213	−0.265	−11.410	−4.264	−0.157	−0.407	−0.651	−0.354
H_3B_3	−0.716	−0.911	−0.716	−0.911	−13.741	−6.456	−0.332	−0.897	−1.451	−0.833
$H_3B_4Li_2$	−0.229	−0.283	−0.238	−0.299	−11.525	−4.328	−0.268	−0.476	−0.679	−0.380
$H_3B_5B_3$	−0.781	−0.990	−0.785	−0.996	−14.351	−7.015	−0.391	−1.010	−1.617	−0.921
$_1H_1He_1$	−0.059	−0.073	−0.055	−0.066	−8.152	−1.531	−0.136	−0.281	−0.424	−0.193
$_1H_1He_4C_3$	−0.301	−0.382	−0.302	−0.385	−10.919	−3.713	−0.299	−0.496	−0.689	−0.439
$_1H_2B_3$	−0.665	−0.844	−0.661	−0.839	−13.147	−5.903	−0.382	−0.843	−1.295	−0.796
$_1H_2Be_2$	−0.111	−0.139	−0.107	−0.133	−9.248	−2.313	−0.089	−0.248	−0.405	−0.173
$_1H_2Li_1$	−0.177	−0.215	−0.140	−0.156	−8.564	−1.842	−0.219	−0.563	−0.899	−0.491
$_1H_2Li_3Be_2$	−0.125	−0.158	−0.123	−0.154	−9.167	−2.256	−0.146	−0.302	−0.455	−0.190
$_1H_3B_2$	−0.501	−0.616	−0.492	−0.601	−9.903	−2.886	−0.334	−0.854	−1.364	−0.679
$_1H_3B_4Li_1$	−0.817	−1.015	−0.696	−0.810	−13.038	−5.604	−0.588	−1.523	−2.439	−1.450
$_1H_3B_5B_3H_0$	−1.812	−2.625	−0.158	−0.192	−16.370	−10.092	−0.151	−0.339	−0.524	−0.262
$_1H_3C_3H_1$	−0.031	−0.036	−0.043	−0.054	−9.620	−2.564	−0.066	−0.118	−0.169	−0.080
$_1H_3C_3$	−0.362	−0.457	−0.358	−0.452	−11.049	−3.904	−0.255	−0.505	−0.750	−0.514
$_1He_2He_2Li_2$	−0.192	−0.240	−0.193	−0.241	−10.861	−3.715	−0.122	−0.324	−0.523	−0.287
$_1Li_2H_3B_3$	−0.742	−0.942	−0.747	−0.949	−14.074	−6.758	−0.399	−0.980	−1.550	−0.874
$_1Li_3Li_2$	−0.212	−0.262	−0.216	−0.269	−11.335	−4.187	−0.157	−0.408	−0.654	−0.347
$_1Li_3Li_3Li_2$	−0.212	−0.262	−0.218	−0.271	−11.492	−4.340	−0.164	−0.413	−0.657	−0.374
$_1Li_3Li_4B_3$	−0.756	−0.957	−0.759	−0.961	−14.266	−6.971	−0.376	−0.981	−1.575	−0.907
$_2B_4H_2B_3$	−0.735	−0.931	−0.732	−0.926	−13.573	−6.326	−0.291	−0.856	−1.411	−0.873
$_2B_5B_3$	−0.782	−0.991	−0.786	−0.997	−14.311	−6.979	−0.383	−1.003	−1.611	−0.921
$_2B_5Be_4B_3$	−0.753	−0.954	−0.753	−0.955	−13.898	−6.611	−0.493	−1.003	−1.504	−0.883
$_2Be_2H_3B_3$	−0.717	−0.911	−0.721	−0.916	−13.810	−6.515	−0.355	−0.917	−1.467	−0.834
$_2Li_2H_2B_3$	−0.622	−0.792	−0.620	−0.789	−12.752	−5.503	−0.389	−0.808	−1.219	−0.768
$_2Li_2He_1$	−0.024	−0.035	−0.019	−0.024	−8.125	−1.404	−0.118	−0.252	−0.383	−0.125
$_2Li_2He_3Be_2$	−0.075	−0.097	−0.076	−0.098	−9.210	−2.223	−0.071	−0.218	−0.363	−0.130
$_2Li_2He_4B_2$	−0.394	−0.495	−0.373	−0.456	−9.819	−2.627	−0.433	−0.853	−1.264	−0.562
$_2Li_3Be_2$	−0.096	−0.123	−0.095	−0.122	−9.040	−2.093	−0.093	−0.248	−0.400	−0.142
$_2Li_3Be_3Li_2$	−0.189	−0.237	−0.190	−0.239	−10.813	−3.659	−0.188	−0.357	−0.523	−0.271
$_2Li_3Li_3Be_2$	−0.118	−0.151	−0.118	−0.149	−9.028	−2.113	−0.068	−0.258	−0.445	−0.170
$_2Li_4B_2$	−0.536	−0.663	−0.521	−0.635	−10.177	−3.091	−0.365	−0.907	−1.440	−0.728
$_2Li_4N_4$	−1.489	−1.923	−1.488	−1.921	−16.247	−8.915	−0.647	−1.711	−2.753	−1.557
$_2Li_5N_3$	−0.400	−0.504	−0.397	−0.501	−9.601	−2.702	−0.183	−0.550	−0.911	−0.456
$_3B_2H_4B_3H_0$	−1.619	−2.389	0.032	0.040	−15.756	−9.457	−0.152	−0.186	−0.218	−0.020
$_3B_3He_4C_3$	−0.265	−0.340	−0.267	−0.342	−10.721	−3.535	−0.284	−0.458	−0.628	−0.395

Table A.8: Correlation contribution to right side electron affinities (in mE_h) for 1D molecules in the G1D test set.

	Total correlation		Intra-domain correlation							
Atom	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1										
$H_2H_2B_3$	−0.063	−0.078	−0.061	−0.074	−0.291	−0.342	−0.020	−0.062	−0.103	−0.077
$H_2He_3B_3$	−0.059	−0.072	−0.058	−0.071	−0.326	−0.375	−0.053	−0.078	−0.102	−0.073
H_2Li_2										
$H_2Li_3Li_2$										
H_3B_3	−0.112	−0.138	−0.112	−0.137	−0.557	−0.597	0.001	−0.093	−0.185	−0.136
$H_3B_4Li_2$	−0.031	−0.037	−0.032	−0.039	−0.163	−0.217	−0.070	−0.056	−0.042	−0.031
$H_3B_5B_3$	−0.075	−0.091	−0.075	−0.092	−0.402	−0.443	0.007	−0.051	−0.107	−0.086
$_1H_1He_1$	0.150	0.205	0	0	−2.326	−3.826	0	0	0	0
$_1H_1He_4C_3$	0.294	0.378	0.135	0.161	−1.932	−3.319	−0.083	−0.002	0.078	0.147
$_1H_2B_3$	−0.025	0.006	−0.276	−0.354	−3.149	−4.393	−0.109	−0.298	−0.484	−0.327
$_1H_2Be_2$	0.515	0.716	0.135	0.149	−2.730	−4.500	0.048	0.138	0.225	0.235
$_1H_2Li_1$	0.691	1.005	0.145	0.164	−3.256	−5.542	0.058	0.173	0.286	0.26
$_1H_2Li_3Be_2$	0.782	1.113	0.232	0.263	−3.064	−5.190	0.090	0.278	0.462	0.427
$_1H_3B_2$	0.923	1.380	0.188	0.225	−3.725	−6.354	0.096	0.237	0.375	0.251
$_1H_3B_4Li_1$	0.932	1.395	0.183	0.217	−3.616	−6.195	0.102	0.231	0.357	0.268
$_1H_3B_5B_3H_0$	0.925	1.394	0.167	0.201	−3.732	−6.374	0.073	0.204	0.332	0.221
$_1H_3C_3H_1$	1.068	1.487	0.463	0.545	−2.958	−5.084	0.188	0.477	0.76	0.658
$_1H_3C_3$	1.001	1.394	0.419	0.489	−3.202	−5.359	0.171	0.448	0.718	0.604
$_1He_2He_2Li_2$	0.301	0.429	−0.045	−0.052	−1.131	−1.670	−0.003	−0.042	−0.08	−0.079
$_1Li_2H_3B_3$										
$_1Li_3Li_2$										
$_1Li_3Li_3Li_2$										
$_1Li_3Li_4B_3$	−0.094	−0.109	−0.095	−0.111	−0.473	−0.526	−0.018	−0.085	−0.151	−0.149
$_2B_4H_2B_3$	−0.130	−0.180	−0.193	−0.273	−2.807	−3.378	−0.025	−0.168	−0.307	−0.085
$_2B_5B_3$	0.100	0.104	0.029	−0.002	−2.251	−2.759	0.001	0.05	0.098	0.201
$_2B_5Be_4B_3$	−0.109	−0.158	−0.176	−0.256	−2.929	−3.488	−0.043	−0.196	−0.346	−0.089
$_2Be_2H_3B_3$	0.980	1.070	0.874	0.910	−2.765	−4.339	0.322	0.774	1.218	0.931
$_2Li_2H_2B_3$	0.261	0.297	0.226	0.244	−5.287	−7.958	0.084	0.251	0.415	0.37
$_2Li_2He_1$	0.458	0.529	0.432	0.492	−4.758	−7.418	0.151	0.471	0.784	0.667
$_2Li_2He_3Be_2$	0.460	0.531	0.435	0.494	−4.753	−7.403	0.150	0.465	0.774	0.665
$_2Li_2He_4B_2$	0.571	0.673	0.543	0.630	−4.474	−7.124	0.133	0.504	0.867	0.742
$_2Li_3Be_2$	0.412	0.487	0.362	0.411	−5.039	−7.767	0.130	0.373	0.612	0.529
$_2Li_3Be_3Li_2$	0.416	0.491	0.367	0.417	−5.044	−7.765	0.066	0.353	0.635	0.543
$_2Li_3Li_3Be_2$	0.335	0.414	0.264	0.298	−5.558	−8.467	0.108	0.273	0.435	0.356
$_2Li_4B_2$	0.353	0.464	0.258	0.305	−6.083	−9.242	0.061	0.235	0.405	0.238
$_2Li_4N_4$	0.330	0.375	0.275	0.288	−5.326	−7.841	−0.012	0.191	0.39	0.469
$_2Li_5N_3$	0.258	0.367	0.146	0.177	−6.691	−9.978	0.150	0.178	0.205	0.039
$_3B_2H_4B_3H_0$	0.746	0.951	0.729	0.924	−7.106	−10.262	0.178	0.787	1.383	0.902
$_3B_3He_4C_3$	0.563	0.727	0.554	0.710	−7.662	−10.836	0.164	0.633	1.093	0.678

Table A.9: Correlation contribution to left side proton affinities (in mE_h) for 1D molecules in the G1D test set. Blank entries indicate that the geometry optimisation could not find a stable proton position.

	Total correlation		Intra-domain correlation							
Atom	MP2	MP3	Intra-MP2	Intra-MP3	LDA	SBLDA	0LDA	αLDA	gLDA	SBgLDA
H_1H_1	0.264	0.378	0	0	−2.687	−4.449	0	0	0	0
$H_2H_2B_3$	0.725	0.925	0.710	0.901	−7.172	−10.343	0.337	0.865	1.383	0.881
$H_2He_3B_3$	0.547	0.708	0.538	0.692	−7.695	−10.879	0.237	0.664	1.082	0.659
H_2Li_2	0.398	0.466	0.360	0.410	−5.028	−7.731	0.134	0.395	0.652	0.542
$H_2Li_3Li_2$	0.325	0.404	0.254	0.287	−5.570	−8.478	0.066	0.248	0.425	0.342
H_3B_3	0.506	0.658	0.496	0.641	−7.819	−10.986	0.340	0.673	1.000	0.602
$H_3B_4Li_2$	0.341	0.451	0.245	0.290	−6.113	−9.272	0.073	0.233	0.390	0.221
$H_3B_5B_3$	0.379	0.510	0.359	0.474	−8.356	−11.488	0.121	0.431	0.736	0.435
$_1H_1He_1$	0.226	0.308	0	0	−0.994	−1.549	0	0	0	0
$_1H_1He_4C_3$	1.018	1.294	0.981	1.229	−5.001	−7.775	0.258	1.020	1.767	1.382
$_1H_2B_3$	0.738	0.941	0.723	0.917	−7.140	−10.320	0.340	0.878	1.406	0.897
$_1H_2Be_2$	0.585	0.702	0.483	0.548	−2.545	−4.164	0.159	0.454	0.744	0.759
$_1H_2Li_1$										
$_1H_2Li_3Be_2$	0.645	0.784	0.526	0.597	−2.471	−4.106	0.169	0.559	0.941	0.872
$_1H_3B_2$	0.158	0.211	0.013	−0.012	−1.629	−2.274	−0.007	−0.089	−0.170	0.102
$_1H_3B_4Li_1$	−0.053	−0.057	−0.053	−0.056	−0.260	−0.380	−0.023	−0.047	−0.070	−0.118
$_1H_3B_5B_3H_0$	−0.016	−0.020	−0.017	−0.020	−0.108	−0.174	0.091	0.041	−0.007	−0.018
$_1H_3C_3H_1$	1.068	1.487	0.463	0.545	−2.958	−5.084	0.188	0.477	0.760	0.658
$_1H_3C_3$	1.102	1.400	1.047	1.310	−4.714	−7.420	0.406	1.092	1.765	1.438
$_1He_2He_2Li_2$	0.457	0.527	0.432	0.491	−4.761	−7.421	0.159	0.474	0.782	0.665
$_1Li_2H_3B_3$	0.410	0.538	0.406	0.530	−8.166	−11.304	0.144	0.490	0.830	0.495
$_1Li_3Li_2$	0.356	0.436	0.287	0.325	−5.475	−8.362	0.090	0.290	0.487	0.402
$_1Li_3Li_3Li_2$	0.318	0.397	0.244	0.275	−5.688	−8.642	0.117	0.264	0.407	0.324
$_1Li_3Li_4B_3$	0.391	0.517	0.376	0.491	−8.172	−11.253	0.198	0.489	0.774	0.487
$_2B_4H_2B_3$	0.771	0.982	0.755	0.955	−6.971	−10.147	0.337	0.884	1.420	0.927
$_2B_5B_3$	0.395	0.528	0.375	0.493	−8.308	−11.446	0.110	0.434	0.751	0.448
$_2B_5Be_4B_3$	0.565	0.729	0.555	0.711	−7.586	−10.760	0.220	0.654	1.080	0.676
$_2Be_2H_3B_3$	0.472	0.615	0.466	0.604	−7.914	−11.081	0.226	0.586	0.940	0.563
$_2Li_2H_2B_3$	0.821	1.044	0.807	1.021	−6.928	−10.123	0.298	0.928	1.545	0.990
$_2Li_2He_1$										
$_2Li_2He_3Be_2$	0.460	0.567	0.355	0.403	−2.969	−4.805	0.120	0.348	0.572	0.554
$_2Li_2He_4B_2$	−0.066	−0.099	−0.133	−0.197	−2.474	−3.155	−0.110	−0.117	−0.124	0.029
$_2Li_3Be_2$	0.503	0.612	0.405	0.459	−2.923	−4.699	0.099	0.392	0.678	0.642
$_2Li_3Be_3Li_2$	0.416	0.491	0.367	0.417	−5.044	−7.765	0.066	0.353	0.635	0.543
$_2Li_3Li_3Be_2$	0.551	0.668	0.453	0.514	−2.737	−4.442	0.202	0.492	0.776	0.729
$_2Li_4B_2$	−0.010	−0.017	−0.091	−0.140	−2.106	−2.738	−0.055	−0.175	−0.293	0.006
$_2Li_4N_4$	0.334	0.518	0.331	0.512	−9.281	−12.228	0.322	0.671	1.013	0.501
$_2Li_5N_3$	0.893	1.146	0.844	1.063	−3.439	−5.190	0.216	0.720	1.213	1.476
$_3B_2H_4B_3H_0$	−0.114	−0.140	−0.112	−0.137	−0.433	−0.507	−0.053	−0.100	−0.146	−0.128
$_3B_3He_4C_3$	0.690	0.905	0.652	0.840	−5.545	−8.427	0.106	0.616	1.116	0.881

Table A.10: Correlation contribution to right side proton affinities (in mE_h) for 1D molecules in the G1D test set. Blank entries indicate that the geometry optimisation could not find a stable proton position.

Bibliography

[1] E. Schrödinger, *Phys. Rev.* **28**, 1049 (1926).

[2] E. Schrödinger, *Ann. Phys.* **384**, 361 (1926).

[3] W. Pauli, *Z. Phys.* **31**, 765 (1925).

[4] M. Born and R. Oppenheimer, *Ann. Phys.* **389**, 457 (1927).

[5] D. R. Hartree, *Math. Proc. Cambridge Philos. Soc.* **24**, 89 (1928).

[6] V. Fock, *Z. Phys.* **61**, 126 (1930).

[7] J. C. Slater, *Phys. Rev.* **34**, 1293 (1929).

[8] A. Szabo and N. S. Ostlund, *Modern quantum chemistry* (McGraw-Hill, New York, 1989).

[9] C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).

[10] G. G. Hall, *Proc. R. Soc. Lond. A* **205**, 541 (1951).

[11] J. A. Pople and R. K. Nesbet, *J. Chem. Phys.* **22**, 571 (1954).

[12] P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980).

[13] P. Pulay, *J. Comp. Chem.* **3**, 556 (1982).

[14] C. Møller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).

[15] J. W. S. Rayleigh, *Theory of Sound* (Macmillan, London, 1894).

[16] P. M. W. Gill, J. A. Pople, L. Radom and R. H. Nobes, *J. Chem. Phys.* **89**, 7307 (1988).

[17] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).

[18] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).

[19] P. F. Loos and P. M. W. Gill, *WIREs Comput. Mol. Sci.* **6**, 410 (2016).

[20] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).

[21] A. D. Becke, *J. Chem. Phys.* **140**, 18A301 (2014).

[22] A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).

[23] A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

[24] J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**, 13244 (1992).

[25] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).

[26] R. Saito, G. Dresselhauss and M. S. Dresselhaus, *Properties of Carbon Nanotubes* (Imperial College Press, London, 1998).

[27] R. Egger and A. O. Gogolin, *Eur. Phys. J. B* **3**, 281 (1998).

[28] M. Bockrath, D. H. Cobden, J. Lu, A. G. Rinzler, R. E. Smalley, L. Balents and P. L. McEuen, *Nature* **397**, 598 (1999).

[29] H. Ishii, H. Kataura, H. Shiozawa, H. Yoshioka, H. Otsubo, Y. Takayama, T. Miyahara, S. Suzuki, Y. Achiba, M. Nakatake, T. Narimura, M. Higashiguchi, K. Shimada, H. Namatame and M. Taniguchi, *Nature* **426**, 540 (2003).

[30] M. Shiraishi and M. Ata, *Sol. State Commun.* **127**, 215 (2003).

[31] A. Schwartz, M. Dressel, G. Grüner, V. Vescoli, L. Degiorgi and T. Giamarchi, *Phys. Rev. B* **58**, 1261 (1998).

[32] V. Vescoli, F. Zwick, W. Henderson, L. Degiorgi, M. Grioni, G. Gruner and L. K. Montgomery, *Eur. Phys. J. B* **13**, 503 (2000).

[33] T. Lorenz, M. Hofmann, M. Grüninger, A. Freimuth, G. S. Uhrig, M. Dumm and M. Dressel, *Nature* **418**, 614 (2002).

[34] M. Dressel, K. Petukhov, B. Salameh, P. Zornoza and T. Giamarchi, *Phys. Rev. B* **71**, 075104 (2005).

[35] T. Ito, A. Chainani, T. Haruna, K. Kanai, T. Yokoya, S. Shin and R. Kato, *Phys. Rev. Lett.* **95**, 246402 (2005).

[36] Z. Hu, M. Knupfer, M. Kielwein, U. K. Rößler, M. S. Golden, J. Fink, F. M. F. de Groot, T. Ito, K. Oka and G. Kaindl, *Eur. Phys. J. B* **26**, 449 (2002).

[37] F. P. Milliken, C. P. Umbach and R. A. Webb, *Sol. State Commun.* **97**, 309 (1996).

[38] S. S. Mandal and J. K. Jain, *Sol. State Commun.* **118**, 503 (2001).

[39] A. M. Chang, *Rev. Mod. Phys.* **75**, 1449 (2003).

[40] A. R. Goñi, A. Pinczuk, J. S. Weiner, J. M. Calleja, B. S. Dennis, L. N. Pfeiffer and K. W. West, *Phys. Rev. Lett.* **67**, 3298 (1991).

[41] O. M. Auslaender, A. Yacoby, R. de Picciotto, K. W. Baldwin, L. N. Pfeiffer and K. W. West, *Phys. Rev. Lett.* **84**, 1764 (2000).

[42] S. V. Zaitsev-Zotov, Y. A. Kumzerov, Y. A. Firsov and P. Monceau, *J. Phys. Condens. Matter* **12**, L303 (2000).

[43] F. Liu, M. Bao, K. L. Wang, C. Li, B. Lei and C. Zhou, *Appl. Phys. Lett.* **86**, 213101 (2005).

[44] H. Steinberg, O. M. Auslaender, A. Yacoby, J. Qian, G. A. Fiete, Y. Tserkovnyak, B. I. Halperin, K. W. Baldwin, L. N. Pfeiffer and K. W. West, *Phys. Rev. B* **73**, 113307 (2006).

[45] H. Monien, M. Linn and N. Elstner, *Phys. Rev. A* **58**, R3395 (1998).

[46] A. Recati, P. O. Fedichev, W. Zwerger and P. Zoller, *J. Opt. B: Quantum Semiclass. Opt.* **5**, S55 (2003).

[47] H. Moritz, T. Stöferle, K. Günter, M. Köhl and T. Esslinger, *Phys. Rev. Lett.* **94**, 210401 (2005).

[48] L. O. Wagner, E. M. Stoudenmire, K. Burke and S. R. White, *Phys. Chem. Chem. Phys.* **14**, 8581 (2012).

[49] E. M. Stoudenmire, L. O. Wagner, S. R. White and K. Burke, *Phys. Rev. Lett.* **109**, 056402 (2012).

[50] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules* (Oxford, Clarendon Press, 1989).

[51] D. J. Doren and D. R. Herschbach, *Chem. Phys. Lett.* **118**, 115 (1985).

[52] J. G. Loeser and D. R. Herschbach, *J. Chem. Phys.* **84**, 3882 (1986).

[53] J. G. Loeser and D. R. Herschbach, *J. Chem. Phys.* **84**, 3893 (1986).

[54] C. M. Rosenthal, *J. Chem. Phys.* **55**, 2474 (1971).

[55] L. L. Foldy, *Am. J. Phys.* **44**, 1192 (1976).

[56] L. L. Foldy, *Am. J. Phys.* **45**, 1230 (1977).

[57] Y. Nogami, M. Vallieres and W. van Dijk, *Am. J. Phys.* **44**, 886 (1976).

[58] J. G. Loeser and D. R. Herschbach, *J. Phys. Chem.* **89**, 3444 (1985).

[59] D. R. Herschbach, *J. Chem. Phys.* **84**, 838 (1986).

[60] D. J. Doren and D. R. Herschbach, *J. Chem. Phys.* **87**, 433 (1987).

[61] J. G. Loeser and D. R. Herschbach, *J. Chem. Phys.* **86**, 3512 (1987).

[62] G. E. Astrakharchik and M. D. Girardeau, *Phys. Rev. B* **83**, 153303 (2011).

[63] R. M. Lee and N. D. Drummond, *Phys. Rev. B* **83**, 245114 (2011).

[64] P. F. Loos and P. M. W. Gill, *Phys. Rev. Lett.* **108**, 083002 (2012).

[65] P. F. Loos, *J. Chem. Phys.* **138**, 064108 (2013).

[66] P. F. Loos and P. M. W. Gill, *J. Chem. Phys.* **138**, 164124 (2013).

[67] P. F. Loos, C. J. Ball and P. M. W. Gill, *J. Chem. Phys.* **140**, 18A524 (2014).

[68] P. F. Loos, *Phys. Rev. A* **89**, 052523 (2014).

[69] R. Loudon, *Am. J. Phys.* **27**, 649 (1959).

[70] K. Burnett, V. C. Reed and P. L. Knight, *J. Phys. B* **26**, 561 (1993).

[71] M. Mayle, B. Hezel, I. Lesanovsky and P. Schmelcher, *Phys. Rev. Lett.* **99**, 113004 (2007).

[72] M. M. Nieto, *Phys. Rev. A* **61**, 034901 (2000).

[73] S. H. Patil, *Phys. Rev. A* **64**, 064902 (2001).

[74] P. M. Platzman and M. I. Dykman, *Science* **284**, 1967 (1999).

[75] M. I. Dykman, P. M. Platzman and P. Seddighrad, *Phys. Rev. B* **67**, 155402 (2003).

[76] M. Andrews, *Am. J. Phys.* **44**, 1064 (1976).

[77] H. N. Núñez-Yépez, C. A. Varga and A. L. Salas-Brito, *Eur. J. Phys.* **8**, 189 (1987).

[78] L. J. Boya, M. Kmiecik and A. Bohm, *Phys. Rev. A* **37**, 3567 (1988).

[79] V. S. Mineev, *Theor. Math. Phys.* **140**, 1157 (2004).

[80] C. R. de Oliveira and A. A. Verri, *Ann. Phys.* **324**, 251 (2009).

[81] G. Abramovici and Y. Avishai, *J. Phys. A* **42**, 285302 (2009).

[82] H. N. Núñez-Yépez, A. L. Salas-Brito and D. A. Solis, *Phys. Rev. A* **83**, 064101 (2011).

[83] T. D. Imbo and U. P. Sukhatme, *Phys. Rev. Lett.* **54**, 2184 (1985).

[84] H. N. Núñez-Yépez and A. L. S. Brito, *Eur. J. Phys.* **8**, 307 (1987).

[85] H. N. Núñez-Yépez, C. A. Varga and A. L. Salas-Brito, *Phys. Rev. A* **39**, 4306 (1989).

[86] M. Moshinsky, *J. Phys. A* **26**, 2445 (1993).

[87] R. G. Newton, *J. Phys. A* **27**, 4717 (1994).

[88] D. Xianxi, J. Dai and J. Dai, *Phys. Rev. A* **55**, 2617 (1997).

[89] K. Connolly and D. J. Griffiths, *Am. J. Phys.* **75**, 524 (2007).

[90] W. Fischer, H. Leschke and P. Müller, *J. Math. Phys.* **36**, 2313 (1995).

[91] P. Kurasov, *J. Phys. A* **29**, 1767 (1996).

[92] W. Fischer, H. Leschke and P. Müller, *J. Phys. A* **30**, 5579 (1997).

[93] C. R. de Oliveira, *Phys. Lett. A* **374**, 2805 (2010).

[94] C. R. de Oliveira and A. A. Verri, *J. Math. Phys.* **53**, 052104 (2012).

[95] L. Mitas, *Phys. Rev. Lett.* **96**, 240402 (2006).

[96] E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).

[97] E. A. Hylleraas, *Adv. Quantum Chem.* **1**, 1 (1964).

[98] H. M. James and A. S. Coolidge, *J. Chem. Phys.* **1**, 825 (1933).

[99] Wolfram Research, Inc., *Mathematica* (Wolfram Research, Inc., Champaign, Illinois, 2015), version 10.2.

[100] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark, editors, *NIST handbook of mathematical functions* (Cambridge University Press, New York, 2010).

[101] J. D. Baker, D. E. Freund, R. Nyden Hill and J. D. Morgan III, *Phys. Rev. A* **41**, 1247 (1990).

[102] P. F. Loos and P. M. W. Gill, *J. Chem. Phys.* **131**, 241101 (2009).

[103] E. A. Hylleraas, *Z. Phys.* **65**, 209 (1930).

[104] J. Linderberg and H. Shull, *J. Mol. Spectrosc.* **5**, 1 (1960).

[105] J. Linderberg, *Phys. Rev.* **121**, 816 (1961).

[106] P. F. Loos and P. M. W. Gill, *Phys. Rev. Lett.* **105**, 113001 (2010).

[107] P. F. Loos and P. M. W. Gill, *Chem. Phys. Lett.* **500**, 1 (2010).

[108] K. R. Lykke, K. K. Murray and W. C. Lineberger, *Phys. Rev. A* **43**, 6104 (1991).

[109] G. Haeffler, D. Hanstorp, I. Kiyan, A. E. Klinkmüller, U. Ljungblad and D. J. Pegg, *Phys. Rev. A* **53**, 4127 (1996).

[110] H. Hotop and W. C. Lineberger, *J. Phys. Chem. Ref. Data* **14**, 731 (1985).

[111] K. T. Andersson, J. Sandstrom, I. Y. Kiyan, D. Hanstorp and D. J. Pegg, *Phys. Rev. A* **62**, 022503 (2000).

[112] P. Frey, F. Breyer and H. Hotop, *J. Phys. B* **11**, L589 (1978).

[113] D. R. Lide, editor, *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, Florida, 2003), chapter 10, 84th edition.

[114] E. Clementi and D. L. Raimond, *J. Chem. Phys.* **38**, 2686 (1963).

[115] E. Clementi, D. L. Raimond and W. P. Reinhardt, *J. Chem. Phys.* **47** (1967).

[116] O. Burrau, *Naturwissenschaften* **15**, 16 (1927).

[117] L. Pauling, *Chem. Rev.* **5**, 173 (1928).

[118] L. Pauling and E. B. Wilson, *Introduction to Quantum Mechnics with Applications to Chemistry* (Dover Publications, Mineola, New York, 1985).

[119] T. C. Scott, M. Aubert-Frecon and J. Grotendorst, *Chem. Phys.* **324**, 323 (2006).

[120] A. V. Turbiner, J. C. Lopez-Vieyra and N. L. Guevara, *Phys. Rev. A* **81**, 042503 (2010).

[121] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure: IV. Constants of diatomic molecules* (Van Nostrand Reinhold Company, New York, 1979).

[122] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **49**, 404 (1968).

[123] L. Pauling, *J. Chem. Phys.* **1**, 56 (1933).

[124] M. Guilhaus, A. Brenton, J. Beynon, M. Rabrenovic and P. R. Schleyer, *J. Phys. B* **17**, L605 (1984).

[125] P. M. W. Gill and L. Radom, *Chem. Phys. Lett.* **136**, 294 (1987).

[126] P. M. W. Gill and L. Radom, *J. Am. Chem. Soc.* **110**, 5311 (1988).

[127] P. M. W. Gill and L. Radom, *Chem. Phys. Lett.* **147**, 213 (1988).

[128] L. Wolniewicz, *J. Chem. Phys.* **43**, 1087 (1965).

[129] J. J. Thomson, *Philos. Mag.* **21**, 225 (1911).

[130] E. Herbst, *Phil. Trans. R. Soc. A* **358**, 2523 (2000).

[131] T. Oka, *Proc. Nat. Am. Soc.* **103**, 12235 (2006).

[132] T. Oka, *Chem. Rev.* **113**, 8738 (2013).

[133] F. Mentch and J. B. Anderson, *J. Chem. Phys.* **80**, 2675 (1984).

[134] C. A. Coulson, *Proc. Cam. Philos. Soc.* **31**, 244 (1935).

[135] G. D. Carney and R. N. Porter, *J. Chem. Phys.* **65**, 3547 (1976).

[136] M. Kertesz, J. Koller and A. Azman, *Theoret. Chim. Acta* **41**, 89 (1976).

[137] J. Hachmann, W. Cardoen and G. K. L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).

[138] T. Tsuchimochi and G. E. Scuseria, *J. Chem. Phys.* **131**, 121102 (2009).

[139] A. V. Sinitskiy, L. Greenman and D. A. Mazziotti, *J. Chem. Phys.* **133**, 014104 (2010).

[140] L. Stella, C. Attaccalite, S. Sorella and A. Rubio, *Phys. Rev. B* **84**, 245117 (2011).

[141] D. Zgid and G. K. L. Chan, *J. Chem. Phys.* **134**, 094115 (2011).

[142] V. R. Saunders, *Faraday Symp. Chem. Soc.* **19**, 79 (1984).

[143] V. R. Saunders, C. Freyria-Fava, R. Dovesi and C. Roetti, *Comput. Phys. Commun.* **84**, 156 (1994).

[144] F. J. M. Rogers, C. J. Ball and P. F. Loos, *Phys. Rev. B* **93**, 235114 (2016).

[145] P. M. W. Gill and L. Radom, *Chem. Phys. Lett.* **132**, 16 (1986).

[146] R. H. Nobes, D. Moncrieff, M. W. Wong, L. Radom, P. M. W. Gill and J. A. Pople, *Chem. Phys. Lett.* **182**, 216 (1991).

[147] A. Deaño, J. Segura and N. M. Temme, *Math. Comput.* **77**, 2277 (2008).

[148] J. Abad and J. Sesma, *J. Comput. Appl. Math.* **78**, 97 (1997).

[149] J. Abad and J. Sesma, *J. Comput. Appl. Math.* **101**, 237 (1999).

[150] W. G. Bickley, L. J. Comrie, J. C. P. Miller, D. H. Sadler and A. J. Thompson, *Bessel Functions. Part II: Functions of Positive Integer Order*, British Association for the Advancement of Science, Mathematical Tables, Volume 10 (Cambridge University Press, Cambridge, 1952).

[151] P. M. W. Gill, B. G. Johnson and J. A. Pople, *Chem. Phys. Lett.* **217**, 65 (1994).

[152] B. Paredes, A. Widera, V. Murg, O. Mandel, S. Folling, I. Cirac, G. V. Shlyapnikov, T. W. Hansch and I. Bloch, *Nature* **429**, 277 (2004).

[153] E. Haller, M. Gustavsson, M. J. Mark, J. G. Danzl, R. Hart, G. Pupillo and H. C. Nägerl, *Science* **325**, 1224 (2009).

[154] P. Schmelcher and L. S. Cederbaum, *Phys. Rev. A* **41**, 4936 (1990).

[155] K. K. Lange, E. I. Tellgren, M. R. Hoffmann and T. Helgaker, *Science* **337**, 327 (2012).

[156] P. Schmelcher, *Science* **337**, 302 (2012).

[157] P. Schmelcher and L. S. Cederbaum, *Int. J. Quantum Chem.* **64**, 501 (1997).

[158] E. I. Tellgren, A. Soncini and T. Helgaker, *J. Chem. Phys.* **129**, 154114 (2008).

[159] E. I. Tellgren, T. Helgaker and A. Soncini, *Phys. Chem. Chem. Phys.* **11**, 5489 (2009).

[160] S. Boblest, C. Schimeczek and G. Wunner, *Phys. Rev. A* **89**, 012505 (2014).

[161] S. Stopkowicz, J. Gauss, K. K. Lange, E. I. Tellgren and T. Helgaker, *J. Chem. Phys.* **143**, 074110 (2015).

[162] B. G. Johnson, P. M. W. Gill, J. A. Pople and D. J. Fox, *Chem. Phys. Lett.* **206**, 239 (1993).

[163] F. London, *Z. Physik. Chem. B* **11**, 222 (1930).

[164] F. London, *Trans. Faraday Soc.* **33**, 8 (1937).

[165] K. S. Pitzer, *J. Chem. Phys.* **23**, 1735 (1955).

[166] A. J. Stone, *The theory of intermolecular forces* (Clarendon Press, Oxford, 1997).

[167] E. Fermi, *Z. Phys.* **36**, 902 (1926).

[168] L. H. Thomas, *Proc. Cam. Phil. Soc.* **23**, 542 (1927).

[169] E. Wigner, *Phys. Rev.* **46**, 1002 (1934).

[170] W. Macke, *Z. Naturforsch. A* **5a**, 192 (1950).

[171] D. Bohm and D. Pines, *Phys. Rev.* **92**, 609 (1953).

[172] D. Pines, *Phys. Rev.* **92**, 626 (1953).

[173] M. Gell-Mann and K. A. Brueckner, *Phys. Rev.* **106**, 364 (1957).

[174] D. F. DuBois, *Ann. Phys.* **7**, 174 (1959).

[175] W. J. Carr, Jr. and A. A. Maradudin, *Phys. Rev.* **133**, A371 (1964).

[176] S. Misawa, *Phys. Rev.* **140**, A1645 (1965).

[177] L. Onsager, L. Mittag and M. J. Stephen, *Ann. Phys.* **18**, 71 (1966).

[178] A. Isihara and D. Y. Kojima, *Z. Phys. B* **21**, 33 (1975).

[179] D. Y. Kojima and A. Isihara, *Z. Phys. B* **25**, 167 (1976).

[180] Y. Wang and J. P. Perdew, *Phys. Rev. B* **43**, 8911 (1991).

[181] G. G. Hoffman, *Phys. Rev. B* **45**, 8730 (1992).

[182] T. Endo, M. Horiuchi, Y. Takada and H. Yasuhara, *Phys. Rev. B* **59**, 7367 (1999).

[183] P. Ziesche and J. Cioslowski, *Physica A* **356**, 598 (2005).

[184] P. F. Loos and P. M. W. Gill, *Phys. Rev. B* **84**, 033103 (2011).

[185] R. K. P. Zia, *J. Phys. C* **6**, 3121 (1973).

[186] M. L. Glasser, *J. Phys. C* **10**, L121 (1977).

[187] A. K. Rajagopal and J. C. Kimball, *Phys. Rev. B* **15**, 2819 (1977).

[188] A. Isihara and T. Toyoda, *Ann. Phys.* **106**, 394 (1977).

[189] A. Isihara and T. Toyoda, *Ann. Phys.* **114**, 497 (1978).

[190] A. Isihara and L. Ioriatti, *Phys. Rev. B* **22**, 214 (1980).

[191] M. L. Glasser, *J. Comp. App. Math.* **10**, 293 (1984).

[192] M. Seidl, *Phys. Rev. B* **70**, 073101 (2004).

[193] S. Chesi and G. F. Giuliani, *Phys. Rev. B* **75**, 153306 (2007).

[194] P. F. Loos and P. M. W. Gill, *Phys. Rev. B* **83**, 233102 (2011).

[195] M. M. Fogler, *Phys. Rev. Lett.* **94**, 056405 (2005).

[196] R. A. Coldwell-Horsfall and A. A. Maradudin, *J. Math. Phys.* **1**, 395 (1960).

[197] W. J. Carr, Jr., *Phys. Rev.* **122**, 1437 (1961).

[198] W. J. Carr, Jr., R. A. Coldwell-Horsfall and A. E. Fein, *Phys. Rev.* **124**, 747 (1961).

[199] G. Meissner, H. Namaizawa and M. Voss, *Phys. Rev. B* **13**, 1370 (1976).

[200] L. Bonsall and A. A. Maradudin, *Phys. Rev. B* **15**, 1959 (1977).

[201] D. Ceperley, *Phys. Rev. B* **18**, 3126 (1978).

[202] B. Tanatar and D. M. Ceperley, *Phys. Rev. B* **39**, 5005 (1989).

[203] Y. Kwon, D. M. Ceperley and R. M. Martin, *Phys. Rev. B* **48**, 12037 (1993).

[204] G. Ortiz and P. Ballone, *Phys. Rev. B* **50**, 1391 (1994).

[205] F. Rapisarda and G. Senatore, *Aust. J. Phys.* **49**, 161 (1996).

[206] Y. Kwon, D. M. Ceperley and R. M. Martin, *Phys. Rev. B* **58**, 6800 (1998).

[207] G. Ortiz, M. Harris and P. Ballone, *Phys. Rev. Lett.* **82**, 5317 (1999).

[208] D. Varsano, S. Moroni and G. Senatore, *Europhys. Lett.* **53**, 348 (2001).

[209] M. W. C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).

[210] C. Attaccalite, S. Moroni, P. Gori-Giorgi and G. B. Bachelet, *Phys. Rev. Lett.* **88**, 256601 (2002).

[211] F. H. Zong, C. Lin and D. M. Ceperley, *Phys. Rev. E* **66**, 036703 (2002).

[212] N. D. Drummond and R. J. Needs, *Phys. Rev. Lett.* **102**, 126402 (2009).

[213] A. Lüchow, R. Petz and A. Schwarz, *Z. Physik. Chem.* **224**, 343 (2010).

[214] J. J. Shepherd, G. Booth, A. Grüneis and A. Alavi, *Phys. Rev. B* **85**, 081103 (2012).

[215] S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).

[216] J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).

[217] J. Sun, J. P. Perdew and M. Seidl, *Phys. Rev. B* **81**, 085123 (2010).

[218] P. M. W. Gill and P. F. Loos, *Theor. Chem. Acc.* **131**, 1069 (2012).

[219] A. J. Pérez-Jiménez, F. Moscardó, J. C. Sancho-García, L. P. Abia, E. San-Fabián and J. M. Pérez-Jordá, *J. Chem. Phys.* **114**, 2022 (2001).

[220] R. Colle and O. Salvetti, *Theor. Chim. Acta* **37**, 329 (1975).

[221] H. Stoll, E. Golka and H. Preuss, *Theor. Chim. Acta* **55**, 29 (1980).

[222] A. D. Becke, *Int. J. Quantum Chem.* **23**, 1915 (1983).

[223] W. L. Luken and J. C. Culberson, *Theor. Chim. Acta* **66**, 279 (1984).

[224] J. F. Dobson, *J. Chem. Phys.* **94**, 4328 (1991).

[225] P. F. Loos and P. M. W. Gill, *Phys. Rev. A* **79**, 062517 (2009).

[226] P. F. Loos and P. M. W. Gill, *Phys. Rev. Lett.* **103**, 123008 (2009).

[227] P. F. Loos and P. M. W. Gill, *Mol. Phys.* **108**, 2527 (2010).

[228] P. F. Loos and P. M. W. Gill, *J. Chem. Phys.* **135**, 214111 (2011).

[229] C. A. Coulson and A. H. Neilson, *Proc. Phys. Soc. (London)* **78**, 831 (1961).

[230] A. J. Coleman, *Int. J. Quantum Chem., Symp.* **1**, 457 (1967).

[231] J. P. Perdew and Y. Wang, *Phys. Rev. B* **46**, 12947 (1992).

[232] J. Cioslowski and G. Liu, *J. Chem. Phys.* **109**, 8225 (1998).

[233] A. M. Lee and P. M. W. Gill, *Chem. Phys. Lett.* **313**, 271 (1999).

[234] P. M. W. Gill, A. M. Lee, N. Nair and R. D. Adamson, *J. Mol. Struct. Theochem.* **506**, 303 (2000).

[235] P. M. W. Gill, D. P. O'Neill and N. A. Besley, *Theor. Chem. Acc.* **109**, 241 (2003).

[236] P. M. W. Gill, D. L. Crittenden, D. P. O'Neill and N. A. Besley, *Phys. Chem. Chem. Phys.* **8**, 15 (2006).

[237] P. M. W. Gill, *Annu. Rep. Prog. Chem., Sect. C* **107**, 229 (2011).

[238] A. J. Proud, M. P. Walker and J. K. Pearson, *Int. J. Quantum Chem.* **113**, 76 (2013).

[239] E. R. Davidson, *Reduced density matrices in quantum chemistry* (Academic Press, New York, 1976).

[240] A. D. Becke, *J. Chem. Phys.* **88**, 1053 (1988).

[241] J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).

[242] R. J. Needs, M. D. Towler, N. D. Drummond and P. L. Rios, *J. Phys. Condens. Matter* **22**, 023201 (2010).

[243] P. J. Knowles and N. C. Handy, *Chem. Phys. Lett.* **111**, 315 (1984).

[244] P. J. Knowles and N. C. Handy, *Comput. Phys. Commun.* **54**, 75 (1989).

[245] J. Cioslowski, *J. Chem. Phys.* **136**, 044109 (2012).

[246] W. Kutzelnigg, *Int. J. Quantum Chem.* **51**, 447 (1994).

[247] L. K. McKemmish and P. M. W. Gill, *J. Chem. Theory Comput.* **8**, 4891 (2012).

[248] W. Kutzelnigg, *Int. J. Quantum Chem.* **113**, 203 (2013).

[249] E. A. Salter, G. W. Trucks and D. S. Cyphert, *Am. J. Phys.* **69**, 120 (2001).

[250] J. Deng and P. M. W. Gill, *J. Chem. Phys.* **134**, 081103 (2011).

[251] J. Noga and W. Kutzelnigg, *J. Chem. Phys* **101**, 7738 (1994).

[252] A. Köhn and D. P. Tew, *J. Chem. Phys.* **133**, 174117 (2010).

[253] N. R. Kestner and O. Sinanoğlu, *Phys. Rev.* **128**, 2687 (1962).

[254] E. Santos, *Anal. R. Soc. Esp. Fis. Quim.* **64**, 177 (1968).

[255] R. J. White and W. Byers Brown, *J. Chem. Phys.* **53**, 3869 (1970).

[256] J. M. Benson and W. Byers Brown, *J. Chem. Phys.* **53**, 3880 (1970).

[257] S. Kais, D. R. Herschbach and R. D. Levine, *J. Chem. Phys.* **91**, 7791 (1989).

[258] M. Taut, *Phys. Rev. A* **48**, 3561 (1993).

[259] S. Ivanov, K. Burke and M. Levy, *J. Chem. Phys.* **110**, 10262 (1999).

[260] J. Cioslowski and K. Pernal, *J. Chem. Phys.* **113**, 8434 (2000).

[261] T. M. Henderson, K. Runge and R. J. Bartlett, *Chem. Phys. Lett.* **337**, 138 (2001).

[262] D. P. O'Neill and P. M. W. Gill, *Phys. Rev. A* **68**, 022505 (2003).

[263] J. Katriel, S. Roy and M. Springborg, *J. Chem. Phys.* **123**, 104104 (2005).

[264] P. M. W. Gill and D. P. O'Neill, *J. Chem. Phys.* **122**, 094110 (2005).

[265] S. Ragot, *J. Chem. Phys.* **128**, 164104 (2008).

[266] J. Cioslowski, *J. Chem. Phys.* **139**, 224108 (2013).

[267] P. F. Loos and P. M. W. Gill, *Mol. Phys.* **110**, 2337 (2012).

[268] T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).

[269] C. Schwartz, *Phys. Rev.* **126**, 1015 (1962).

[270] R. N. Hill, *J. Chem. Phys.* **83**, 1173 (1985).

[271] W. Kutzelnigg and J. D. Morgan III, *J. Chem. Phys.* **96**, 4484 (1992).

[272] B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.* **98**, 5612 (1993).

[273] A. W. Overhauser, *Phys. Rev. Lett.* **3**, 414 (1959).

[274] A. W. Overhauser, *Phys. Rev.* **128**, 1437 (1962).

[275] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari and L. A. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989).

[276] J. A. Pople, P. M. W. Gill and B. G. Johnson, *Chem. Phys. Lett.* **199**, 557 (1992).

[277] R. Neumann, R. H. Nobes and N. C. Handy, *Mol. Phys.* **87**, 1 (1996).

[278] C. Adamo, M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **112**, 2643 (2000).

[279] A. V. Arbuznikov and M. Kaupp, *Chem. Phys. Lett.* **381**, 495 (2003).

[280] R. T. Sharp and G. K. Horton, *Phys. Rev.* **90**, 317 (1953).

[281] E. Engel and R. M. Dreizler, *J. Comput. Chem.* **20**, 31 (1999).

[282] M. Ernzerhof, S. N. Maximoff and G. E. Scuseria, *J. Chem. Phys.* **116**, 3980 (2002).

[283] I. G. Ryabinkin and V. N. Staroverov, *J. Chem. Phys.* **137**, 164113 (2012).

[284] R. Gaudoin and K. Burke, *Phys. Rev. Lett.* **93**, 173001 (2004).