Mathematical Analysis of Quantum Chemical Models for Small Atoms

Benjamin David Goddard

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Mathematics Department, University of Warwick

August 2007

THE UNIVERSITY OF
WARWICK

Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor Gero Friesecke, for initially introducing me to the subject of mathematical quantum chemistry, for his advice and guidance during my Ph.D., and for the numerous stimulating conversations.

I am indebted to many fellow graduate students for support, advice, and necessary mathematical and extracurricular distractions throughout the course of this research. In lexicographic order; Abi Davies, Amanda Window, Caroline Poos, Greg Franklin, James Smith, Jon Potts, Matthew Knight, Nicholas Jackson, Oliver Tearne, Simon Lloyd, and Tim Sullivan.

Danke schön to Anna Dall'Acqua, Jannis Giannoulis, and Mario Koppen for helping to make my stay in Munich such a pleasant experience.

I would also like to thank my examiners, Peter Gill and Alex Sobolev, and my examination coordinator, Volker Betz, for their useful comments and suggestions.

My deepest thanks to Christina Fell for her unwavering faith, understanding, and support throughout the last four years; I owe you my sanity.

Last, but by no means least, a huge thank you to Mum, Dad, and my brother Alan for their loving support, encouragement, and inspiration.

This Ph.D. was partially supported by the EPSRC.

Declaration

I declare that, to the best of my knowledge, the material contained in this thesis is original and my own work except where otherwise indicated, cited, or commonly known.

The material in this thesis is submitted for the degree of Ph.D. to the University of Warwick only, and has not been submitted to any other university.

Abstract

Quantum chemical methods are of great use in predicting chemical and physical properties of atoms and molecules. Whilst many of the methods are built upon a rigorous mathematical foundation, there are very few chemically-specific mathematical results, and many aspects, such as the choice of optimal basis sets, are poorly understood.

This thesis centres on the choice of a minimal basis for the second row atoms, motivated by perturbation theory and physical screening arguments, for which a full Configuration Interaction (CI) calculation is performed. This calculation requires no numerical integration and can essentially be performed ‘by hand’. As well as demonstrating that this method leads to an accurate prediction of the atomic spectra, a number of computationally useful results concerning the Hamiltonian matrix and the spectra themselves are rigorously proven.

A rigorous rate of convergence for the CI method applied to the ground state of the Helium atom, with a basis ordered by angular momentum, is also derived. This includes a derivation of the leading order constant, enabling extrapolation of computational results.

Another interesting area of computational chemistry is the prediction of molecular geometries. This thesis investigates the AH_2 trimers, where A is a second row atom. The relationship between the maximum of the pair density of the central atom and the bond angle is investigated, using the canonical ground state wavefunctions derived in the CI calculations. The non-numerical results are independent of the radial parts of the wavefunctions and inserting the CI wavefunctions leads to excellent qualitative and reasonably accurate quantitative numerical predictions.

Chapter 1

Introduction

1.1 The Schrödinger Equation

In January 1926, Schrödinger published his wave equation [Sch26] for quantum mechanics, giving a 'derivation' for time independent systems. Furthermore, he showed that it gave the correct energy eigenvalues for the Hydrogen-like atom. Whilst these solutions for a one-electron, one-nucleus problem can be found explicitly, this is not possible for other atoms or molecules.

Therefore, effort must be concentrated upon finding accurate approximate solutions to the many-electron equation. Almost all of these methods begin with the non-relativistic, Born-Oppenhiemer approximation [BO27] to the Schrödinger equation. In atomic units ($m = e = \hbar = a_0 = 1$), for an atom with N electrons, this is given by

$$H\Psi = E\Psi$$

where H is the operator

$$H = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i - \sum_{i=1}^{N}\frac{Z}{r_i} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}},$$

Z is the nuclear charge, r_i is the distance between the nucleus and electron i and r_{ij} is the distance between the two electrons i and j.

The Born-Oppenheimer approximation is based upon the idea that nuclei are at least 1800 times heavier than electrons. This multi-scale effect leads to an uncoupling of the motions of the nuclei and electrons, allowing the nuclei to be treated as stationary point particles when considering the electronic motion. For the majority of atomic and molecular systems, this approximation introduces only very small errors.

We note here that there are no empirical parameters in the Schrödinger equation; the solutions are completely *ab initio*, predicting physical values such as energy and molecular geometry without any reference to experimental data.

1.2 The Hartree-Fock Method

One of the most widely used theories in the prediction of atomic ground states is Hartree-Fock theory. The origins of this theory are due to Hartree [Har28] who claimed that a good approximation to an atomic ground state could be obtained by replacing the inter-electron repulsion term V_{ee} in the Hamiltonian by the average repulsion between the electrons. With this assumption the Hamiltonian becomes separable and the wavefunction may be written as a product of one-electron orbitals known as a Hartree Product:

$$\Psi_{HP}(x_1, x_2, \ldots, x_N) = \psi_1(x_1)\psi_2(x_2)\ldots\psi_N(x_N).$$

This form is obviously very simple and convenient for calculation, however, as well as ignoring the inter-electron term it has another shortcoming - it fails to satisfy the antisymmetry condition.

The next stage in the approximation was suggested by Fock [Foc30] who noted the lack of antisymmetry and suggested that the wavefunction be written instead as a Slater determinant. This gives an antisymmetric wavefunction in which the electrons are indistinguishable. Due to the variational principle the energy of this single Slater determinant is an upper bound to the true energy of the system.

In order to find the best upper bound for the energy it is necessary to find the set of single electron spin orbitals which minimizes the energy, i.e. to find the best single determinant. Since determinants are invariant under application of any (non-singular) linear transformation, we may choose the orbitals to be orthonormal. In order to minimize the upper bound this single determinant is inserted into the full Hamiltonian and we introduce a Lagrange multiplier to impose orthonormality and minimize with respect to the ψ_i, obtaining the set of equations [SO96]

$$f\psi_i = \epsilon_i\psi_i, \quad i = 1, \ldots, N,$$

where the Fock operator f is given by

$$f = h_{core} + \sum_{j=1}^{N}(J_j - K_j),$$

where, for electron positions (in both space and spin) x_i and nuclei of charge Z_A at position R_A, the core Hamiltonian h_{core}, Coulomb operator J_j and exchange operator K_j are defined by

$$\begin{aligned} h_{core}\psi(x) &= \left(-\frac{1}{2}\Delta - \sum_{A=1}^{M} \frac{Z_A}{|x - R_A|}\right)\psi(x), \\ J_j\psi(x) &= \psi(x)\int dx' \frac{|\psi_j(x')|^2}{|x - x'|}, \\ K_j\psi(x) &= \psi_j(x)\int dx' \frac{\psi_j^*(x')\psi(x')}{|x - x'|}. \end{aligned}$$

These equations are known as the Hartree-Fock (HF) equations. Although they look superficially simple, they are highly non-trivial to solve; they are non-linear and so must be solved self-consistently. Hence HF theory is called a self-consistent-field (SCF) approach. In practice this means that an initial guess of the orbitals must be made, the HF equations are used to compute a new set of orbitals, and the process repeated until the energy does not change by more than a given threshold.

Before we discuss the various computational methods used to solve the HF equations we make some important observations about the equations and solutions. The three terms in the Hamiltonian can be explained as follows: h_{core} simply gives the kinetic and nuclear potential energy of the individual electrons. J_j is the electrostatic potential arising from the charge distribution of all N electrons, including the unphysical self interaction term. This term is cancelled in the final K_j term called the exchange term. This non-local operator results from the antisymmetry condition and causes same-spin electrons to avoid each other.

The ϵ_i, which were originally the Lagrange multipliers, are known as the orbital energies, a definition which is consistent with the fact that the orbitals are eigenfunctions of the Fock operator. The total HF energy is simply given by the sum of orbital energies (plus the nuclear-nuclear terms in a molecular system). In addition, Koopmans' theorem [SO96] states that the ionization energy of a system is given by the negative of the highest orbital energy.

Another point of note is that the HF method does not include the energy that comes from the interaction of different Slater determinants within a true eigenfunction of the full Hamiltonian. The energy which this cannot capture is known as the correlation energy and is defined [SO96] as

$$E_{corr} = E_0 - E_{HF} \tag{1.1}$$

where E_0 is the exact non-relativistic energy.

1.2.1 Solution of the Hartree-Fock equations

As noted in the previous section, solution of the non-linear HF equations is highly non-trivial. One method of solution is purely numerical, see e.g. [FF77] and is known as exact HF. Another method solves the equations in a space spanned by a set of basis functions and is known as the Hartree-Fock-Roothan (RHF) [Roo60] method. Assuming sufficient computing power, one would expect that exact HF will give the better numerical results, however the RHF wavefunctions can prove useful in themselves, for example as initial wavefunctions in a configuration interaction calculation [BBBC92].

In order to make the equations more easily solved a number of approximations are commonly made. We will discuss only the case of atoms in which the first assumption is that the one-electron orbitals can be split into a spatial part and a spin part $\psi_i(x) = \psi_i(r, s) = \phi(r)\chi(s)$. The second assumption is that the spatial parts of the wavefunctions are solutions to a central-potential (i.e. atomic) Hamiltonian. This leads to the chemically well-known idea of *s*-, *p*-, *d*-orbitals and so on [AdP01].

One further choice concerns the spatial parts of spin-paired orbitals. In restricted HF (often referred to as simply HF), the spatial orbitals of spin-paired electrons are taken to be the same. In unrestricted HF (UHF), the orbitals are allowed to vary independently, thus leading to an increase in accuracy but at the expense of extra computation.

Once a basis $\phi(r)$ has been chosen, the spatial part of the one electron orbitals can be expanded as $\psi_i = \sum_{j=1}^{J} c_{ij}\phi_j$, leading to a matrix equation. The problem is then reduced to linear algebra and many powerful matrix-diagonalization techniques can be used. There is wide range of specialized software available for such computations.

The two most common choices of the ϕ_j are Gaussians and Slater-type orbitals (STO's). The main advantage of Gaussians is that they lead to comparatively simple integrals, especially for the molecular case. The main disadvantage is that a much larger number of Gaussians are needed than STO's to model a physically realistic wavefunction, especially at a cusp [Cle65].

1.2.2 Extensions of the Hartree-Fock Method

Whilst the HF method gives a good approximation for wavefunctions of closed shell systems (where each spatial orbital is doubly occupied), which are well-represented by a single Slater determinant, the accuracy is not as high for open shell systems. A number of approaches choose to extend the method to use a linear combination of Slater determinants, reducing the correlation energy and allowing the formation of spin and angular momentum eigenfunctions.

Again, the two most popular methods are RHF, extended to multiple determinants for open shells (MDRHF), and numerical (exact) Multi-Configuration HF (MCHF) the most prominent example being the code of Froese Fischer [FF69, FF72]. In open shell RHF it is usual to form spin and angular momentum eigenfunctions from Slater determinants containing orbitals of the same general form, e.g. $|1s^2 2s^m 2p^n\rangle$, by taking different spin combinations along with different choices of orbitals from the symmetry groups, e.g. $\frac{1}{\sqrt{3}}(|p_1 \uparrow p_1 \downarrow\rangle + |p_2 \uparrow p_2 \downarrow\rangle + |p_3 \uparrow p_3 \downarrow\rangle)$.

MCHF has two distinct differences to true HF, firstly, as in MDRHF, it allows the use of multiple determinants for open shells. Furthermore, it can be extended to include correlation between different wavefunctions of the same symmetry, the most common example being to replace $2s \uparrow 2s \downarrow$ with $2p \uparrow 2p \downarrow$ when possible [VM70]. This method clearly produces a better approximation to the true wavefunction and gives a lower energy, but moves much further away from the simple HF idea of a single determinant. With MCHF it is possible to capture at least some of the correlation energy missing in true HF, thus negating one of the major errors of the original theory. This Multi-Configuration approach is also possible with RHF but seems to be much less frequently utilized.

In fact, with multiple configurations, the method would be better described as a Multi-Configuration Self-Consistent-Field method, the only connection with HF being the model potential used. There are other possible methods which use a self-consistent potential and wavefunction, although the Hartree-Fock potential seems to be most widely used.

Unfortunately, there seems to be little standardization of notation in the literature and it is often unclear which methods are being used, especially when a piece of code is referenced without explicitly stating which options (e.g. HF/RHF/MCHF) were used. It is also the case that RHF is used to refer to multi-determinant approximations without explicit statement of this and the results are then referred to as 'Hartree-Fock' when in fact they are nothing of the sort [VM70].

The use of multiple determinants is most obvious in the literature dealing with excited states of atoms, for which spin and angular momentum eigenfunctions are required. For results such as these (which do not use multiple configurations) we will adopt the notation Multi-determinant Hartree-Fock (MDHF) whilst RHF will simply denote the method of calculation through choice of a basis.

The term 'Hartree-Fock' seems to have become more synonymous with replacing the true $1/r_{12}$ term in the Hamiltonian by an average inter-electron potential than with its original definition as a single Slater determinant approximation. In summary, ordered by increasing complexity (and also increasing accuracy) we have HF < MDHF < MCHF with RHF and exact (numerical) HF simply being different computational methods.

1.3 Post Hartree-Fock Methods

We now move on to describe some of the more accurate methods currently used. In general, the improvement in these methods comes from modelling the correlation energy (1.1), which is the main barrier to HF calculations achieving chemical accuracy.

1.3.1 Configuration Interaction

The first of these is Configuration Interaction (CI) (see e.g. [SO96]) in which the exact eigenfunctions of H are expanded as

$$\Psi_j = \sum_i c_{ij} |\Phi_i\rangle$$

where $|\Phi_i\rangle$ are a basis of Slater determinants. However, this basis is infinite [Fri02] which prohibits working with the whole space and thus the expansion is truncated

$$\Psi_j = \sum_i^I c_{ij} |\Phi_i\rangle.$$

The matrix of H is constructed by $H_{ij} = \langle\Phi_i|H|\Phi_j\rangle$ for $i, j = 1, \dots, I$. The aim of CI is to diagonalize this matrix and thus find the eigenfunctions of H. Due to the variational theorem, the energy of the truncated wavefunction converges to the exact energy as $I \to \infty$.

The connection with HF is that the $|\Phi_i\rangle$ can be written as excitations from the HF reference determinant $|\Phi_0\rangle$:

$$\Psi = c_0|\Phi_0\rangle + \sum_{r,a} c_a^r |\Phi_a^r\rangle + \sum_{a<b,r<s} c_{ab}^{rs} |\Phi_{ab}^{rs}\rangle + \cdots$$

where $|\Phi_a^r\rangle$ denotes replacing orbital a by orbital r and so on.

Every $|\Phi_i\rangle$ can be described by the one electron orbitals it contains, which are often called a configuration. As well as taking a truncated expansion of Ψ it is also computationally necessary to take a finite set of one-electron orbitals from which to form the $|\Phi_i\rangle$. The quality of the one-electron basis set can be tested by considering the improvement in energy when the basis set is expanded.

One easy simplification when computing, for example, a particular spin and angular momentum eigenfunction is that it is only necessary to use linear combinations of the $|\Phi_i\rangle$ which are also eigenfunctions with the same eigenvalues. If the matrix calculations are performed using all possible $|\Phi\rangle$ (possibly symmetry restricted) formed from a given basis of one-electron orbitals then this is called full CI. The theoretical method of using a complete one-electron basis is called complete CI.

Even with a relatively incomplete basis of one-electron orbitals, the size of the basis of N-electron Slater determinants grows as $N!$ and so the full CI approach is not computationally feasible. Perhaps the most common truncation is to restrict the number of excitations from the HF reference determinant $|\Phi_0\rangle$. Since the Hamiltonian contains only one- and two-electron terms only singly and doubly excited determinants can interact directly with $|\Phi_0\rangle$. This is the minimal practical level of excitation and is known as CISD, allowing triple excitations it becomes CISDT and so on. In small molecules at their equilibrium geometries, CISD captures around 95% and CISDT around 99% of the correlation energy [HH83].

In fact, due to Brillouin's theorem [SO96] the singly excited Slater determinants do not interact with $|\Phi_0\rangle$ if it is the HF solution. Although they have a very small effect on the energy, through coupling with doubly excited terms which couple directly to the ground state, the singly excited states are normally included in a CI calculation due to their relatively small number and their effect on one-electron properties such as dipole moment.

One of the major challenges in CI is to choose a good basis of both one- and N-electron wavefunctions. For example, Brillouin's theorem, spin and angular momentum symmetries, and many other properties allow one to determine *a priori* that some matrix elements will be zero.

A standard approximation is to assume a 'frozen core' model in which the inner shell electron wavefunctions are fixed. For the second row atoms of Lithium to Neon this frozen core is the two $1s$ electrons. It can be shown that, for doubly occupied core orbitals, these orbitals can be deleted from the $|\Phi_i\rangle$ giving $|\overline{\Phi}_i\rangle$ and the matrix elements become $\langle\Phi_i|H|\Phi_j\rangle = \langle\overline{\Phi}_i|H_{FC}|\overline{\Phi}_j\rangle$ where

$$H_{FC} = E_c + \sum_{i=1}^{N-N_c} h_c(i) + \sum_{i<j}^{N-N_c} r_{ij}^{-1},$$

N_c is the number of core orbitals, E_c the 'frozen core energy' (expected value of the determinant formed from the N_c core orbitals), and $h_c(i)$ is the one-electron Hamiltonian for electron i in the average field of the core electrons [HDGP75].

1.3.2 Møller-Plesset Perturbation Theory

Møller-Plesset Perturbation theory (MPPT) (see e.g. [SO96]), along with CI and MCHF is one of the most widely used correlated electron approaches to quantum chemistry. Essentially, a small perturbation to the HF Hamiltonian is introduced with a parameter λ and the resulting wavefunction and energy are expressed as an asymptotic series in λ [MP34]. Zero order gives the sum of orbital energies, first order the HF energy and orders above that improve on the HF method by capturing correlation effects. Typically calculations are truncated at fourth order due to increasing computational complexity, but higher orders are possible for isolated calculations [KNB89, HC96a, HC96b].

Recent studies of higher orders have shown that, in a wide range of cases, the method is not necessarily convergent and may converge slowly, diverge at high orders or show erratic behaviour [LASS00]. The results at a given order are also strongly dependent on the chemical system and choice of one-orbital basis. It therefore appears that the assumption that the series converge is not valid. It is important to note that, unlike HF and CI, MPPT is not a variational method at any given level of perturbation, so a lower energy may not be a better approximation to the true energy.

1.3.3 Coupled Cluster Theory

Coupled Cluster (CC) (see e.g. [SO96]) makes the ansatz that the wavefunction of the ground state can be written as $\Psi = e^T|\Phi_0\rangle$ where $|\Phi_0\rangle$ is a Slater determinant, usually the HF ground state. T is an excitation operator which acts on $|\Phi_0\rangle$ by the expansion $e^T = 1 + T + \frac{T^2}{2!} + \dots$. Further, T can be expanded as $T = T_1 + T_2 + T_3 + \dots$ where T_1 causes single excitations, T_2 double excitations and so on. Each of the T_i can then be expanded in a linear combination of excitations, normally expressed in second quantization, and the goal of the theory is to find the minimizing expansion coefficients.

This is much in the spirit of the CI expansions however, a lower level of excitation is needed as the exponential operator causes non-linearity ($e^T = 1 + T_1 + T_2 + \frac{1}{2}T_1^2 + T_1T_2 + \frac{1}{2}T_2^2 + \cdots$) meaning that an expansion truncated after k excitations can generate Slater determinants with more than k excited orbitals. This means that CC will in general produce a more accurate wavefunction for a given level of excitation than CI. Notation for a given level of excitation is of the form CCSD(T) where terms in parentheses are calculated using perturbation theory, the case shown being single and double excitations with triple excitations calculated with perturbation theory. This is the most common level used as it offers a good balance between accuracy and computational expense.

Using CCSD for molecules at their equilibrium geometries, around 95% of the correlation energy is recovered and this is increased to around 99% when using CCSD(T) [HG96]. This means that CCSD(T) is sufficient to achieve chemical accuracy in these cases.

CC theory can be expanded to deal with excited states but its main drawback, as with CI and other expansions based on the HF determinant, is when trying to model bond breaking and other situations when it is clear that the HF solution is not valid. In such cases the wavefunction must be described as a sum of Slater determinants of orbitals centered on each separated atom. This is less of a problem in methods like CI where these linear combinations of determinants are possible.

1.3.4 Density Functional Theory

Density Functional Theory (DFT) is almost certainly the most widely used technique for all but the smallest atomic and molecular calculations. The reason for this can be seen first by considering the computational complexity of methods such as CI. Suppose we take a representation of Ψ and wish to discretize each coordinate using 10 mesh points (a very small number). For N electrons Ψ is a function of $3N$ (ignoring spin) variables and hence 10^{3N} values are needed. It is clear that for systems with more than about ten electrons this is computationally unfeasible.

DFT overcomes this problem by replacing the many-body wavefunction by the electron density which, regardless of the number of electrons, is a function of only three spatial positions. In a sense this means that DFT is an extension of Thomas-Fermi methods. The development of DFT was led by Walter Kohn, for which he received the Nobel Prize in Chemistry in 1998. In 1964, Kohn and Hohenberg [HK64] proved that a given ground state density can, in principle, be used to calculate the ground state wavefunction. In other words, the ground state wavefunction is a unique functional of the density. It follows that all other properties, including the energy, are functionals of the density. In 1965 Kohn and Sham [KS65] showed how to minimize the energy functional using a self-consistent method. This formulation can be extended to excited states using Time Dependent DFT [BWG05].

The main problem with DFT is that the exact functionals for exchange and correlation are not known for atomic or molecular systems. Thus approximate functionals such as B3LYP [Bec93] which combines a DFT exchange functional (BLYP from the names Becke, Lee, Yang and Parr [LYP88]) with the exact exchange functional from HF. This scheme includes three parameters, which are chosen by fitting to trial functions of small molecules. Schemes with larger numbers of parameters are also used (for example [VVS98] has 21 parameters). This empirical fitted DFT functionals do not really qualify as *ab initio* methods. However, the results of DFT calculations are incredibly good, for example predicting energies of small molecules to within 0.01 au and bond lengths to within 0.001nm [PTSS04].

We do not discuss DFT in any more detail as it has few similarities with our proposed method and there are a number of reviews on the subject such as [PY95a]. In particular, for atoms and small molecules, DFT produces almost exact energies and is more accurate than any computationally feasible basis-dependent approximation.

1.4 Computational Cost, Accuracy, and Understanding

Modern theoretical chemists have access to enormous computing power and many of the most intensive computations are in the field of quantum chemistry. In fact, as of June 2006, the most computationally intensive programme ever was a quantum mechanical simulation of 1000 molybdenum atoms under high pressure [Kni06] . Full CI calculations for systems involving a few billion determinants are possible [OJK$^+$96]. However, due to the factorial growth of full CI, even small molecules are beyond its reach when chemical accuracy is required. In practice, CCSD(T) is the most reliable *ab initio* method that is computationally tractable for medium sized molecules. For much larger systems, Møller-Plesset to second order (MP2) remains the best option. Formally, HF, DFT, MP2, and CCSD(T) scale as N^4, N^4, N^5, and N^7 respectively. For practical calculations it is possible to reduce the order of HF to N^3 or even N^2, DFT to similar order, and MP2 to N^3, and even further to N^2 at long range [SACKH95, PY95b, Fri05].

From the previous discussion we see that there are two main sources of error in the energy, that which comes from correlation effects and that which is introduced by truncation of the basis sets used. Whilst the above post-HF methods and large amounts of computing power allow theoretical chemists to make very accurate predictions (at least on small systems), this method of solution leads to very little intuitive understanding of the solutions.

Two specific examples are numerical HF and DFT. For large basis sets in numerical HF the intuitive idea of orbitals ordered by quantum number, e.g. $1s, 2s, 2p, \ldots$ becomes blurred. It is hard to extract these one-electron orbitals from the vast amounts of numerical data. In DFT, orbitals have to be extracted from the density and hence are 'natural' but not 'intuitive'. Also, in DFT with experimentally fitted parameters there seems to be very little physical understanding of why a particular parameter should have the chosen value and exactly how this value affects the solution.

This situation contrasts sharply with calculations performed before large amounts of computing power were widely available (e.g. [HH36]), which focussed much more on the physical description of orbitals and how to choose small, accurate basis sets which allow easy computation. This is an area which seems to have great potential for application in methods such as CI. If a suitable basis can be chosen such that the necessary integrals are easy to perform and give analytic solutions (which would allow efficient and accurate storage, reducing numerical errors), and also cause the full CI matrix to be relatively sparse (leading to easy diagonalisation), then this would vastly improve the computational efficiency of such methods.

Although many of the methods described above are built upon a sound mathematical foundation, and in general will converge in the limit of complete basis sets (note the exception of the limit of MPPT), there is little mathematical understanding of the more subtle aspects of these methods. These aspects include the rate of convergence of a given method in a given basis, how to choose optimal small basis sets for different problems, the structure of the Hamiltonian matrix in a given basis set, and rigorous (basis independent) spectral orderings.

As well as being mathematically interesting, these questions are relevant for computational calculations. As well as examples discussed previously, knowledge of the rate (and leading constant) of convergence for a method would allow extrapolation of computational results to higher accuracy.

Chapter 2

The Many Electron Energy Matrix

2.1 Introduction

We wish to use first-order perturbation theory to derive a ‘pen and paper’ method which will accurately predict a number of properties of atoms in the second period. Basic such properties relevant to chemical behaviour are given by the spatial and spin symmetries of the electronic ground state, as captured mathematically by its total angular momentum and total spin quantum numbers. The first property we are concerned with is the spectrum of the atom, including the ground state energy, ordering of spectral lines associated with spin-angular momentum eigenspaces, and their spectral gap energies.

The key point distinguishing our calculations from the literature is that neither empirical shell ordering rules, nor Hund’s rules, nor statistical independence approximations, nor numerical simulations are invoked. Furthermore, in Section 4.3, we will derive a rigorous spectral ordering and compare this to the empirical Hund’s rules.

2.2 Perturbation Theory

2.2.1 The Schrödinger Equation

We begin with the non-relativistic, Born-Oppenhiemer approximation to the Schrödinger equation, in atomic units ($m = e = \hbar = a_0 = 1$), for an atom with N electrons [SO96]

$$H\Psi = E\Psi$$

where H is the operator

$$H = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i - \sum_{i=1}^{N}\frac{Z}{r_i} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}, \tag{2.1}$$

Z the nuclear charge, r_i the distance between the nucleus and electron i, and r_{ij} the distance between electrons i and j.

We note that the our main interest will be in the case of neutral atoms but our analysis applies equally to the case $Z \geq N$, which will be used in Section 5.3 to compute predicted first ionization energies.

As is usual in (simple) wavefunction-based atomic calculations, we choose to work with a many-electron basis formed from antisymmetric products of one-electron atomic orbitals.

2.2.2 One-electron orbitals

Firstly we note the well-known result that the one-electron Hydrogen equation $H_0\psi = E\psi$, $H_0 = -\frac{1}{2}\Delta - 1/r$ is exactly soluble (see e.g. [Gri95]). We may characterize the eigenfunctions by three quantum numbers; the principle quantum number n (which corresponds to the energy in the non-interacting case); the orbital angular momentum quantum number ℓ (which corresponds to the total angular momentum); and the magnetic quantum number m (which corresponds to the z-component of angular momentum, where z is an arbitrary fixed axis). There is also a fourth quantum number for multi-electron atoms that is not determined by the spatial Hydrogen equation, the spin s.

The angular momentum quantum numbers are determined by the four operators

$$\begin{aligned}
L_1 &= \frac{1}{i}\left(x_2\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_2}\right), \\
L_2 &= \frac{1}{i}\left(x_3\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_3}\right), \\
L_3 &= \frac{1}{i}\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right), \text{ and} \\
\underline{L}^2 &= (L_1)^2 + (L_2)^2 + (L_3)^2.
\end{aligned}$$

Each of the L_i commute with $\underline{L}^2$ but not with each other, and all four operators commute with the Hamiltonian [Sch01]. Hence it is possible to jointly solve

$$H_0\psi = E\psi, \quad \underline{L}^2 = \lambda\psi, \text{ and } L_3\psi = \mu\psi,$$

where the choice of L_i is arbitrary.

Solution of these equations leads to the constraints $n \in \mathbb{N}$, $\ell = 0, 1, \ldots, n$, and $m = -\ell, -\ell+1, \ldots, 0, \ldots, \ell-1, \ell$. The energy of an orbital with principal quantum number n is $E = -\frac{1}{2n^2}$, regardless of the values of the other quantum numbers. The other two eigenvalues are $\lambda = \ell(\ell+1)$ and $\mu = m$.

Introducing a spin variable which takes values $s \in \{-\frac{1}{2}, \frac{1}{2}\} \cong \mathbb{Z}_2$, we denote the normalized eigenfunctions by $\psi_i(x,s)$, where $x \in \mathbb{R}^3$, ordered such that the corresponding eigenvalues E_i satisfy $E_i \leq E_{i+1}$. For energy $-\frac{1}{2n^2}$, the degenerate eigenspaces are of dimension $2n^2$ and are spanned by

$$\begin{aligned} V_n :&= \{\psi_i(x,s) \mid i = 1 + (n-1)n(2n-1)/3, \ldots, n(n+1)(2n+1)/3\} \\ &= \{\phi_{n\ell m s}(x,s) \mid \ell = 0, \ldots, n-1,\ m = -\ell, \ldots, \ell,\ s = \pm 1/2\}, \end{aligned}$$

where $\phi_{n\ell m s}$ is the unique normalized eigenfunction of H_0, $\underline{L}^2$ and L_3 with eigenvalues $-\frac{1}{2n^2}$, $\ell(\ell+1)$ and m and with spin s.

2.2.3 Many-electron eigenfunctions

A many-electron wavefunction must be antisymmetric under interchange of the spatial and spin parts of two different one-electron orbitals. The simplest form of such a wavefunction formed from N one-electron wavefunctions ψ_i, $i = 1, \ldots, N$, is

$$|\psi_i \ldots \psi_N\rangle := \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1,s_1) & \psi_2(x_1,s_1) & \ldots & \psi_N(x_1,s_1) \\ \psi_1(x_2,s_2) & \psi_2(x_2,s_2) & \ldots & \psi_N(x_2,s_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(x_N,s_N) & \psi_2(x_N,s_N) & \ldots & \psi_N(x_N,s_N) \end{vmatrix},$$

and is called a Slater determinant. The antisymmetry condition follows from the fact that interchanging two columns of a matrix changes the sign of the determinant. The prefactor of $\frac{1}{\sqrt{N!}}$ ensures normalization when the one electron orbitals are orthonormal.

The following theorem describes how the ground state eigenfunctions for the non-interacting many-electron system are formed from the one-electron eigenfunctions.

Theorem 2.2.1. *[Fri0X]*

1. *The lowest eigenvalue of* $H_0(N,Z) := -\frac{1}{2}\sum_{i=1}^N \Delta_i - \sum_{i=1}^N \frac{Z}{r_i}$ *on the space* $\Psi : (\mathbb{R}^3 \times \mathbb{Z}_2)^N \to \mathbb{C}$ *satisfying the antisymmetry condition is* $E = Z^2 \sum_{i=1}^N E_i$.

2. *The corresponding eigenspace is*

$$\mathit{Span}\big\{|\psi_1^Z \dots \psi_{d_-}^Z \psi_{i_1}^Z \dots \psi_{i_{N-d_-}}^Z\rangle \\ \qquad\qquad \mid i_1 < \dots < i_{N-d_-},\ i_1, \dots, i_{N-d_-} \in \{d_- + 1, \dots, d_+\}\big\},$$

where $\psi_i^Z := Z^{3/2}\psi_i(Zx, s)$ *and* $d_\pm$ *are given by*

$$d_- := \max\{d \le N \mid E_d < E_{d+1}\}, \quad d_+ := \min\{d \ge N \mid E_d < E_{d+1}\}.$$

This result shows that the eigenfunctions of the non-interacting Hamiltonian are linear combinations of Slater determinants formed from the first d_+ of the ψ_i^Z. Hence there is, in general, a degeneracy for the non-interacting case. The only uniquely-defined cases are when $E_N < E_{N+1}$.

The eigenvalues of the above ψ_i^Z are given by $E_n^Z = -\frac{Z^2}{2n^2}$, $n \in \mathbb{N}$ with corresponding $2n^2$-dimensional eigenspaces spanned by

$$V_n^Z := \{Z^{3/2}\psi_i(Zx, s) \mid \psi_i(x, s) \in V_n\}.$$

In particular, let d be the largest integer such that $N_* := \sum_{n=1}^{d} 2n^2 \le N$, then

$$\text{GS of } H_0(N, Z) = \text{Span}\{|\psi_1^Z \dots \psi_{N_*}^Z \chi_1^Z, \dots \chi_{N-N_*}^Z\rangle\}$$

where $\{\psi_1^Z, \dots, \psi_{N_*}^Z\} = \bigcup_{n=1}^{d} V_n^Z$ and $\{\chi_1^Z, \dots, \chi_{N-N_*}^Z\}$ is any $(N - N_*)$-element subset of V_{d+1}^Z. We see that, in the traditional chemistry terminology, d denotes the last completely filled electron shell and the χ_i^Z are in the valence shell. The dimension is then given by

$$\text{dim GS of } H_0 = \begin{cases} 1 & \text{if } N_* = N \\ \binom{2(d+1)^2}{N-N_*} & \text{if } N_* < N. \end{cases} \tag{2.2}$$

These values for the second row atoms are shown in Table 2.1.

Atom	He	Li	Be	B	C	N	O	F	Ne
N	2	3	4	5	6	7	8	9	10
dim GS	1	8	28	56	70	56	28	8	1

Table 2.1: Dimensions of degenerate H_0 ground states, as given by (2.2).

This degeneracy shows that the experimentally observed energy splittings between different angular momentum and spin states must result from the inter-electron term in the Hamiltonian. We introduce the notation $\psi_i(x) := \psi_i^Z(x, 1/2)$ and $\overline{\psi_i}(x) := \psi_i^Z(x, -1/2)$, taking the choice of Z to be implicit in the wavefunctions. We do not determine explicit forms for the orbitals here, instead considering the method in general; we will later study a number of basis sets.

2.2.4 First Order Perturbation Theory

Our motivation for the use of perturbation theory comes from the following result concerning the ground state energy. We introduce a small parameter $\lambda \in [0, 1]$ as a coupling constant for the interelectron term, i.e.

$$H_\lambda = H_0 + \lambda H',$$

where

$$H_0 = \sum_{i=1}^{N} \left(-\frac{1}{2}\Delta_{x_i} - \frac{Z}{|x_i|} \right), \quad H' = \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}.$$

Physically, this parameter corresponds to rescaling the equation in Z, which produces a $1/Z$ coefficient for H', i.e. making λ arbitrarily small is equivalent to fixing the number of electrons and letting the nuclear charge tend to infinity.

Denoting the bottom eigenvalue of H_λ by E_λ we will prove that, despite the infinite dimensional nature of the problem, E_λ can be asymptotically expanded at $\lambda = 0$ and the $\mathcal{O}(\lambda)$ term is defined by a finite-dimensional auxiliary problem.

Theorem 2.2.2 (First Order Energy Correction)**.** *Let*

$$H_\lambda = H_0 + \lambda H'$$

where H_0 and H_λ are linear operators such that H_0 has a spectral gap between the two lowest lying eigenvalues (the eigenspaces of which may be degenerate) and H' is non-negative. Let E_λ denote the lowest eigenvalue of H_λ, then

$$E_\lambda = E_0 + \lambda \tilde{E}_1 + \mathcal{O}(\lambda^2) \text{ as } \lambda \to 0,$$

where $E_0 = \langle H_0 \rangle_{\psi_0}$ and

$$\tilde{E}_1 := \min_{\psi_0 \in \text{ GS of } H_0} \frac{\langle H' \rangle_{\psi_0}}{\langle 1 \rangle_{\psi_0}}.$$

Proof. Upper Bound From the definition of E_λ (specifically it being the infimum over all ψ) we see that

$$\begin{aligned}\frac{E_\lambda - E_0}{\lambda} &\leq \frac{1}{\lambda}\left(\min_{\psi_0 \in \text{ GS of } H_0} \frac{\langle H_0 + \lambda H'\rangle_{\psi_0}}{\langle 1\rangle_{\psi_0}} - \frac{\langle H_0\rangle_{\psi_0}}{\langle 1\rangle_{\psi_0}}\right)\\ &= \frac{1}{\lambda}\left(\lambda \min_{\psi_0 \in \text{ GS of } H_0} \frac{\langle H'\rangle_{\psi_0}}{\langle 1\rangle_{\psi_0}}\right) = \tilde{E}_1.\end{aligned}$$

Lower Bound By Zhislin's Theorem [Zhi60] (or [RS78, Theorem XIII.7]), $\forall \lambda \in [0,1]$, the infimum in E_λ is attained. Let ψ_λ be a minimizer and let $\epsilon := \|P_{(\text{GS of } H_0)^\perp}\psi_\lambda\|_{L^2}$, i.e.

$$\psi_\lambda = \sqrt{1-\epsilon^2}\psi_0^{(\lambda)} + \epsilon\psi_\perp^{(\lambda)}$$

for some $\psi_0^{(\lambda)} \in$ GS of H_0 and some $\psi_\perp^{(\lambda)} \in$ orthogonal complement of GS of H_0.

Using this minimizer we have

$$\begin{aligned}\frac{E_\lambda - E_0}{\lambda} &= \frac{1}{\lambda}\Big(\langle H_0 + \lambda H'\rangle_{\sqrt{1-\epsilon^2}\psi_0^{(\lambda)}+\epsilon\psi_\perp^{(\lambda)}} - \langle H_0\rangle_{\psi_0}\Big)\\ &= \frac{1}{\lambda}\Big((1-\epsilon^2)\langle H_0\rangle_{\psi_0} + \epsilon^2\langle H_0\rangle_{\psi_\perp^{(\lambda)}} + \langle \lambda H'\rangle_{\sqrt{1-\epsilon^2}\psi_0^{(\lambda)}+\epsilon\psi_\perp^{(\lambda)}} - \langle H_0\rangle_{\psi_0}\Big)\\ &\geq \frac{\epsilon^2}{\lambda}(E^1 - E_0) + \langle H'\rangle_{\sqrt{1-\epsilon^2}\psi_0^{(\lambda)}+\epsilon\psi_\perp^{(\lambda)}}\\ &= \frac{\epsilon^2}{\lambda}\Delta + \langle H'\rangle_{\sqrt{1-\epsilon^2}\psi_0^{(\lambda)}+\epsilon\psi_\perp^{(\lambda)}} \qquad (2.3)\end{aligned}$$

The inequality comes from the fact that $\psi_\perp^{(\lambda)}$ is not in the ground state of H_0 and so must have expectation at least that of the first excited state, E^1, and having defined $\Delta := E^1 - E_0 > 0$. Expanding the second term in (2.3) we have

$$\langle H'\rangle_{\sqrt{1-\epsilon^2}\psi_0^{(\lambda)}+\epsilon\psi_\perp^{(\lambda)}} = (1-\epsilon^2)\langle H'\rangle_{\psi_0^{(\lambda)}} + 2\sqrt{1-\epsilon^2}\epsilon\mathrm{Re}\langle\psi_0^{(\lambda)}|H'|\psi_\perp^{(\lambda)}\rangle + \epsilon^2\langle H'\rangle_{\psi_\perp^{(\lambda)}}.$$

Clearly the last term is non-negative and bounding the other ϵ-dependent terms by constants independent of λ we have, substituting into (2.3),

$$\frac{E_\lambda - E_0}{\lambda} \geq \frac{\epsilon^2}{\lambda}\Delta + \langle H'\rangle_{\psi_0^{(\lambda)}} - C\epsilon$$

where, for example,

$$C := 2\sup_{\lambda\in[0,1]}\|H'\psi_0^{(\lambda)}\|_{L^2} + \sup_{\lambda\in[0,1]}\langle H'\rangle_{\psi_0^{(\lambda)}}.$$

We now wish to minimize this value over ϵ, which is the solution of

$$\frac{2\epsilon\Delta}{\lambda} - C = 0,$$

giving the minimum at $\epsilon = \frac{C\lambda}{2\Delta}$, with a value of $-\frac{C^2\lambda}{4\Delta}$.

Hence we have

$$\tilde{E_1} \geq \frac{E_\lambda - E_0}{\lambda} \geq \tilde{E_1} - \frac{C^2\lambda}{4\Delta}$$

and so

$$0 \geq E_\lambda - (E_0 + \lambda\tilde{E_1}) \geq -\frac{C^2\lambda^2}{4\Delta},$$

which proves the lemma. □

Note that this result differs from many in the standard literature, e.g [Kat67], in that we do not assume that the ground state is non-degenerate.

It follows that, at least for the ground state, up to first order we need only compute the two energy matrices $\langle H_0 \rangle$ and $\langle H' \rangle$ on the ground state of H_0. Further, by Theorem 2.2.1 and the orthonormality of Slater determinants, we see that $\langle H_0 \rangle$ will be a multiple of the identity and it remains to consider the matrix for H'. Of further interest is that the first order wavefunction Ψ is the solution of a PDE, namely $(H_0 - E_0)\Psi = 0$.

The idea that the inter-electron term provides the first order correction to the energy is not a new one. Sharma and Coulson [SC62] showed that, for a two electron Hamiltonian, assuming an expansion of both the energy and the wavefunctions in powers of $\frac{1}{Z}$, $E = E_0 + \frac{1}{Z}E_1 + \frac{1}{Z^2}E_2 + \cdots$ and $\Psi = \Psi_0 + \frac{1}{Z}\Psi_1 + \frac{1}{Z^2}\Psi_2 + \cdots$ with suitable orthogonality conditions on the Ψ_i, then E_0 is the energy of the non-interacting Hamiltonian and $E_1 = \langle \Psi_0 | \frac{1}{r_{12}} | \Psi_0 \rangle$. It is conceivable that this extends to any number of electrons, however, the main advantage of our proof is that we do not need to assume that asymptotic expansions of this kind exist. In fact, as can be seen from analysis of Møller-Plesset Perturbation results [LASS00], expansions of the energy and wavefunction are not well-behaved.

2.3 Simplification of the Energy Matrix Entries

2.3.1 Expected Values of the Hamiltonian Operator

Our goal now is to find the eigenvalues of $\langle \Psi | H' | \Psi \rangle$, where Ψ is in the ground state of H_0, the lowest of which will correspond to the first order perturbation theory ground state of H_λ.

The first problem of note is that the matrix elements are $3N$-dimensional spatial integrals and a sum over the $2N$ different spins. This is obviously not numerically computationally feasible when N is much greater than 5; taking only 10 points in each spatial dimension gives a total of 10^{15} points, which is beyond current numerics. Hence we wish to reduce the number of integrals required and also avoid any numerical integration in order to obtain closed-form analytic results with no numerical error.

Consider the action of the inter-electron repulsion term $H' = \sum_{i=1}^{N} \sum_{j>i}^{N} r_{ij}^{-1}$ where $r_{ij} = \frac{1}{|x_i - x_j|}$. It is clear that the operator H' acts on the individual positions of the electrons, and is a 2-particle operator. We may therefore use Slater's rules (Theorem 2.3.1) to split the integral into a sum of two-coordinate (6-dimensional plus spin) integrals.

The first step is to ensure that the two Slater determinants involved are in maximum coincidence, in other words, they differ in as few columns as possible. For example, given $|14\bar{1}2\rangle$ and $|\bar{1}132\rangle$ it would seem that the two determinants differ in all but the last column. However, noting that

$$|14\bar{1}2\rangle = -|1\bar{1}42\rangle = |1\bar{1}24\rangle \text{ and } |\bar{1}132\rangle = -|1\bar{1}32\rangle = |1\bar{1}23\rangle,$$

we see that they in fact differ only in one column.

2.3.2 Spin Orbital Integrals

We now introduce some standard integrals which will form the H' matrix elements. Consider a general two-particle operator

$$H' := \sum_{i=1}^{N} \sum_{j>i}^{N} h(x_i, s_i, x_j, s_j),$$

and denote, as in [SO96],

$$\begin{aligned}\langle ij|k\ell\rangle &= \langle \chi_i \chi_j | \chi_k \chi_\ell \rangle \\ &= \sum_{s_1, s_2 \in Z_2} \int_{\mathbb{R}^6} dx_1 dx_2 \chi_i^*(x_1, s_1) \chi_j^*(x_2, s_2) h(x_1, s_1, x_2, s_2) \chi_k(x_1, s_1) \chi_\ell(x_2, s_2)\end{aligned}$$

and

$$\begin{aligned}[][ij|k\ell] &= [\chi_i \chi_j | \chi_k \chi_\ell] \\ &= \sum_{s_1, s_2 \in Z_2} \int_{\mathbb{R}^6} dx_1 dx_2 \chi_i^*(x_1, s_1) \chi_j(x_1, s_1) h(x_1, s_1, x_2, s_2) \chi_k^*(x_2, s_2) \chi_\ell(x_2, s_2),\end{aligned}$$

where the χ_i's are simply a renumbering of the ψ_i's and $\overline{\psi_i}$'s by $\chi_{2n+1}(x,s) := \psi_n(x)$, $\chi_{2n}(x,s) := \overline{\psi_n}(x)$. We note that $\langle ij|k\ell\rangle = [ik|j\ell]$ and their use is a matter of taste as much as anything. We choose to use the square bracket notation, which follows the intuitive idea of repulsion between two charge distributions $\chi_i(x_1)\chi_k(x_1)$ and $\chi_j(x_2)\chi_\ell(x_2)$. Furthermore, for real orbitals, permuting the indices on either side of the bar does not change the integral. Finally, we will later see that the transition from spin integrals to purely spatial integrals is simpler in this notation.

Theorem 2.3.1 (Slater's Rules [SO96])**.** *Let $|K\rangle$ and $|L\rangle$ be two Slater determinants formed from N orthogonal one-electron orbitals, and let*

$$H' := \sum_{i=1}^{N}\sum_{j>i}^{N} h(x_i, s_i, x_j, s_j).$$

The matrix element $\langle K|H'|L\rangle$ is given by

1. $|K\rangle = |L\rangle = |\cdots mn\cdots\rangle$:

$$\langle K|H'|K\rangle = \frac{1}{2}\sum_{m}^{N}\sum_{n}^{N}\langle mn|mn\rangle - \langle mn|nm\rangle$$

2. $|K\rangle = |\cdots mn\cdots\rangle$, $|L\rangle = |\cdots pn\cdots\rangle$:

$$\langle K|H'|L\rangle = \frac{1}{2}\sum_{n}^{N}\langle mn|pn\rangle + \langle nm|np\rangle - \langle mn|np\rangle - \langle nm|pn\rangle$$

3. $|K\rangle = |\cdots mn\cdots\rangle$, $|L\rangle = |\cdots pq\cdots\rangle$:

$$\langle K|H'|L\rangle = \frac{1}{2}\big(\langle mn|pq\rangle + \langle nm|qp\rangle - \langle mn|qp\rangle - \langle nm|pq\rangle\big).$$

We may further simplify these expressions when $H' = V_{ee}$, the inter-electron repulsion, by noting that V_{ee} is symmetric under permutation of the x_i and that it commutes with the χ_i. Hence, by the identities above, we have

Corollary 2.3.2 (Slater's Rules for V_{ee} [SO96])**.** *When $H' = V_{ee} = \sum_{i=1}^{N}\sum_{j>i}^{N} r_{ij}^{-1}$ in Theorem 2.3.1 we have the simplified form*

1. $|K\rangle = |L\rangle = |\cdots mn\cdots\rangle$:

$$\langle K|V_{ee}|K\rangle = \frac{1}{2}\sum_{m}^{N}\sum_{n}^{N}[mm|nn] - [mn|nm]$$

2. $|K\rangle = |\cdots mn \cdots\rangle$, $|L\rangle = |\cdots pn \cdots\rangle$*:*

$$\langle K|V_{ee}|L\rangle = \sum_{n}^{N} [mp|nn] - [mn|np]$$

3. $|K\rangle = |\cdots mn \cdots\rangle$, $|L\rangle = |\cdots pq \cdots\rangle$*:*

$$\langle K|V_{ee}|L\rangle = [mp|nq] - [mq|np].$$

Note that there is no self-interaction in the identities above, the only term which could cause this is in the $|K\rangle = |L\rangle$ case and terms corresponding $m = n$ cancel.

Since V_{ee} is independent of spin, and the orbitals on either side of $[\cdot\cdot|\cdot\cdot]$ are functions of the same variable, any integral that has a single bar on either side of the square bracket notation gives zero after summing over spin. If the integral is non-zero the bars can be removed as the integrals are equivalent. (Note that the analogous conditions for the $\langle ij|kl\rangle$ notation are less simple.) Once again following [SO96], we introduce notation for the purely spatial integrals:

$$(ij|k\ell) := (\psi_i\psi_j|\psi_k\psi_\ell) = \int_{\mathbb{R}^6} dx_1 dx_2 \psi_i^*(x_1)\psi_j(x_1) r_{12}^{-1} \psi_k^*(x_2)\psi_\ell(x_2). \tag{2.4}$$

These integrals lead to the idea of Coulomb and exchange integrals, of the form $(ii|jj)$ and $(ij|ji)$ respectively. Coulomb integrals are due to the repulsive potentials between any two electrons, whereas exchange integrals are a purely quantum effect which occur only between orbitals having the same spin.

2.3.3 Evaluation of the Integrals

The next stage is to analytically evaluate these 6-dimensional spatial integrals, for which a number of methods already exist in the literature. Perhaps the most common method is to expand both the products of wavefunctions and the r_{12}^{-1} term in Legendre functions,

$$r_{12}^{-1} = \sum_{k=0}^{\infty} \frac{r_<^k}{r_>^{k+1}} P_k(\cos\theta_{12})$$

where $r_<$ and $r_>$ are the shorter and longer of r_1 and r_2 respectively and θ_{12} is the angle between the two electrons.

The orthogonality property of the Legendre functions is then used to reduce the calculation to a finite sum of two-dimensional integrals, which must then be split along the line $r_1 = r_2$ to utilise $r_<$ and $r_>$. The simplest example of this method is the Helium atom which is a standard text books calculation (e.g. [Sch01]).

This method is not easily extended to wavefunctions that, for example, include powers of r_{12} explicitly, and is also relatively complicated when computing integrals for higher n and ℓ Hydrogen-like orbitals.

Another method, which does not seem to be widely used, is that of Calais and Löwdin [CL62] which applies to integrals containing more general powers of r_{12}. This method first integrates over the angular terms, using spherical harmonics and then uses a specific choice of coordinates to enable the radial integral to be split into a product of three independent integrals. Interestingly, this method shows directly that the integrals required are rational functions of Z. The disadvantage of this method is that it applies only to Slater type orbitals with additional powers of r_{12} (due to the choice of coordinate system) and it seems unlikely that there are analogous coordinate systems for other types of orbital.

Ideally we wish to use a method that is adaptable to a wide range of basis sets and requires only simple calculations, preferably avoiding infinite expansions and the like. In particular, we would like a method that works equally well for Slater type orbitals (the ‘natural’ choice for atoms) and also for Gaussians, which are widely used in numerical molecular calculations due to relatively easy evaluation of multi-center integrals.

Due to the physical restrictions on the one-electron wavefunctions (finite modulus at all points, square integrable, decay exponentially), the spatial part of the product of two one-electron wavefunctions lies in $L^1(\mathbb{R}^3)$ and hence we may compute the Fourier transform.

Definition 2.3.3. *Let* $f \in L^1(\mathbb{R}^n)$, *then the* Fourier transform *of* f, *denoted* $\hat{f}$ *is given by*

$$\hat{f}(k) := \int_{\mathbb{R}^n} f(x) e^{-ik\cdot x} dx.$$

However, the Coulomb repulsion $1/|x|$ does not lie in $L^1(\mathbb{R}^3)$ (nor in $L^2(\mathbb{R}^3)$.) Nevertheless, we claim that

Lemma 2.3.4. *Let* $f, g \in L^1(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ *such that* $\hat{f}, \hat{g} \in L^1(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$, *then*

$$\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} dx dy \frac{1}{|x-y|} f^*(x) g(y) = \frac{1}{2\pi^2} \int_{\mathbb{R}^3} dk \frac{1}{|k|^2} \hat{f}^*(k) \hat{g}(k).$$

Proof If all functions were in L^1 then this would simply be a consequence of Plancherel's theorem and the fact that the Fourier transform of a convolution is the product of the Fourier transforms. However, since the Coulomb repulsion term is not in $L^1(\mathbb{R}^3)$ we give a rigorous proof in Appendix A. □

Note that, although the fact that such integrals can be calculated using Fourier transforms is not new, this proof is in contrast to 'proofs' readily available in the literature, e.g. [SO96], which simply state that the Fourier transform of $1/|x|$ is $4\pi/|k|^2$.

Now, under the these assumptions (which hold for the orbitals in (2.7) and their Fourier transforms in Table 2.2), we have

$$(ij|k\ell) =: \int_{\mathbb{R}^3} \int (\mathbb{R}^3) dx dy \frac{1}{|x-y|} f^*(x) g(y) = \frac{1}{2\pi^2} \int_{\mathbb{R}^3} dk \frac{1}{|k|^2} \hat{f}^*(k) \hat{g}(k). \tag{2.5}$$

As an aside, it is interesting to note that this method gives an easy proof that the exchange integrals are positive. The same proof applies for any potential which has a positive Fourier transform.

We have now simplified the original $3N$-dimensional integrals into a linear combination of 3-dimensional integrals. However, from (2.2) and the resulting Table 2.1, we see that this still leads to a very large number of integrals. The next stage will be to simplify the V_{ee} matrix in order to reduce the number of integrals necessary. However, before we do this we investigate the matrix elements in a specific one-electron basis.

2.4 Explicit One-electron Basis

In order to obtain numerical results using this method, we must choose a suitable one-electron basis from which to form the Slater determinants. We begin by considering the general form for the Hydrogen atom eigenfunctions (see e.g. [Gri95]):

$$\psi_{n,l,m}(r,\theta,\phi) = R_{n,\ell}(r) Y_{\ell,m}(\theta,\phi) \tag{2.6}$$

where

$$R_{n,\ell}(r) := \left(\left(\frac{2}{n}\right)^3 \frac{(n-\ell-1)!}{2n[(n+\ell)!]} \right)^{1/2} e^{-r/n} \left(\frac{2r}{n}\right)^{\ell} L^{2\ell+1}_{n-\ell-1}\left(\frac{2r}{n}\right),$$

where $L^k_n(x)$ is a generalized Laguerre polynomial [AS72]:

$$L^k_n(x) := \sum_{m=0}^{n} (-1)^m \frac{(n+k)!}{(n-m)!(k+m)!m!} x^m$$

and $Y_{\ell,m}(\theta,\phi)$ is a spherical harmonic [AS72]:

$$Y_{\ell,m}(\theta,\phi) := \left(\frac{(2\ell+1)!(\ell+m)!}{4\pi(\ell+m)!}\right)^{1/2} P_\ell^m(\cos\theta)e^{-im\phi}$$

with $P_\ell^m(x)$ an associated Legendre function

$$P_\ell^m(x) := \frac{(-1)^m}{2^\ell \ell!}(1-x^2)^{m/2}\frac{d^{\ell+m}}{dx^{\ell+m}}(x^2-1)^\ell.$$

The first five such Hydrogen orbitals are given by

$$\begin{aligned}
\psi_{1,0,0} &= \frac{1}{\sqrt{\pi}}e^{-r}, \\
\psi_{2,0,0} &= \frac{1}{\sqrt{8\pi}}\left(1-\frac{r}{2}\right)e^{-r/2}, \\
\psi_{2,1,0} &= \frac{1}{\sqrt{32\pi}}r\cos\theta e^{-\frac{r}{2}} = \frac{1}{\sqrt{32\pi}}x_3e^{-r/2}, \\
\psi_{2,1,1} &= \frac{1}{\sqrt{64\pi}}r\sin\theta e^{-r/2}e^{i\phi}, \text{ and} \\
\psi_{2,1,-1} &= \frac{1}{\sqrt{64\pi}}r\sin\theta e^{-r/2}e^{-i\phi}.
\end{aligned}$$

We begin by restricting our basis to orbitals with $n = 1$ and 2. Along with the result of Theorem 2.2.1, this can be justified by considering the expected value of H_0 for orbitals with higher principal quantum number. The relevant value is $E = -\frac{1}{2n^2}$ and we see that the energy of the higher n terms will be much too high be the ground state. Furthermore, for the same reason, we restrict to Slater determinants containing both $\psi_{1,0,0}$ orbitals.

This restriction is equivalent to considering the operator PHP where H is the Hamiltonian in (2.1) and P is the orthogonal projection onto the space of Slater determinants formed from both $\psi_{1,0,0}$ orbitals and $N-2$ of $\psi_{2,0,0}$ to $\overline{\psi_{2,1,-1}}$.

The remaining freedom in our basis concerns linear combinations of the p-eigenfunctions, which are still angular momentum eigenfunctions. One option would be to construct three $2p$-orbitals that are each eigenfunctions of L_3 (with eigenvalues -1, 0 and $+1$); this would allow easier computation of the angular momentum eigenspaces. However, we choose to use a spatially symmetric set of $2p$-orbitals. This is beneficial as it reduces the number of different integrals required (by rotational symmetry) and also simplifies the form of $\underline{L}^2$ eigenfunctions.

We define two new eigenfunctions:

$$\psi_{2p_x} = \frac{\psi_{2,1,1} - \psi_{2,1,-1}}{i\sqrt{2}} = \frac{1}{\sqrt{32\pi}} x_1 e^{-\frac{r}{2}}$$
$$\psi_{2p_y} = \frac{\psi_{2,1,1} + \psi_{2,1,-1}}{\sqrt{2}} = \frac{1}{\sqrt{32\pi}} x_2 e^{-\frac{r}{2}}.$$

Finally we use Theorem 2.2.1, and in particular $\psi_i^Z = Z^{3/2}\psi_i(Zx, s)$, which allows us to write these orbitals for general Z:

$$\begin{aligned}
\psi_1 &:= \psi_{1s} := \frac{Z^{\frac{3}{2}}}{\sqrt{\pi}} e^{-Zr}, \\
\psi_2 &:= \psi_{2s} := \frac{Z^{\frac{3}{2}}}{\sqrt{8\pi}} \big(1 - Zr/2\big) e^{-\frac{Zr}{2}}, \\
\psi_3 &:= \psi_{2p_3} := \frac{Z^{\frac{5}{2}}}{\sqrt{32\pi}} x_3 e^{-\frac{Zr}{2}}, \\
\psi_4 &:= \psi_{2p_1} := \frac{Z^{\frac{5}{2}}}{\sqrt{32\pi}} x_1 e^{-\frac{Zr}{2}}, \text{ and} \\
\psi_5 &:= \psi_{2p_2} := \frac{Z^{\frac{5}{2}}}{\sqrt{32\pi}} x_2 e^{-\frac{Zr}{2}},
\end{aligned} \tag{2.7}$$

which are taken to have spin up (or 1/2), the same spatial orbitals with spin down (or $-1/2$), are denoted by $\overline{\psi_i}$. These orbitals are all orthogonal (either through the spatial or spin parts) and normalized.

2.5 Evaluation of Coulomb and Exchange Integrals

2.5.1 Fourier Transforms

The next task it to explicitly compute the Fourier transforms of products of one-electron wavefunctions involved in the Coulomb and exchange integrals. Recall that the general integrals from Slater's rules will be of the form $(ij|kl)$, and hence we need to compute the Fourier transforms for all pairs of spatial orbitals $\psi_i\psi_j$, $i, j = 1 \dots 5$.

We begin with the simplest case, $(\psi_1\psi_1)(x)$, which is given (up to normalization) by $e^{-\lambda|x|}$. We note that λ will vary with Z and hence compute a general form for this Fourier transform.

Lemma 2.5.1. *The Fourier transform of* $e^{-\lambda|x|}$ *is* $\frac{8\lambda\pi}{(\lambda^2+|k|^2)^2}$.

Proof The Fourier transform is defined to be $\widehat{e^{-\lambda|x|}} = \int_{\mathbb{R}^3} e^{-\lambda|x|}e^{-ik\cdot x}dx$. If $k = 0$, converting to spherical polar coordinates, we have

$$\widehat{e^{-\lambda|x|}} = \int_0^\infty \int_0^\pi \int_0^{2\pi} e^{-\lambda r} r^2 \sin\theta d\phi d\theta dr = 4\pi \int_0^\infty e^{-\lambda r} r^2 dr = 4\pi \frac{2}{\lambda^3} = \frac{8\pi\lambda}{\lambda^4},$$

where the penultimate equality follows from integrating by parts twice. For $k \neq 0$,

$$\begin{aligned}\widehat{e^{-\lambda|x|}} &= \int_0^\infty \int_0^\pi \int_0^{2\pi} e^{-\lambda r} e^{-i|k|r\cos\theta} r^2 \sin\theta d\phi d\theta dr \\ &= 2\pi \int_0^\infty \int_{-1}^1 r^2 e^{-\lambda r} e^{-i|k|rt} dt dr = \frac{4\pi}{|k|} \int_0^\infty r e^{-\lambda r} \left(\frac{e^{i|k|r} - e^{-i|k|r}}{2i} \right) dr \\ &= \frac{4\pi}{|k|} \int_0^\infty \left(\frac{r e^{r(i|k|-\lambda)}}{2i} - \frac{r e^{r(-i|k|-\lambda)}}{2i} \right) dr\end{aligned}$$

Integrating by parts we have

$$\widehat{e^{-\lambda|x|}} = -\frac{4\pi}{|k|} \int_0^\infty \left(\frac{e^{r(i|k|-\lambda)}}{2i(i|k|-\lambda)} - \frac{e^{r(-i|k|-\lambda)}}{2i(-i|k|-\lambda)} \right) dr,$$

which trivially integrates to give

$$\widehat{e^{-\lambda|x|}} = -\frac{4\pi}{|k|} \left[\frac{e^{r(i|k|-\lambda)}}{2i(i|k|-\lambda)^2} - \frac{e^{r(-i|k|-\lambda)}}{2i(-i|k|-\lambda)^2} \right]_0^\infty = \frac{8\pi\lambda}{(\lambda^2+|k|^2)^2}.$$

□

Applying this result to (11|11) with $\lambda = 2Z$, gives

$$\psi_1\psi_1 = \frac{1}{\pi} Z^3 e^{-2Z|x|} \Rightarrow \widehat{\psi_1\psi_1} = \frac{16Z^4}{(4Z^2+|k|^2)^2}.$$

In order to calculate the remaining Fourier transforms we note that there are a number of standard forms, namely $x_1^{n_1} x_2^{n_2} x_3^{n_3} |x|^{n_4} e^{-\lambda|x|}, n_i \in \mathbb{N}$. For our functions the highest degree $\sum_{i=1}^4 n_i$ will be equal to two.

The $|x|^n$ term can be calculated from the result in Lemma 2.5.1 by noting that differentiation with respect to the parameter λ commutes with the Fourier transform. The other terms can be calculated using the standard rules for differentiation of Fourier transforms. We then have the following results:

Lemma 2.5.2.

$$\begin{aligned}
\widehat{|x|e^{-\lambda|x|}} &= \frac{32\lambda^2\pi}{(\lambda^2+|k|^2)^3} - \frac{8\pi}{(\lambda^2+|k|^2)^2},\\
\widehat{|x|^2e^{-\lambda|x|}} &= \frac{192\lambda^3\pi}{(\lambda^2+|k|^2)^4} - \frac{96\lambda\pi}{(\lambda^2+|k|^2)^3},\\
\widehat{x_je^{-\lambda|x|}} &= -\frac{32i\lambda\pi k_j}{(\lambda^2+|k|^2)^3},\\
\widehat{x_j^2e^{-\lambda|x|}} &= \frac{32\lambda\pi}{(\lambda^2+|k|^2)^3} - \frac{192\lambda\pi k_j^2}{(\lambda^2+|k|^2)^4},\\
\widehat{x_\ell x_je^{-\lambda|x|}} &= -\frac{192\lambda\pi k_jk_\ell}{(\lambda^2+|k|^2)^4} \quad (j\neq\ell),\ \textit{and}\\
\widehat{|x|x_je^{-\lambda|x|}} &= \frac{32i\pi k_j}{(\lambda^2+|k|^2)^3} - \frac{192i\lambda^2\pi k_j}{(\lambda^2+|k|^2)^4}.
\end{aligned}$$

Proof We begin with the identity

$$|x|e^{-\lambda|x|} = -\frac{d}{d\lambda}e^{-\lambda|x|},$$

and note that differentiation with respect to λ commutes with the Fourier transform. We then have

$$\widehat{|\cdot|e^{-\lambda|\cdot|}}(k) = -\frac{d}{d\lambda}\widehat{e^{-\lambda|\cdot|}}(k),$$

and using the normal rules for differentiation we find

$$\widehat{|\cdot|e^{-\lambda|\cdot|}}(k) = -\frac{d}{d\lambda}\left(\frac{8\lambda\pi}{(\lambda^2+|k|^2)^2}\right) = \frac{32\lambda^2\pi}{(\lambda^2+|k|^2)^3} - \frac{8\pi}{(\lambda^2+|k|^2)^2}.$$

Similarly, we have that

$$\begin{aligned}
\widehat{|x|^2e^{-\lambda|x|}} &= \frac{d^2}{d\lambda^2}e^{-\lambda|x|} = \frac{d}{d\lambda}\left(\frac{8\pi}{(\lambda^2+|k|^2)^2} - \frac{32\lambda^2\pi}{(\lambda^2+|k|^2)^3}\right)\\
&= \frac{192\pi\lambda^3}{(\lambda^2+|k|^2)^4} - \frac{96\pi\lambda}{(\lambda^2+|k|^2)^3}.
\end{aligned}$$

For the second set of Fourier transforms, we recall the well known differentiation identities for Fourier transforms of a function $f \in L^1(R^n)$

$$\frac{d}{dk_j}\widehat{f}(k) = -i\widehat{x_jf}(k) \ \Rightarrow\ \widehat{x_jf}(k) = i\frac{d}{dk_j}\widehat{f}(k) \text{ and } \widehat{x_\ell x_jf}(k) = -\frac{d^2}{dk_\ell dk_j}\widehat{f}(k).$$

Taking $f(x) = e^{-\lambda|x|}$, we have $\widehat{f}(k) = \frac{8\pi\lambda}{(\lambda^2+|k|^2)^2}$ and noting that

$$\frac{d}{dk_j}|k| = \frac{d}{dk_j}(k_1^2+k_2^2+k_3^2)^{1/2} = \frac{k_j}{|k|},$$

gives

$$(\widehat{x_j e^{-\lambda|x|}}) = i\frac{d}{dk_j}\widehat{f}(k) = -\frac{32\pi i\lambda k_j}{(\lambda^2+|k|^2)^3}.$$

Similarly, we have

$$\begin{aligned}\widehat{x_j^2 e^{-\lambda|x|}} &= -\frac{d^2}{dk_j^2}\widehat{f}(k) = \frac{32\pi\lambda}{(\lambda^2+|k|^2)^3} - \frac{192\pi\lambda k_j^2}{(\lambda^2+|k|^2)^4}, \text{ and}\\ \widehat{x_j x_\ell e^{-\lambda|x|}} &= -\frac{d^2}{dk_j dk_\ell}\widehat{f}(k) = -\frac{192\pi\lambda k_j k_\ell}{(\lambda^2+|k|^2)^4}.\end{aligned}$$

It remains to consider

$$(\widehat{|x|x_j e^{-\lambda|x|}}) = -\frac{d}{d\lambda}(\widehat{x_j e^{-\lambda|x|}}) = \frac{d}{d\lambda}\frac{32i\pi\lambda k_j}{(\lambda^2+|k|^2)^3} = \frac{32ik_j\pi}{(\lambda^2+|k|^2)^3} - \frac{192ik_j\pi\lambda^2}{(\lambda^2+|k|^2)^4}.$$

This final result can equivalently be obtained by considering the differentiation with respect to k_j of the Fourier transform of $|x|e^{-\lambda|x|}$ which, of course, leads to the same result. ☐

We are now in a position to calculate the remaining orbital product Fourier transforms and begin with the case of two $2s$-orbitals:

$$\psi_2\psi_2 = \frac{Z^3}{8\pi}\big(1-\tfrac{1}{2}Z|x|\big)^2 e^{-Z|x|} = \frac{Z^3}{8\pi}\big(1 - Z|x| + \tfrac{1}{4}Z^2|x|^2\big)e^{-Z|x|}.$$

Using the linearity of the Fourier transform, the results from Lemmas 2.5.1 and 2.5.2, and noting that in this case $\lambda = Z$, we have

$$\begin{aligned}(\widehat{\psi_2\psi_2}) &= \frac{1}{8\pi}Z^3\bigg[\frac{8\pi Z}{(Z^2+|k|^2)^2} - Z\bigg(\frac{32Z^2\pi}{(Z^2+|k|^2)^3} - \frac{8\pi}{(Z^2+|k|^2)^2}\bigg)\\ &\qquad\qquad + \frac{Z^2}{4}\bigg(\frac{192\pi Z^3}{(Z^2+|k|^2)^4} - \frac{96\pi Z}{(Z^2+|k|^2)^3}\bigg)\bigg]\\ &= \frac{2Z^4}{(Z^2+|k|^2)^2} - \frac{7Z^6}{(Z^2+|k|^2)^3} + \frac{6Z^8}{(Z^2+|k|^2)^4}.\end{aligned}$$

For purely s-states, the remaining Fourier transform we need is that of $\psi_1\psi_2 = \frac{1}{\sqrt{8\pi}}Z^3(1-Z|x|/2)e^{-3Z|x|/2}$, which is required for the exchange integrals. Note that, for this case, $\lambda = 3Z/2$, and we have

$$\begin{aligned}\widehat{(\psi_1\psi_2)} &= \frac{Z^3}{\sqrt{8}\pi}\left[\frac{8\pi(\frac{3}{2}Z)}{\big((\frac{3}{2}Z)^2+|k|^2\big)^2} - \frac{Z}{2}\left(\frac{32(\frac{3}{2}Z)^2\pi}{\big((\frac{3}{2}Z)^2+|k|^2\big)^3} - \frac{8\pi}{\big((\frac{3}{2}Z)^2+|k|^2\big)^2}\right)\right]\\ &= \frac{4\sqrt{2}Z^4}{\big((\frac{3}{2}Z)^2+|k|^2\big)^2} - \frac{9\sqrt{2}Z^6}{\big((\frac{3}{2}Z)^2+|k|^2\big)^3}.\end{aligned}$$

We now move on to considering cases which involve p-orbitals and by spatial symmetry need only consider a general p_j. For this we have $\psi_{p_j}\psi_{p_j} = \frac{1}{32\pi}Z^5x_j^2e^{-Z|x|}$ and so, with $\lambda = Z$,

$$\begin{aligned}\widehat{(\psi_{p_j}\psi_{p_j})} &= \frac{Z^5}{32\pi}\left(\frac{32\pi Z}{(Z^2+|k|^2)^3} - \frac{192\pi Z k_j^2}{(Z^2+|k|^2)^4}\right)\\ &= \frac{Z^6}{(Z^2+|k|^2)^3} - \frac{6Z^6k_j^2}{(Z^2+|k|^2)^4}.\end{aligned}$$

For the next integral, we consider a mixed term of ψ_1 and ψ_{p_j}, giving $\psi_1\psi_{p_j} = \frac{1}{\sqrt{32}\pi}Z^4x_je^{-3Z|x|/2}$ and so, with $\lambda = 3Z/2$,

$$\widehat{(\psi_1\psi_{p_j})} = \frac{1}{\sqrt{32}\pi}Z^4\left(-\frac{32i\pi(\frac{3}{2}Z)k_j}{\big((\frac{3}{2}Z)^2+|k|^2\big)^3}\right) = -\frac{6\sqrt{2}iZ^5k_j}{\big((\frac{3}{2}Z)^2+|k|^2\big)^3}.$$

We next consider the case $\psi_2\psi_{p_j} = \frac{Z^4}{16\pi}(1-Z|x|/2)x_je^{-Z|x|}$, giving

$$\begin{aligned}\widehat{(\psi_2\psi_{p_j})} &= \frac{1}{16\pi}Z^4\left[-\frac{32i\pi Zk_j}{(Z^2+|k|^2)^3} - \frac{Z}{2}\left(\frac{32ik_j\pi}{(Z^2+|k|^2)^3} - \frac{192ik_j\pi Z^2}{(Z^2+|k|^2)^4}\right)\right]\\ &= \frac{6Z^7ik_j}{(Z^2+|k|^2)^4} - \frac{3Z^5ik_j}{(Z^2+|k|^2)^3}.\end{aligned}$$

The final case is when we have two different p-orbitals, leading to $\psi_{p_j}\psi_{p_\ell} = \frac{Z^5}{32\pi}x_jx_\ell e^{-Z|x|}$, which has $\lambda = Z$, giving

$$\widehat{(\psi_{p_j}\psi_{p_\ell})} = \frac{Z^5}{32\pi}\left(-\frac{192\pi Zk_jk_\ell}{(Z^2+|k|^2)^4}\right) = -\frac{6k_jk_\ell Z^6}{(Z^2+|k|^2)^4}.$$

For ease of reference we collect the above results in Table 2.2. These Fourier transforms appear to be a new result and are not found in any of the standard literature.

Function	Fourier Transform
$\psi_1\psi_1$	$\frac{16Z^4}{(4Z^2+\lvert k\rvert^2)^2}$
$\psi_2\psi_2$	$\frac{2Z^4}{(Z^2+\lvert k\rvert^2)^2} - \frac{7Z^6}{(Z^2+\lvert k\rvert^2)^3} + \frac{6Z^8}{(Z^2+\lvert k\rvert^2)^4}$
$\psi_1\psi_2$	$\frac{4\sqrt{2}Z^4}{((\frac{3}{2}Z)^2+\lvert k\rvert^2)^2} - \frac{9\sqrt{2}Z^6}{((\frac{3}{2}Z)^2+\lvert k\rvert^2)^3}$
$\psi_{p_j}\psi_{p_j}$	$\frac{Z^6}{(Z^2+\lvert k\rvert^2)^3} - \frac{6Z^6k_j^2}{(Z^2+\lvert k\rvert^2)^4}$
$\psi_1\psi_{p_j}$	$-\frac{6\sqrt{2}iZ^5k_j}{((\frac{3}{2}Z)^2+\lvert k\rvert^2)^3}$
$\psi_2\psi_{p_j}$	$\frac{6Z^7ik_j}{(Z^2+\lvert k\rvert^2)^4} - \frac{3Z^5ik_j}{(Z^2+\lvert k\rvert^2)^3}$
$\psi_{p_j}\psi_{p_\ell}$	$-\frac{6k_jk_\ell Z^6}{(Z^2+\lvert k\rvert^2)^4}$

Table 2.2: Fourier Transforms necessary for Coulomb and Exchange Integrals, using the one-electron orbitals in (2.7).

2.5.2 Coulomb and Exchange Integrals

We are now ready to compute explicit forms for the required Coulomb and exchange integrals in Theorem 2.3.1. At the present moment we have not computed the necessary eigenspaces that will allow us to determine the required integrals. However, for completeness, we compute all Coulomb and exchange integrals involving $\psi_1 \dots \psi_5$. We will later see that all other integrals generated by Slater's rules are zero for our choice of one electron basis and restriction on the form of the Slater determinants.

Recalling the integral notation from (2.4) we begin with $(11|11)$, which is the only integral required for $|\psi_1\overline{\psi_1}\rangle$, the ground state for the Helium atom. This integral is well known in the literature and has the value $\frac{5}{8}Z$ [Sch01].

Recalling the Fourier transform of $\psi_1\psi_1$ given in (2.2) and the general form of the integral (2.5), we have

$$(11|11) = \frac{1}{2\pi^2}\int_{\mathbb{R}^3} \frac{1}{|k|^2}\left(\frac{16Z^4}{(4Z^2+|k|^2)^2}\right)^2 dk.$$

We now convert to spherical polar coordinates and integrate over the angular terms, giving

$$\begin{aligned}(11|11) &= \frac{1}{2\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \frac{1}{r^2}\left(\frac{16Z^4}{(4Z^2+r^2)^2}\right)^2 r^2 \sin\theta dr d\theta d\phi \\ &= \frac{512Z^8}{\pi}\int_0^{\infty} \frac{1}{(4Z^2+r^2)^4} dr.\end{aligned}$$

The final one-dimensional integral can be easily evaluated using a computer package such as Maple or Mathematica. Alternatively, if we desire a purely pen-and-paper method, then these integrals can be evaluated using complex contour integration. However, evaluating the one-dimensional integrals gives us little useful insight so we use the electronic method. In this case we find

$$(11|11) = \frac{512Z^8}{\pi}\frac{5\pi}{4096Z^7} = \frac{5}{8}Z.$$

The remaining integrals are calculated in the same way, although any involving p-orbitals will have non-trivial angular integrals. As an example we compute $(33|44)$ and the other integral calculations, along with useful standard angular integrals, are given in Appendix B. The resulting Coulomb and exchange integrals are given in Table 2.3.

These integral values are not new, in that they can be reconstructed from the values of the F and G integrals of [LS60]. These F and G integrals (see [CS35]) are purely radial, having integrated over the angular terms, and thus to reconstruct our results it is necessary only to determine the angular contributions.

For $(33|44)$, by rotational symmetry of the p-orbitals it suffices to consider p_1 and p_3, where $k_1 = r\sin\theta\sin\phi$ and $k_3 = r\cos\theta$.

$$\begin{aligned}
(33|44) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\left(\frac{Z^6}{(Z^2+|k|^2)^3} - \frac{6Z^6k_3^2}{(Z^2+|k|^2)^4}\right) \\
&\qquad\qquad \times\left(\frac{Z^6}{(Z^2+|k|^2)^3} - \frac{6Z^6k_1^2}{(Z^2+|k|^2)^4}\right)dk \\
&= \frac{1}{2\pi^2}\int_0^\infty\int_0^\pi\int_0^{2\pi}\left(\frac{Z^6}{(Z^2+r^2)^3} - \frac{6Z^6r^2\cos^2\theta}{(Z^2+r^2)^4}\right) \\
&\qquad\qquad \times\left(\frac{Z^6}{(Z^2+r^2)^3} - \frac{6Z^6r^2\sin^2\theta\sin^2\phi}{(Z^2+|k|^2)^4}\right)\sin\theta d\phi d\theta dr.
\end{aligned}$$

Performing the angular integrals using the standard forms in Appendix B and the radial integral using Maple gives

$$\begin{aligned}
(33|44) &= \frac{1}{2\pi^2}\int_0^\infty\int_0^\pi\left(\frac{2Z^{12}}{(Z^2+r^2)^6} - \frac{6Z^{12}r^2(2\cos^2\theta+\sin^2\theta)}{(Z^2+r^2)^7}\right. \\
&\qquad\qquad \left. + \frac{36Z^{12}r^4\cos^2\theta\sin^2\theta}{(Z^2+r^2)^8}\right)\sin\theta d\theta dr \\
&= \frac{1}{2\pi}\int_0^\infty\left(\frac{4Z^{12}}{(Z^2+r^2)^6} - \frac{16Z^{12}r^2}{(Z^2+r^2)^7} + \frac{48Z^{12}r^4}{5(Z^2+r^2)^8}\right)dr \\
&= \frac{1}{2\pi}\frac{447Z\pi}{1280} = \frac{447}{2560}Z.
\end{aligned}$$

Integral	(11|11)	(11|22)	(12|21)	(22|22)	(11|33)	(13|31)
Result	$\frac{5}{8}Z$	$\frac{17}{81}Z$	$\frac{16}{729}Z$	$\frac{77}{512}Z$	$\frac{59}{243}Z$	$\frac{112}{6561}Z$
Integral	(22|33)	(23|32)	(33|33)	(33|44)	(34|43)	
Result	$\frac{83}{512}Z$	$\frac{15}{512}Z$	$\frac{501}{2560}Z$	$\frac{447}{2560}Z$	$\frac{27}{2560}Z$	

Table 2.3: Values of Coulomb and Exchange Integrals for the scaled atomic wavefunctions given in (2.7).

As mentioned previously, all of the integrals are rational multiples of Z, removing any numerical errors from the calculations. It it also immediately noticeable that the exchange terms are much smaller than the Coulomb terms. However, as we will see later, the exchange terms play a very important role in energy level splitting.

We now show that all other possible integrals generated by Slater's rules given in Corollary 2.3.2 are zero. One reason for this is that each Slater determinant is constrained to contain ψ_{1s} and $\overline{\psi_{1s}}$, if this were not the case then we could have terms such as $(1s2s|p_ip_i)$ which are non-zero.

There are three cases to consider, the first is $(mp|nn)$ where $m, p \neq 1s, m \neq n \neq p$, which leads to $(2sp_i|p_jp_j)$, $(2sp_i|1s1s)$, $(p_ip_j|p_kp_k)$, $(p_ip_j|1s1s)$ and $(p_ip_j|2s2s)$. The second case is $(mn|np)$ where $m, p \neq 1s, m \neq n \neq p$ which gives $(2s1s|1sp_i)$, $(p_i1s|1sp_j)$, $(p_i2s|2sp_j)$ and $(2sp_i|p_ip_j)$. Finally we have the case $(mn|pq)$ where $m \neq n \neq p \neq q$, $m, n, p, q \neq 1s$, which results in $(2sp_i|p_jp_k)$.

By the results of Table 2.2 the Fourier transforms satisfy $\widehat{sp_i} = f_1(|k|)k_i$, $\widehat{p_ip_j} = f_2(|k|)k_ik_j$ and $\widehat{p_ip_i} = f_3(|k|) + f_4(|k|)k_j^2$. By (2.5) we see that the only angular dependence comes from these terms, and noting that $\int_0^{2\pi} \sin\phi d\phi = \int_0^{2\pi} \sin\phi\cos\phi d\phi = 0$, it is trivial that all of the above integrals are zero.

Chapter 3

Simplification of the Energy Matrix

3.1 Angular Momentum and Spin Eigenspaces

We are now in a position to evaluate the V_{ee} matrix in the basis chosen in Section 2.4, using the integrals derived in Section 2.5.2. However, we first aim to use extra information about the chosen wavefunctions to further simplify the matrix and reduce the number of matrix entries we need to calculate.

Our aim is to show that a suitable basis of linear combinations of Slater determinants can be chosen which results in the V_{ee} matrix having block diagonal form. In order to do this, we wish to find operators which commute, and thus have simultaneous eigenspaces with, H. The important point is that these eigenspaces will be much smaller than the ground state of non-interacting Hamiltonian (see Table 2.1). The two most useful physical properties for this are angular momentum $\underline{L}$ and spin $\underline{S}$, which are vector operators given by

$$\underline{L} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} := \frac{1}{i} \begin{pmatrix} x_2\partial_3 - x_3\partial_2 \\ x_3\partial_1 - x_1\partial_3 \\ x_1\partial_2 - x_2\partial_1 \end{pmatrix}, \tag{3.1}$$

$$\underline{S} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}; \quad S_1 := \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_2 := \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_3 := \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.2}$$

where S_j are the Pauli spin matrices. It is well known (see e.g. [Sch01]) that the largest set of commuting operators formed from the L_i and S_i is $\{\underline{L}^2, \underline{S}^2, L_3, S_3\}$, where taking the third component is merely a convention. $\underline{L}^2$ and $\underline{S}^2$ are known as the total (spatial) angular momentum and total spin operators respectively and are defined as

$$\underline{L}^2 := L_1^2 + L_2^2 + L_3^2 \text{ and } \underline{S}^2 := S_1^2 + S_2^2 + S_3^2.$$

We now use the nontrivial fact (see Lemma 3.1.4 below) that for $V :=$ span{GS of H_0} (as given in Theorem 2.2.1) the operator PHP, where P is the orthogonal projection onto V, commutes with $\underline{L}$ and $\underline{S}$. Hence the eigenspace of PHP on V can be factored into simultaneous eigenspaces of $\underline{S}^2$ and $\underline{L}^2$.

When acting on a multi-electron system these operators can be written in terms of the one-electron angular momentum and spin operators. Denote

$$L_i := \sum_{j=1}^{N} L_i(j), \quad S_i := \sum_{j=1}^{N} S_i(j)$$

where $L_i(j)$ and $S_i(j)$ are the ith components of $\underline{L}$ and $\underline{S}$ acting on the jth coordinate, and $\underline{L}(j)$ and $\underline{S}(j)$ are the vectors formed from these operators. This gives

$$\underline{L}^2 = \sum_{i,j=1}^{N} \underline{L}(i) \cdot \underline{L}(j), \quad \underline{S}^2 = \sum_{i,j=1}^{N} \underline{S}(i) \cdot \underline{S}(j),$$

and we now wish to derive identities for the action of these operators upon a Slater determinant.

3.1.1 One-particle Operators on Slater determinants

In this section we will derive rules for the application of L_3 and S_3 to a Slater determinant.

Lemma 3.1.1 ([Fri0X]). *Consider a general bounded linear operator B acting on $L^2(\mathbb{R}^3 \times \mathbb{Z}_2; \mathbb{C})$ and the N-particle analogue, $B_N := \sum_{i=1}^{N} B(i)$, where $B(i)$ is B acting on the ith coordinate, which acts on $L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N; \mathbb{C})$. Let $\psi_1 \dots \psi_N \in L^2(\mathbb{R}^3 \times \mathbb{Z}_2; \mathbb{C})$, then*

$$B_N|\psi_1 \dots \psi_N\rangle = |(B\psi_1)\psi_2 \dots \psi_N\rangle + \dots + |\psi_1 \dots (B\psi_N)\rangle.$$

Proof B_N is invariant under permutation of coordinates, hence it commutes with the antisymmetrizer $\mathcal{A}_N$ of the Slater determinant. Hence

$$B_N|\psi_1 \dots \psi_N\rangle = \sqrt{N!}\mathcal{A}_N\Big(\sum_{i=1}^{N} B(i)\Big)\psi_1 \otimes \dots \otimes \psi_N.$$

Now,

$$B(i)\psi_1 \otimes \cdots \otimes \psi_N = \psi_1 \otimes \cdots \otimes \psi_{i-1} \otimes (B\psi_i) \otimes \psi_{i+1} \otimes \cdots \otimes \psi_N$$

and the result follows. ☐

In particular, if $\psi_1, \ldots, \psi_N$ are all eigenfunctions of the operator B then the result simplifies even further:

Corollary 3.1.2. *If, further to the assumptions of Lemma 3.1.1, we have* $B\psi_i = \lambda_i \psi_i$ *then*

$$B_N|\psi_1 \ldots \psi_N\rangle = \Big(\sum_{i=1}^{N} \lambda_i\Big)|\psi_1 \ldots \psi_N\rangle.$$

This result is the main argument for choosing a basis of p-orbitals that are all eigenfunctions of L_3, for example $p_3, \frac{1}{\sqrt{2}}(p_1 + ip_2), \frac{1}{\sqrt{2}}(p_1 - ip_2)$. However, computing the L_3 eigenspaces for $\{p_1, p_2, p_3\}$ (for which the Coulomb and exchange integrals are the same) is easier than computing the necessary Coulomb and exchange integrals and $\underline{L}^2$ eigenspaces for $\{p_3, p_+, p_-\}$.

Lemma 3.1.1 and its corollary appear to be new, although the result of Corollary 3.1.2 is often used in the literature for S_3 and L_3 eigenfunctions, i.e. it is well known that the total eigenvalue is the sum of the one-particle eigenvalues.

3.1.2 Two-particle Operators on Slater Determinants

We now wish to derive an analogous formula for the action of a general two-electron operator of the form $B_N := \sum_{i,j=1}^{N} B(i) \cdot B(j)$, of which both $\underline{L}^2$ and $\underline{S}^2$ are specific examples. Unfortunately, due to the two-particle nature, there is no simple form as in the one-electron case. We may however show that the operator acts as the sum of actions of $B(i) \cdot B(j)$ for each electron pair (i, j).

Lemma 3.1.3. *Consider a general bounded linear operator B acting on $L^2(\mathbb{R}^3 \times \mathbb{Z}_2; \mathbb{C})$ and the extended action on pairs of electrons,*

$$B_N := \sum_{i,j=1}^{N} B(i) \cdot B(j),$$

where $B(k)$ is B acting on the kth coordinate, which acts on $L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N; \mathbb{C})$. Let the set $\{\psi_1, \ldots, \psi_K \mid \psi_i \in L^2(\mathbb{R}^3 \times \mathbb{Z}_2, \mathbb{C})\}$ be invariant under the action of B and let $c_{ij}^{ab} \in \mathbb{C}$ be the expansion coefficients of $(B(1) \cdot B(2))\psi_i(1)\psi_j(2)$ in pairs

$\psi_a(1)\psi_b(2)$ *for* $i \neq j$, *i.e.*

$$(B(1)\cdot B(2))\psi_i \otimes \psi_j =: \sum_{a,b=1}^{K} c_{ij}^{ab}\psi_a \otimes \psi_b.$$

Then for all $\alpha_1, \dots, \alpha_N \in \{1, \dots, K\}$ *we have*

$$B_N|\psi_{\alpha_1}\dots\psi_{\alpha_N}\rangle = \sum_{i=1}^{N} |\psi_{\alpha_1}\dots(B\cdot B\psi_{\alpha_i})\dots\psi_{\alpha_N}\rangle + \sum_{\substack{i,j=1\\ i\neq j}}^{N}\sum_{a,b=1}^{K} c_{\alpha_i\alpha_j}^{ab}|\psi_{\alpha_1}\dots\psi_a\dots\psi_b\dots\psi_{\alpha_N}\rangle,$$

where ψ_a *is at position* i *and* ψ_b *at position* j.

Proof As in the one particle case, B_N is invariant under permutation and so commutes with the antisymmetrizer $\mathcal{A}_N$. We now partition B_N into one-electron ($i = j$) and two-electron ($i \neq j$) terms and, by Lemma 3.1.1, we find that for a Slater determinant $|\psi_{\alpha_1}\dots\psi_{\alpha_N}\rangle$ the one-electron terms give $\sum_{i=1}^{N}|\psi_{\alpha_1}\dots(B\cdot B\psi_{\alpha_i})\dots\psi_{\alpha_N}\rangle$. The effect of the two-electron terms is less trivial and we find, setting

$$(B(1)\cdot B(2))\psi_i \otimes \psi_j =: \sum_{a,b=1}^{K} c_{ij}^{ab}\psi_a \otimes \psi_b,$$

and $\tilde{B_N} := \sum_{\substack{i,j=1\\ i\neq j}}^{N} B(i)\cdot B(j)$, that

$$\tilde{B_N}|\psi_{\alpha_1}\dots\psi_{\alpha_i}\dots\psi_{\alpha_N}\rangle = \sqrt{N!}\mathcal{A}_N\Bigg(\sum_{\substack{i,j=1\\ i\neq j}}^{N} B(i)\cdot B(j)\Bigg)\psi_{\alpha_1}\otimes\dots\otimes\psi_{\alpha_N}$$

$$= \sqrt{N!}\mathcal{A}_N\sum_{\substack{i,j=1\\ i\neq j}}^{N}\sum_{a,b=1}^{K} c_{\alpha_i\alpha_j}^{ab}\psi_{\alpha_1}\otimes\dots\otimes\psi_a\otimes\dots\otimes\psi_b\otimes\dots\otimes\psi_{\alpha_N}$$

where ψ_a and ψ_b are at positions i and j respectively. Taking the antisymmetrizer (which is linear) inside the sums we find

$$\tilde{B_N}|\psi_{\alpha_1}\dots\psi_{\alpha_i}\dots\psi_{\alpha_N}\rangle = \sum_{\substack{i,j=1\\ i\neq j}}^{N}\sum_{a,b=1}^{K} c_{\alpha_i\alpha_j}^{ab}|\psi_{\alpha_1}\dots\psi_a\dots\psi_b\dots\psi_{\alpha_N}\rangle$$

and the result follows. □

As with the one-particle case these results appear to be new. Unlike the one-particle case, when calculating $\underline{S}^2$ eigenfunctions, the consensus in the literature seem to be to use diagrammatic methods such as Young tableaux or branching diagrams [Pau79]. These methods, whilst relatively simple to apply, are generally not proved, presumably due to their dependence on more technical results from areas such as Lie algebra representation theory. Our approach, which requires only the application of the operator to each pair and then matrix diagonalisation to find the eigenfunctions seems preferable, and certainly more rigorous. $\underline{L}^2$ is generally not considered in the literature as it is not a well defined observable for molecules (due to lack of rotational symmetry).

3.1.3 Angular Momentum, Spin, Parity, and PHP

We now return to showing that the spin and angular momentum operators commute with the projected Hamiltonian PHP.

Along with these operators, we introduce the parity operator $\hat{R}$, which is given by inversion of all coordinates through the origin:

$$\hat{R}\Psi(x_1, s_1, x_2, s_2, \ldots, x_N, s_N) = \Psi(-x_1, s_1, -x_2, s_2, \ldots, -x_N, s_N). \tag{3.3}$$

This clearly commutes with spin operators and also, since inversion at the origin commutes with any rotation leaving the origin fixed, with angular momentum (which is generated by infinitesimal rotations).

Lemma 3.1.4. *Let $V :=$ GS of H_0 and let P be the orthogonal projection onto V. Then the operators L_3, S_3, $\underline{L}^2$, and $\underline{S}^2$ and $\hat{R}$ commute with PHP.*

Proof We prove the lemma in two parts, we first show that V and its orthogonal complement $V^\perp$ are invariant under each operator. We do this by showing that the span of one-electron orbitals U_n with principle quantum number n is invariant under the one-particle operators L_3, S_3 and $\hat{R}$ for each n.

We write a one-particle spin-up eigenfunction of $-\frac{1}{2}\Delta - Z/r$ as $\psi_{n\ell m} = R_{n,\ell}(r)Y_{\ell,m}(\theta,\phi)\alpha$, and analogously for the spin down with $\alpha \mapsto \beta$.

The invariance of U_n under L_3 and $\underline{L}^2$ follows from the fact that each L_i is independent of spin, gives zero on any function of r, and hence acts only on the $Y_{\ell,m}$. Since the spherical harmonics are eigenfunctions of L_3 and $\underline{L}^2$ (with eigenvalues m and $\ell + 1$) the result is trivial.

Analogously the S_i (and hence $\underline{S}^2$) act only on the spin and trivially this leaves n, ℓ and m invariant, since both spin states of each orbital are present in U_n the result follows.

Invariance under $\hat{R}$ follows from the fact that inversion does not affect the radial or spin parts and at most changes the sign of $Y_{\ell,m} = cP_\ell^m(\cos\theta)e^{im\phi}$: Note that $\theta \mapsto \theta + \pi$ and $\phi \mapsto \phi$ is an inversion, and this trivially leaves $e^{im\phi}$ invariant. Further, since $P_\ell^m(\cos\theta)$ is a linear combination of terms $\cos^a\theta \sin^b\theta$, with $a+b$ odd or even as ℓ is odd or even, using $\cos(\theta+\pi) = -\cos\theta$ and $\sin(\theta+\pi) = -\sin\theta$ gives the result.

Using Lemmas 3.1.1 and 3.1.3, we therefore see that V and $V^\perp$ are invariant under the N-electron operators L_3, S_3, $\underline{L}^2$, and $\underline{S}^2$. Invariance under the N-electron version of $\hat{R}$ follows since it acts simultaneously on all coordinates.

It follows that P commutes with each of these operators: Let A be a general linear operator such that $AV \subseteq V$ and $AV^\perp \subseteq V^\perp$. Defining $\Psi_V := P\Psi$, $\Psi_V^\perp := (I-P)\Psi$ shows

$$PA\Psi = PA(\Psi_V + \Psi_V^\perp) = P(A\Psi_V + A\Psi_V^\perp) = A\Psi_V = AP\Psi.$$

The second part of the proof is to show that the N-electron operators L_3, S_3, $\underline{L}^2$, $\underline{S}^2$ and $\hat{R}$ commute with H. The result for the spin operators follows trivially from the fact that H does not depend on spin. The result for the angular momentum operators follows from the fact that H is invariant under simultaneous rotation of all electron coordinates about any axis. Hence, since the angular momentum operators are generated by infinitesimal rotations, they must commute with H and the result follows. For the parity, we note that the H is invariant under inversion through the origin and so the result holds. □

Note that this result extends to any space V for which V and $V^\perp$ are invariant under each of the operators. In particular, as long as orthogonality is maintained, and the same radial part is used for each orbital with the same values of n and ℓ, the radial parts of the orbitals can be freely varied. Further, although this invariance is assumed for such subspaces in much of the standard literature (namely that spin and angular momentum eigenfunctions can be formed within such subspaces), there does not seem to be a rigorous proof readily available in the literature.

3.1.4 Spin of Slater Determinants

We first consider the effects of the spin operators on a Slater determinant. Using the Pauli matrices (3.2) we see that, denoting spin up and down by the corresponding arrows [Fri0X],

$$\underline{S}(\uparrow) = \begin{pmatrix} \frac{1}{2}\downarrow \\ \frac{i}{2}\downarrow \\ \frac{1}{2}\uparrow \end{pmatrix}, \quad \underline{S}(\downarrow) = \begin{pmatrix} \frac{1}{2}\uparrow \\ -\frac{i}{2}\uparrow \\ -\frac{1}{2}\downarrow \end{pmatrix}.$$

Hence we see that $S_3\psi(r,s) = \pm\frac{1}{2}\psi(r,s)$ for $s = \uparrow, \downarrow$ respectively. By Corollary 3.1.2 we have

$$S_3|\psi_1 \dots \psi_N\rangle = s|\psi_1 \dots \psi_N\rangle$$

where $s = \frac{1}{2}(\#\uparrow \text{orbitals} - \#\downarrow \text{orbitals})$.

The one-electron terms in $\underline{S}^2$ are given by $\underline{S} \cdot \underline{S} = \sum_{\ell=1}^{3} S_\ell S_\ell$. For $\psi(x,\uparrow)$ we have

$$\sum_{\ell=1}^{3} S_\ell S_\ell \psi(x,\uparrow) = S_1\tfrac{1}{2}\psi(x,\downarrow) + S_2\tfrac{i}{2}\psi(x,\downarrow) + S_3\tfrac{1}{2}\psi(x,\uparrow) = \tfrac{3}{4}\psi(x,\uparrow),$$

and the spin down case is analogous, giving $\frac{3}{4}\psi(r,\downarrow)$. Hence by Lemma 3.1.3 we see that the one electron term simply gives

$$\sum_{i=1}^{N} |\psi_{\alpha_1} \dots (\underline{S} \cdot \underline{S}\psi_{\alpha_i}) \dots \psi_{\alpha_N}\rangle = \tfrac{3N}{4}|\psi_{\alpha_1} \dots \psi_{\alpha_N}\rangle \tag{3.4}$$

For the two-electron term we consider the action on a general tensor product of two one-electron orbitals. Consider two spatial wavefunctions $a, b \in L^2(\mathbb{R}^3)$, there are four cases depending on the spin states of the orbitals [Fri0X]:

$$\begin{aligned}
(\underline{S}(1) \cdot \underline{S}(2))a\uparrow \otimes b\downarrow &= \tfrac{1}{4}a\downarrow \otimes b\uparrow + \tfrac{1}{4}a\downarrow \otimes b\uparrow - \tfrac{1}{4}a\uparrow \otimes b\downarrow \\
&= \tfrac{1}{2}a\downarrow \otimes b\uparrow - \tfrac{1}{4}a\uparrow \otimes b\downarrow \\
(\underline{S}(1) \cdot \underline{S}(2))a\downarrow \otimes b\uparrow &= \tfrac{1}{2}a\uparrow \otimes b\downarrow - \tfrac{1}{4}a\downarrow \otimes b\uparrow \\
(\underline{S}(1) \cdot \underline{S}(2))a\uparrow \otimes b\uparrow &= \tfrac{1}{4}a\uparrow \otimes b\uparrow \\
(\underline{S}(1) \cdot \underline{S}(2))a\downarrow \otimes b\downarrow &= \tfrac{1}{4}a\downarrow \otimes b\downarrow\,.
\end{aligned} \tag{3.5}$$

Note that, in particular, a spin pair of the same spatial orbitals in a Slater determinant does not contribute to the total spin.

3.1.5 Angular Momentum of Slater Determinants

As with spin we start by calculating the effect of $\underline{L}$ on our basis functions. From the derivative form of $\underline{L}$ in (3.1) we see that $\underline{L}\psi(|x|, s) = 0$, which follows from

$$\partial_i f(|x|) = f'(|x|)\partial_i(|x|) = f'(|x|)\tfrac{x_i}{|x|},$$

and hence $(x_j\partial_i - x_i\partial_j)f(|x|) = 0$. For three general p-orbitals defined as $p_i(x, s) = f(|x|)x_i\chi(s)$ we find [Fri0X] that the spatial parts satisfy

$$\underline{L}p_1 = \begin{pmatrix} 0 \\ -ip_3 \\ ip_2 \end{pmatrix}, \quad \underline{L}p_2 = \begin{pmatrix} ip_3 \\ 0 \\ -ip_1 \end{pmatrix}, \quad \underline{L}p_3 = \begin{pmatrix} -ip_2 \\ ip_1 \\ 0 \end{pmatrix}. \tag{3.6}$$

Which, if the indices are understood modulo three, is equivalent to

$$L_j p_j = 0, \quad L_{j+1}p_j = -ip_{j-1}, \quad L_{j-1}p_j = ip_{j+1},$$

and in particular

$$L_3 p_1 = ip_2, \quad L_3 p_2 = -ip_1, \quad L_3 p_3 = 0.$$

For the one-electron terms of $\underline{L}^2$ we find that

$$\sum_{\ell=1}^{3} L_\ell L_\ell p_j(x)\chi(s) = iL_{j-1}p_{j+1}\chi(s) - iL_{j+1}p_{j-1}\chi(s) = 2p_j\chi(s).$$

Hence by Lemma 3.1.3 we see that the one electron term simply gives

$$\sum_{i=1}^{N} |\psi_{\alpha_1} \dots (\underline{L} \cdot \underline{L}\psi_{\alpha_i}) \dots \psi_{\alpha_k}\rangle = 2P|\psi_{\alpha_1} \dots \psi_{\alpha_N}\rangle, \tag{3.7}$$

where P is the number of p-orbitals in the Slater determinant.

Finally we need to compute the two-electron terms and, since $\underline{L}$ is independent of spin, we have [Fri0X]

$$\begin{aligned}
(\underline{L}(1) \cdot \underline{L}(2))p_i\alpha \otimes p_i\beta &= -(p_{i-1}\alpha \otimes p_{i-1}\beta + p_{i+1}\alpha \otimes p_{i+1}\beta) \\
(\underline{L}(1) \cdot \underline{L}(2))p_i\alpha \otimes p_j\beta &= p_j\alpha \otimes p_i\beta \quad (i \neq j),
\end{aligned} \tag{3.8}$$

for $\alpha, \beta \in \{\uparrow, \downarrow\}$.

We are now in a position to use Lemmas 3.1.1 and 3.1.3 to calculate the $\underline{L}^2$-$\underline{S}^2$-

L_3-S_3-$\hat{R}$ joint eigenspaces of Lithium to Neon. However, before we do so, there is one more result which will reduce the amount of computation needed. The motivation for the following comes mainly from the sizes of the ground states of H_0, which can be seen in Table 2.1. It is clear that the dimension of a space with p valence electrons is the same as that with $8-p$ valence electrons, which can be seen as a property of binomial coefficients. We propose that there is also a deeper symmetry, which is analogous to a particle-hole duality result.

3.1.6 Particle-Hole Duality

We begin by recalling the definition of the Hodge Star operator from differential geometry [Jos02]. Let M be a n-dimensional Riemannian manifold with inner product g. The Hodge star operator (denoted by $*$) is a linear operator mapping k-forms on M to $(n-k)$-forms on M, i.e.

$$* : \Omega^k(M^n) \to \Omega^{n-k}(M^n).$$

This operator may be defined in a coordinate-free manner by the condition

$$u \wedge *v = g(u,v)\,\mathbf{Vol}(g), \tag{3.9}$$

where the notation $g(u,v)$ denotes the inner product on k-forms and $\mathbf{Vol}(g)$ is the unit volume form associated to the metric.

In our case the n-dimensional manifold is analogous to the space of Slater determinants with n valence orbitals, and thus the k- and $(n-k)$-forms are respectively the spaces of Slater determinants containing k and $(n-k)$ valence orbitals. We wish to define an operator analogous to $*$ which will then be used to prove a duality result.

We begin by defining the volume element for the valence shell by

$$|1_v\rangle := |2s\overline{2s}p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle,$$

which is unique up to sign. If we were to use the definition of $*$ given by (3.9), the left hand side would not necessarily be a Slater determinant as it may not be antisymmetric. Hence we define the operator

$$\tilde{\mathcal{A}}_n : \tilde{S}_{n-k} \otimes \tilde{S}_k \to \tilde{S}_n,$$

where $\tilde{S}_m$ is the space of *normalized* Slater determinants containing m orbitals, to be the standard antisymmetrizer which also normalizes the resulting Slater determinant. We now extend $\tilde{\mathcal{A}}_n$ to the space of non-normalized Slater determinants S_n by

$$\mathcal{A}_n\big(\alpha|\Psi_1\rangle \otimes \beta|\Psi_2\rangle\big) := \alpha\beta\tilde{\mathcal{A}}_n\big(|\Psi_1\rangle \otimes |\Psi_2\rangle\big),$$

with $|\Psi_1\rangle \in \tilde{S}_{n-k}$ and $|\Psi_2\rangle \in \tilde{S}_k$.

We finally define the dual (star) operator $* : S_k \to S_{n-k}$ by

$$\mathcal{A}_n\big(\alpha|\Psi_1\rangle \otimes *(\beta|\Psi_2\rangle)\big) := \langle\alpha\Psi_1|\beta\Psi_2\rangle|1_v\rangle, \tag{3.10}$$

where $\alpha|\Psi_1\rangle \in S_{n-k}$, $\beta|\Psi_2\rangle \in S_k$, and $\langle\cdot|\cdot\rangle$ denotes the standard $L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^{n-k};\mathbb{C})$ inner product on S_{n-k}.

It is clear from this form, and the fact that all Slater determinants which differ by at least one one-electron orbital are orthogonal, that $*$ gives a non-zero result only when $|\Psi_1\rangle = |\Psi_2\rangle$ and then (3.10) returns a scalar multiple of $|1_v\rangle$. This suggests that $*|\Psi\rangle$ should be a Slater determinant formed from the orbitals of $|1_v\rangle$ that are not present in $|\Psi\rangle$. It is clear that this result can be extended to Slater determinants that all contain the same core orbitals by suitably modifying the tensor product operator to retain only one set of core orbitals, and extending the inner product to include the core orbitals.

An equivalent definition for our Slater determinants, in terms of creation and annihilation operators is given by:

Definition 3.1.5. *The dual of a Slater determinant* $\alpha|\Psi\rangle = \alpha|1s\overline{1s}\phi_1 \dots \phi_k\rangle$, $\alpha \in \mathbb{C}$, *denoted by* $*(\alpha|\Psi\rangle)$, *is given by*

$$*\big(\alpha|\Psi\rangle\big) := \alpha^* a(\phi_k) \dots a(\phi_1)|1\rangle, \tag{3.11}$$

where $|1\rangle := |1s\overline{1s}2s\overline{2s}p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$. *Note that* $*$ *extends linearly to linear combinations of Slater determinants.*

We note that the complex conjugation of α is necessary to ensure consistency with the conjugate linear property of the inner product. It is easy to check that, after extending the inner product to include the core orbitals, and modifying the tensor product to retain only one set of the core orbitals, that this definition is consistent with (3.10).

This formulation of the dual seems to be novel and, although there are cases where the idea is mentioned in passing (e.g. in the discussion of [LS60]), there seems to be no formal investigation in the literature. Hence the remaining results in this section are new.

Lemma 3.1.6 (Anti-commutativity of $\underline{L}$ and $*$)**.** *Let $\underline{L}$ denote the angular momentum operator, which, in second quantization, is given by*

$$\underline{L} := i \begin{bmatrix} a_3^\dagger a_2 - a_2^\dagger a_3 + a_{\overline{3}}^\dagger a_{\overline{2}} - a_{\overline{2}}^\dagger a_{\overline{3}} \\ a_1^\dagger a_3 - a_3^\dagger a_1 + a_{\overline{1}}^\dagger a_{\overline{3}} - a_{\overline{3}}^\dagger a_{\overline{1}} \\ a_2^\dagger a_1 - a_1^\dagger a_2 + a_{\overline{2}}^\dagger a_{\overline{1}} - a_{\overline{1}}^\dagger a_{\overline{2}} \end{bmatrix}$$

where $a_i := a(p_i)$, $a_{\overline{i}} := a(\overline{p_i})$, etc. Then $\underline{L}$ and $*$ anticommute on

$$\mathit{Span}\big\{ |1s\overline{1s}\phi_1\phi_2\ldots\phi_k\rangle \mid \phi_j \in \{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\} \big\}.$$

Proof We begin by considering the term $i(a_m^\dagger a_n - a_n^\dagger a_m), m \neq n$ and show that this anticommutes with $*$ for a general Slater determinant $\alpha a^\dagger(\phi_1)\ldots a^\dagger(\phi_k)|0\rangle$ (where $|0\rangle = |1s\overline{1s}\rangle$). For the first term, if none of the ϕ_i are p_n, we see that $i\alpha a_m^\dagger a_n a^\dagger(\phi_1)\ldots a^\dagger(\phi_k)|0\rangle = 0$ as, in this case, a_n anticommutes up to the vacuum. It is also zero if one of the ϕ_i is p_m as $a_m^\dagger a_m^\dagger = 0$. Else we must have $\phi_i = p_n$ for some i so anticommuting both $a_m^\dagger$ and a_n up to this point (which results in no change of sign) gives

$$i\alpha a^\dagger(\phi_1)\ldots a^\dagger(\phi_{i-1}) a_m^\dagger a_n a_n^\dagger a^\dagger(\phi_{i+1})\ldots a^\dagger(\phi_k)|0\rangle.$$

Rewriting $a_n a_n^\dagger = 1 - a_n^\dagger a_n$ and anticommuting a_n up to the vacuum leaves only the term from the 1, and hence we have

$$\begin{aligned} &i\alpha a_m^\dagger a_n a^\dagger(\phi_1)\ldots a^\dagger(\phi_k)|0\rangle \\ &\quad= \begin{cases} 0 & \text{if } \phi_i \neq p_n \forall i \text{ or } \phi_j = p_m, \\ i\alpha a^\dagger(\phi_1)\ldots a^\dagger(\phi_{i-1}) a_m^\dagger a^\dagger(\phi_{i+1})\ldots a^\dagger(\phi_k)|0\rangle & \text{if } \phi_j \neq p_m \forall j \text{ and } \phi_i = p_n. \end{cases} \end{aligned}$$

Using (3.11), this gives

$$\begin{aligned} &* \, i\alpha a_m^\dagger a_n a^\dagger(\phi_1)\ldots a^\dagger(\phi_k)|0\rangle \\ &\quad= \begin{cases} 0 & \text{if } \phi_i \neq p_n \forall i \text{ or } \phi_j = p_m, \\ (i\alpha)^* a(\phi_k)\ldots a(\phi_{i+1}) a_m a(\phi_{i-1})\ldots a(\phi_1)|1\rangle & \text{if } \phi_j \neq p_m \forall j \text{ and } \phi_i = p_n. \end{cases} \end{aligned}$$

We therefore see that

$$
\begin{aligned}
&* \left(i(a_m^\dagger a_n - a_n^\dagger a_m)\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle\right) \qquad (3.12)\\
&= \begin{cases} (i\alpha)^* a(\phi_k)\dots a(\phi_{i+1})a_m a(\phi_{i-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_j \neq p_m \forall j \text{ and } \phi_i = p_n,\\ -(i\alpha)^* a(\phi_k)\dots a(\phi_{j+1})a_n a(\phi_{j-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_i \neq p_n \forall i \text{ and } \phi_j = p_m,\\ 0 & \text{else.} \end{cases}
\end{aligned}
$$

We now wish to compute $ia_m^\dagger a_n(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle)$. Applying $*$ gives $i\alpha^* a_m^\dagger a_n a(\phi_k)\dots a(\phi_1)|1\rangle$ and, as before, we note that this is zero if one of the ϕ_i is p_n or if none of the ϕ_i are p_m. (Note that it is here that we use that $|1\rangle$ is a closed shell, i.e. contains all orbitals involved in $\underline{L}$.) Else, following a similar method to before, we find that

$$
\begin{aligned}
&i\alpha^* a_m^\dagger a_n a(\phi_k)\dots a(\phi_{j+1})a_m a(\phi_{j-1})\dots a(\phi_1)|1\rangle\\
&\qquad = -i\alpha^* a(\phi_k)\dots a(\phi_{j+1})a_n a(\phi_{j-1})\dots a(\phi_1)|1\rangle.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&ia_m^\dagger a_n * \alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle\\
&= \begin{cases} 0 & \text{if } \phi_j \neq p_m \forall j \text{ or } \phi_i = p_n,\\ -i\alpha^* a(\phi_k)\dots a(\phi_{j+1})a_n a(\phi_{j-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_i \neq p_n \forall i \text{ and } \phi_j = p_m. \end{cases}
\end{aligned}
$$

We therefore see that

$$
\begin{aligned}
&i(a_m^\dagger a_n - a_n^\dagger a_m)(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle) \qquad (3.13)\\
&= \begin{cases} -(i\alpha)^* a(\phi_k)\dots a(\phi_{i+1})a_m a(\phi_{i-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_j \neq p_m \forall j \text{ and } \phi_i = p_n,\\ (i\alpha)^* a(\phi_k)\dots a(\phi_{j+1})a_n a(\phi_{j-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_i \neq p_n \forall i \text{ and } \phi_j = p_m,\\ 0 & \text{else,} \end{cases}
\end{aligned}
$$

where we have used the identity $(i\alpha)^* = -(i\alpha)^*$ for $\alpha \in \mathbb{C}$.

It follows from (3.12) and (3.13) that, on a general Slater determinant, $i(a_m^\dagger a_n - a_n^\dagger a_m)* = - * i(a_m^\dagger a_n - a_n^\dagger a_m)$, which implies $\underline{L}* = - * \underline{L}$. The result follows by the linearity of $\underline{L}$ and $*$. □

The identities in the proof will be useful when investigating the relationship between $*\underline{S}$ and $\underline{S}*$ so, for ease of reference, we have:

Corollary 3.1.7. *For* $\chi_n \neq \chi_m$

$$\begin{aligned}
&* \beta a^\dagger(\chi_m)a(\chi_n)\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle \\
&= \begin{cases} 0 & \text{if } \phi_i \neq \chi_n \forall i \text{ or } \phi_j = \chi_m, \\ (\alpha\beta)^* a(\phi_k)\dots a(\phi_{i+1})a(\chi_m)a(\phi_{i-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_j \neq \chi_m \forall j \text{ and } \phi_i = \chi_n. \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
&\beta a^\dagger(\chi_m)a(\chi_n)(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle) \\
&= \begin{cases} 0 & \text{if } \phi_j \neq \chi_m \forall j \text{ or } \phi_i = \chi_n, \\ -\beta\alpha^* a(\phi_k)\dots a(\phi_{j+1})a(\chi_n)a(\phi_{j-1})\dots a(\phi_1)|1\rangle & \text{if } \phi_i \neq \chi_n \forall i \text{ and } \phi_j = \chi_m. \end{cases}
\end{aligned}$$

Lemma 3.1.8 (Anti-commutativity of $\underline{S}$ and $*$). *Let* $\underline{S}$ *denote the spin angular momentum operator, which, in second quantization, is given by*

$$\underline{S} := \frac{1}{2}\sum_j \begin{bmatrix} a^\dagger(\overline{\chi_j})a(\chi_j) + a^\dagger(\chi_j)a(\overline{\chi_j}) \\ i\big(a^\dagger(\overline{\chi_j})a(\chi_j) - a^\dagger(\chi_j)a(\overline{\chi_j})\big) \\ a^\dagger(\chi_j)a(\chi_j) - a^\dagger(\overline{\chi_j})a(\overline{\chi_j}) \end{bmatrix}$$

where χ_j *are the spatial orbitals which, when coupled with the spin functions corresponding to up and down, form the spin orbitals* ϕ_j *and* $\overline{\phi_j}$*. Then* $\underline{S}$ *and* $*$ *anticommute on*

$$Span\big\{|1s\overline{1s}\phi_1\phi_2\dots\phi_k\rangle \mid \phi_j \in \{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\}\big\}.$$

Proof First note that each component of $\underline{S}$ gives zero when applied to a spin pair so we need only consider the action on the valence orbitals, i.e. $a(\phi_j)|0\rangle = 0$ for all j.

From the second quantized form of $\underline{S}$ it is clear that, along with the results of Corollary 3.1.7, we also need $*\beta a^\dagger(\chi)a(\chi)\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle$ and $\beta a^\dagger(\chi)a(\chi)(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle)$. We begin with

$$\beta a^\dagger(\chi)a(\chi)\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle$$

and note that this is clearly zero if none of the ϕ_i are χ (by acting $a(\chi)$ on $|0\rangle$) and else is $\beta\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle$ (by noting that $a^\dagger(\chi)a(\chi)a^\dagger(\chi) = (1 -$

$a(\chi)a^\dagger(\chi))a^\dagger(\chi) = a^\dagger(\chi))$. Hence we see that

$$*\beta a^\dagger(\chi)a(\chi)\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle = \begin{cases} 0 & \text{if } \phi_i \neq \chi \forall i, \\ (\alpha\beta)^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_i = \chi. \end{cases}$$

It remains to compute $\beta a^\dagger(\chi)a(\chi)(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle)$ which, after applying $*$ is $\beta\alpha^* a^\dagger(\chi)a(\chi)a(\phi_k)\dots a(\phi_1)|1\rangle$, which is clearly zero if $\phi_i = \chi$ and else is $\beta\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle$ by anticommuting both $a^\dagger(\chi)$ and $a(\chi)$ up to $|1\rangle$ which gives no sign change and noting that $a^\dagger(\chi)a(\chi)|1\rangle = |1\rangle$. We therefore have

$$\beta a^\dagger(\chi)a(\chi)(*\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle) = \begin{cases} 0 & \text{if } \phi_i = \chi, \\ \beta\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_i \neq \chi \forall i. \end{cases}$$

Considering $S_3 = \frac{1}{2}\sum_j a^\dagger(\chi_j)a(\chi_j) - a^\dagger(\overline{\chi_j})a(\overline{\chi_j})$ and in particular one compoenent $\tilde{S}_3 := \frac{1}{2}a^\dagger(\chi)a(\chi) - a^\dagger(\overline{\chi})a(\overline{\chi})$, we see that

$$*\tilde{S}_3\alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle = \begin{cases} \frac{1}{2}\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_i = \chi, \phi_j \neq \overline{\chi}\forall j, \\ -\frac{1}{2}\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_j = \overline{\chi}, \phi_i \neq \chi \forall i, \\ 0 & \text{else.} \end{cases}$$

and

$$\tilde{S}_3 * \alpha a^\dagger(\phi_1)\dots a^\dagger(\phi_k)|0\rangle = \begin{cases} -\frac{1}{2}\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_i = \chi, \phi_j \neq \overline{\chi}\forall j, \\ \frac{1}{2}\alpha^* a(\phi_k)\dots a(\phi_1)|1\rangle & \text{if } \phi_j = \overline{\chi}, \phi_i \neq \chi \forall i, \\ 0 & \text{else.} \end{cases}$$

from which it is clear that $S_3* = -*S_3$.

The proofs that $S_1* = -*S_1$ and $S_2* = -*S_2$ are analogous and use the formulae from Corollary 3.1.7, (and are in fact virtually identical to the arguments of Lemma 3.1.6). Hence, on Slater determinants, $\underline{S}* = -*\underline{S}$ and the result follows by the linearity of $\underline{S}$ and $*$. ☐

Lemma 3.1.9 (Commutativity of $\hat{R}$ and $*$). *Let* $\Psi = \sum_{i=1}^N \alpha_i|\Psi_i\rangle$ *with*

$$|\Psi_i\rangle \in \big\{|1s\overline{1s}\phi_1\phi_2\dots\phi_k\rangle \mid \phi_j \in \{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\}\big\}.$$

be such that $\hat{R}\Psi = r\Psi$, $r \in \{-1, +1\}$, *(i.e.* Ψ *has well defined parity) then* $\hat{R}(*\Psi) = r(*\Psi)$.

Proof For a single Slater determinant, the result is trivial since $|1\rangle$ contains an even number of p-orbitals and the parity is simply equivalent to the parity of the number of p-orbitals in a Slater determinant. The result follows by the linearity of $*$ and $\hat{R}$. $\square$

It is clear that, for the obvious extension of $*$ to higher valence shells, Lemma 3.1.8 extends to any set of Slater determinants that all contain the same core of spin-paired orbitals and the remaining orbitals chosen from the valence shell. The extensions for Lemmas 3.1.6 and 3.1.9 would require determining the action of $\underline{L}$ and $\hat{R}$ on higher angular momentum one-electron orbitals.

Theorem 3.1.10 (Particle-Hole Duality). *Let* $\Psi = \sum_{i=1}^{I} \alpha_i |\Psi_i\rangle$ *be a linear combination of Slater determinants, where* $|\Psi_i\rangle = |1s\overline{1s}\phi_1\phi_2 \dots \phi_k\rangle$ *with*

$$\phi_j \in \{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\},$$

which satisfies $\underline{L}^2\Psi = \mathcal{L}\Psi$ *and* $\underline{S}^2\Psi = \mathcal{S}\Psi$*. Then for the dual operator given by Definition 3.1.5,* $\underline{L}^2(*\Psi) = \mathcal{L}(*\Psi)$ *and* $\underline{S}^2(*\Psi) = \mathcal{S}(*\Psi)$*. Furthermore, if* $L_3\Psi = \ell\Psi$*,* $S_3\Psi = s\Psi$ *and* $\hat{R}\Psi = r\Psi$*, then* $L_3(*\Psi) = -\ell(*\Psi)$*,* $S_3(*\Psi) = -s(*\Psi)$ *and* $\hat{R}(*\Psi) = r(*\Psi)$*.*

Proof The proof follows directly from Lemmas 3.1.6, 3.1.8 and 3.1.9. $\square$

This result is very useful as it follows that we need only compute the eigenspaces for Li-C (1 to 4 valence electrons) and the others follow immediately. We also note that the above results are not restricted to the $n = 2$ valence shell and, assuming analogous results exist for higher angular momentum orbitals, would be even more advantageous for the larger n shells.

3.2 Explicit Many Electron Eigenspaces

3.2.1 Calculation of $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ Eigenspaces

Recall from Lemma 3.1.4 that $\underline{L}^2$, $\underline{S}^2$, L_3, S_3, $\hat{R}$ and PHP commute, hence we may find simultaneous eigenspaces of these operators.

In order to simplify the analysis we note that each Slater determinant is already an eigenfunction of S_3, and that the space spanned by $|\Psi_i\rangle$ with S_3 eigenvalue s is isomorphic (by flipping all spins) to that spanned by $|\tilde{\Psi}_i\rangle$ with S_3 eigenvalue $-s$. This can be seen from the fact that $\underline{L}^2$ and $\underline{S}^2$ commute with flipping the spins of all electrons; the result for $\underline{L}^2$ follows trivially from (3.7) and (3.8) and that for $\underline{S}^2$ from (3.4) and (3.5). It follows that it is sufficient to consider cases where $S_3\Psi = s\Psi$ with $s \geq 0$.

Further we see that the contribution of the pair $1s\overline{1s}$ to S_3, $\underline{S}^2$, L_3 and $\underline{L}^2$ is zero, due to it being a closed spin shell of radial functions. Further, the eigenvalue of $\hat{R}$ on a Slater determinant is the product of the eigenvalues of the one-electron orbitals, which are independent of spin. Hence we may ignore the core $1s$-orbitals for the purpose of these calculations.

We begin by choosing a basis of single Slater determinants for each S_3 eigenspace and forming the corresponding $\underline{S}^2$ matrix. The eigenfunctions and eigenvalues are then found by explicitly diagonalization, giving a basis of each S_3-$\underline{S}^2$ eigenspace with eigenvalues s and $\mathcal{S}$ respectively.

For each of these S_3-$\underline{S}^2$ bases we form the $\underline{L}^2$ matrix and again diagonalize to give a basis for each S_3-$\underline{S}^2$-$\underline{L}^2$ eigenspace with eigenvalues s, $\mathcal{S}$ and $\mathcal{L}$ respectively.

Finally we note that the energy of an $\underline{L}^2$-$\underline{S}^2$ eigenfunction is independent of the choice of S_3 and L_3 eigenfunction, and hence we need only choose a specific $\underline{L}^2$-$\underline{S}^2$-S_3 eigenspace (normally that with highest s eigenvalue as this gives the smallest number of S_3 basis elements), and construct the L_3 eigenfunctions in the same manner. The L_3 eigenvalue is always chosen to be zero as this allows the results of Theorem 3.3.1 to be more easily demonstrated.

The $\underline{L}^2$-$\underline{S}^2$-S_3-L_3-$\hat{R}$ eigenspaces are then trivial to identify by inspection.

The calculations themselves are somewhat lengthy and are given in Appendix C. It is worth recalling that Theorem 3.1.10 means we need only calculate the eigenspaces up to Carbon, the rest being given by duality.

As mentioned previously, although diagrammatic methods to compute $\underline{S}^2$ eigenfunctions are well known, there are no such results for $\underline{L}^2$ eigenfunctions, and these tables seem to be the first explicit formulation of the eigenfunctions.

3.2.2 $\underline{L}^2$-$\underline{S}^2$ Simultaneous Eigenspaces

The $\underline{L}^2$-$\underline{S}^2$ eigenspaces for Lithium-Neon are given in Tables 3.1-3.7, which use the following conventions: 1) The two $1s$ orbitals present in every Slater determinant are not shown. 2) The eigenfunctions are not normalized. 3) In all cases, it is assumed that $i, j, k \in \{1, 2, 3\}$, $i \neq j \neq k$. Further, the eigenspace contains all possible linear combinations of Slater determinants for which this rule holds, e.g. $|p_i p_j\rangle \rightsquigarrow \{|p_1p_2\rangle, |p_1p_3\rangle, |p_2p_3\rangle\}$. 4) Eigenfunctions of the form $a|\Psi_1\rangle+b|\Psi_2\rangle+c|\Psi_3\rangle$ are such that $a + b + c = 0$ and are hence two-dimensional eigenspaces.

Before looking at the refined $\underline{L}^2$-$\underline{S}^2$-L_3-S_3 eigenspaces, it is useful to note that

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\rangle$ $\lvert \overline{s}\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert p_i\rangle$ $\lvert \overline{p_i}\rangle$

Table 3.1: Lithium $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\overline{s}\rangle$ $\lvert p_1\overline{p_1}\rangle + \lvert p_2\overline{p_2}\rangle + \lvert p_3\overline{p_3}\rangle$	
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{p_i}\rangle - \lvert \overline{s}p_i\rangle$	$\lvert sp_i\rangle$ $\lvert p_ip_j\rangle$ $\lvert \overline{sp_i}\rangle$ $\lvert \overline{p_ip_j}\rangle$ $\lvert s\overline{p_i}\rangle + \lvert \overline{s}p_i\rangle$ $\lvert p_i\overline{p_j}\rangle + \lvert \overline{p_i}p_j\rangle$
$\langle \underline{L}^2 \rangle = 6$	$\lvert p_i\overline{p_j}\rangle - \lvert \overline{p_i}p_j\rangle$ $a\lvert p_1\overline{p_1}\rangle + b\lvert p_2\overline{p_2}\rangle + c\lvert p_3\overline{p_3}\rangle$	

Table 3.2: Beryllium $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$	$\langle \underline{S}^2 \rangle = \frac{15}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert sp_1\overline{p_1}\rangle + \lvert sp_2\overline{p_2}\rangle + \lvert sp_3\overline{p_3}\rangle$ $\lvert \overline{s}p_1\overline{p_1}\rangle + \lvert \overline{s}p_2\overline{p_2}\rangle + \lvert \overline{s}p_3\overline{p_3}\rangle$	$\lvert p_1p_2p_3\rangle$ $\lvert \overline{p_1p_2p_3}\rangle$ $\lvert p_1p_2\overline{p_3}\rangle + \lvert p_1\overline{p_2}p_3\rangle + \lvert \overline{p_1}p_2p_3\rangle$ $\lvert \overline{p_1p_2}p_3\rangle + \lvert \overline{p_1}p_2\overline{p_3}\rangle + \lvert p_1\overline{p_2p_3}\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{s}p_i\rangle$ $\lvert s\overline{sp_i}\rangle$ $\lvert p_ip_j\overline{p_j}\rangle + \lvert p_ip_k\overline{p_k}\rangle$ $\lvert \overline{p_i}p_j\overline{p_j}\rangle + \lvert \overline{p_i}p_k\overline{p_k}\rangle$ $2\lvert \overline{s}p_ip_j\rangle - \lvert s\overline{p_i}p_j\rangle - \lvert sp_i\overline{p_j}\rangle$ $2\lvert s\overline{p_ip_j}\rangle - \lvert \overline{s}p_i\overline{p_j}\rangle - \lvert \overline{sp_i}p_j\rangle$	$\lvert sp_ip_j\rangle$ $\lvert \overline{sp_ip_j}\rangle$ $\lvert sp_i\overline{p_j}\rangle + \lvert s\overline{p_i}p_j\rangle + \lvert \overline{s}p_ip_j\rangle$ $\lvert s\overline{p_ip_j}\rangle + \lvert \overline{s}p_i\overline{p_j}\rangle + \lvert \overline{sp_i}p_j\rangle$
$\langle \underline{L}^2 \rangle = 6$	$\lvert sp_i\overline{p_j}\rangle - \lvert s\overline{p_i}p_j\rangle$ $\lvert \overline{s}p_i\overline{p_j}\rangle - \lvert \overline{sp_i}p_j\rangle$ $a\lvert sp_1\overline{p_1}\rangle + b\lvert sp_2\overline{p_2}\rangle + c\lvert sp_3\overline{p_3}\rangle$ $a\lvert \overline{s}p_1\overline{p_1}\rangle + b\lvert \overline{s}p_2\overline{p_2}\rangle + c\lvert \overline{s}p_3\overline{p_3}\rangle$ $\lvert p_ip_j\overline{p_j}\rangle - \lvert p_ip_k\overline{p_k}\rangle$ $\lvert \overline{p_i}p_j\overline{p_j}\rangle - \lvert \overline{p_i}p_k\overline{p_k}\rangle$ $a\lvert \overline{p_3}p_1p_2\rangle + b\lvert p_3p_1\overline{p_2}\rangle + c\lvert p_3\overline{p_1}p_2\rangle$ $a\lvert p_3\overline{p_1p_2}\rangle + b\lvert \overline{p_3}p_1\overline{p_2}\rangle + c\lvert \overline{p_3p_1}p_2\rangle$	

Table 3.3: Boron $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$	$\langle \underline{S}^2 \rangle = 6$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\overline{s}p_1\overline{p_1}\rangle + \lvert s\overline{s}p_2\overline{p_2}\rangle + \lvert s\overline{s}p_3\overline{p_3}\rangle$ $\lvert p_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert p_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert p_2\overline{p_2}p_3\overline{p_3}\rangle$	$3\lvert \overline{s}p_1p_2p_3\rangle - \lvert sp_1p_2\overline{p_3}\rangle - \lvert s\overline{p_1}p_2p_3\rangle - \lvert sp_1\overline{p_2}p_3\rangle$ $3\lvert s\overline{p_1p_2p_3}\rangle - \lvert \overline{sp_1p_2}p_3\rangle - \lvert \overline{s}p_1\overline{p_2p_3}\rangle - \lvert \overline{sp_1}p_2\overline{p_3}\rangle$ $\lvert sp_1\overline{p_2p_3}\rangle + \lvert s\overline{p_1}p_2\overline{p_3}\rangle + \lvert s\overline{p_1p_2}p_3\rangle$ $-\lvert \overline{sp_1}p_2p_3\rangle - \lvert \overline{s}p_1\overline{p_2}p_3\rangle - \lvert \overline{s}p_1p_2\overline{p_3}\rangle$	$\lvert sp_1p_2p_3\rangle$ $\lvert \overline{sp_1p_2p_3}\rangle$ $\lvert \overline{s}p_1p_2p_3\rangle + \lvert sp_1p_2\overline{p_3}\rangle + \lvert sp_1\overline{p_2}p_3\rangle + \lvert s\overline{p_1}p_2p_3\rangle$ $\lvert s\overline{p_1p_2p_3}\rangle + \lvert \overline{sp_1p_2}p_3\rangle + \lvert \overline{sp_1}p_2\overline{p_3}\rangle + \lvert \overline{s}p_1\overline{p_2p_3}\rangle$ $\lvert sp_1\overline{p_2p_3}\rangle + \lvert s\overline{p_1}p_2\overline{p_3}\rangle + \lvert s\overline{p_1p_2}p_3\rangle$ $+\lvert \overline{sp_1}p_2p_3\rangle + \lvert \overline{s}p_1\overline{p_2}p_3\rangle + \lvert \overline{s}p_1p_2\overline{p_3}\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{p_i}p_j\overline{p_j}\rangle - \lvert \overline{s}p_ip_j\overline{p_j}\rangle + \lvert s\overline{p_i}p_k\overline{p_k}\rangle - \lvert \overline{s}p_ip_k\overline{p_k}\rangle$	$\lvert s\overline{s}p_ip_j\rangle$ $\lvert s\overline{sp_ip_j}\rangle$ $\lvert s\overline{s}p_i\overline{p_j}\rangle + \lvert s\overline{sp_i}p_j\rangle$ $\lvert p_1p_2p_3\overline{p_i}\rangle$ $\lvert \overline{p_1p_2p_3}p_i\rangle$ $\lvert p_i\overline{p_i}p_j\overline{p_k}\rangle + \lvert p_i\overline{p_ip_j}p_k\rangle$ $\lvert sp_ip_j\overline{p_j}\rangle + \lvert sp_ip_k\overline{p_k}\rangle$ $\lvert \overline{sp_i}p_j\overline{p_j}\rangle + \lvert \overline{sp_i}p_k\overline{p_k}\rangle$ $\lvert s\overline{p_i}p_j\overline{p_j}\rangle + \lvert \overline{s}p_ip_j\overline{p_j}\rangle + \lvert s\overline{p_i}p_k\overline{p_k}\rangle + \lvert \overline{s}p_ip_k\overline{p_k}\rangle$	
$\langle \underline{L}^2 \rangle = 6$	$\lvert s\overline{s}p_i\overline{p_j}\rangle - \lvert s\overline{sp_i}p_j\rangle$ $a\lvert s\overline{s}p_1\overline{p_1}\rangle + b\lvert s\overline{s}p_2\overline{p_2}\rangle + c\lvert s\overline{s}p_3\overline{p_3}\rangle$ $\lvert p_i\overline{p_i}p_j\overline{p_k}\rangle - \lvert p_i\overline{p_ip_j}p_k\rangle$ $a\lvert p_1\overline{p_1}p_2\overline{p_2}\rangle + b\lvert p_1\overline{p_1}p_3\overline{p_3}\rangle + c\lvert p_2\overline{p_2}p_3\overline{p_3}\rangle$ $\lvert s\overline{p_i}p_j\overline{p_j}\rangle - \lvert \overline{s}p_ip_j\overline{p_j}\rangle - \lvert s\overline{p_i}p_k\overline{p_k}\rangle + \lvert \overline{s}p_ip_k\overline{p_k}\rangle$ $a(\lvert s\overline{p_1p_2}p_3\rangle + \lvert \overline{s}p_1p_2\overline{p_3}\rangle) + b(\lvert sp_1\overline{p_2p_3}\rangle + \lvert \overline{sp_1}p_2p_3\rangle)$ $+c(\lvert s\overline{p_1}p_2\overline{p_3}\rangle + \lvert \overline{s}p_1\overline{p_2}p_3\rangle)$	$\lvert sp_ip_j\overline{p_j}\rangle - \lvert sp_ip_k\overline{p_k}\rangle$ $\lvert \overline{sp_i}p_j\overline{p_j}\rangle - \lvert \overline{sp_i}p_k\overline{p_k}\rangle$ $a\lvert sp_1p_2\overline{p_3}\rangle + b\lvert sp_1\overline{p_2}p_3\rangle + c\lvert s\overline{p_1}p_2p_3\rangle$ $a\lvert \overline{sp_1p_2}p_3\rangle + b\lvert \overline{sp_1}p_2\overline{p_3}\rangle + c\lvert \overline{s}p_1\overline{p_2p_3}\rangle$ $\lvert s\overline{p_i}p_j\overline{p_j}\rangle + \lvert \overline{s}p_ip_j\overline{p_j}\rangle - \lvert s\overline{p_i}p_k\overline{p_k}\rangle - \lvert \overline{s}p_ip_k\overline{p_k}\rangle$ $a(\lvert s\overline{p_1p_2}p_3\rangle - \lvert \overline{s}p_1p_2\overline{p_3}\rangle) + b(\lvert sp_1\overline{p_2p_3}\rangle - \lvert \overline{sp_1}p_2p_3\rangle)$ $+c(\lvert s\overline{p_1}p_2\overline{p_3}\rangle - \lvert \overline{s}p_1\overline{p_2}p_3\rangle)$	

Table 3.4: Carbon $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$	$\langle \underline{S}^2 \rangle = \frac{15}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert sp_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert sp_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert sp_2\overline{p_2}p_3\overline{p_3}\rangle$ $\lvert \overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert \overline{s}p_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert \overline{s}p_2\overline{p_2}p_3\overline{p_3}\rangle$	$\lvert s\overline{s}p_1p_2p_3\rangle$ $\lvert s\overline{sp_1p_2p_3}\rangle$ $\lvert s\overline{s}p_1\overline{p_2}p_3\rangle + \lvert s\overline{sp_1}p_2p_3\rangle + \lvert s\overline{s}p_1p_2\overline{p_3}\rangle$ $\lvert s\overline{s}p_1\overline{p_2p_3}\rangle + \lvert s\overline{sp_1}p_2\overline{p_3}\rangle + \lvert s\overline{sp_1p_2}p_3\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{s}p_ip_j\overline{p_j}\rangle + \lvert s\overline{s}p_ip_k\overline{p_k}\rangle$ $\lvert s\overline{sp_i}p_j\overline{p_j}\rangle + \lvert s\overline{sp_i}p_k\overline{p_k}\rangle$ $\lvert p_1p_2p_3\overline{p_jp_k}\rangle$ $\lvert \overline{p_1p_2p_3}p_jp_k\rangle$ $2\lvert \overline{s}p_i\overline{p_i}p_jp_k\rangle - \lvert sp_i\overline{p_ip_j}p_k\rangle - \lvert sp_i\overline{p_i}p_j\overline{p_k}\rangle$ $2\lvert sp_i\overline{p_ip_jp_k}\rangle - \lvert \overline{s}p_i\overline{p_i}p_j\overline{p_k}\rangle - \lvert \overline{s}p_i\overline{p_ip_j}p_k\rangle$	$\lvert sp_1p_2p_3\overline{p_i}\rangle$ $\lvert \overline{sp_1p_2p_3}p_i\rangle$ $\lvert sp_i\overline{p_i}p_j\overline{p_k}\rangle + \lvert sp_i\overline{p_ip_j}p_k\rangle + \lvert \overline{s}p_i\overline{p_i}p_jp_k\rangle$ $\lvert sp_i\overline{p_ip_jp_k}\rangle + \lvert \overline{s}p_i\overline{p_i}p_j\overline{p_k}\rangle + \lvert \overline{s}p_i\overline{p_ip_j}p_k\rangle$
$\langle \underline{L}^2 \rangle = 6$	$\lvert s\overline{s}p_ip_j\overline{p_j}\rangle - \lvert s\overline{s}p_ip_k\overline{p_k}\rangle$ $\lvert s\overline{sp_i}p_j\overline{p_j}\rangle - \lvert s\overline{sp_i}p_k\overline{p_k}\rangle$ $a\lvert s\overline{s}p_1p_2\overline{p_3}\rangle + b\lvert s\overline{s}p_1\overline{p_2}p_3\rangle + c\lvert s\overline{sp_1}p_2p_3\rangle$ $a\lvert s\overline{sp_1p_2}p_3\rangle + b\lvert s\overline{s}p_1\overline{p_2p_3}\rangle + c\lvert s\overline{sp_1}p_2\overline{p_3}\rangle$ $\lvert sp_i\overline{p_i}p_j\overline{p_k}\rangle - \lvert sp_i\overline{p_ip_j}p_k\rangle$ $\lvert \overline{s}p_i\overline{p_i}p_j\overline{p_k}\rangle - \lvert \overline{s}p_i\overline{p_ip_j}p_k\rangle$ $a\lvert sp_1\overline{p_1}p_2\overline{p_2}\rangle + b\lvert sp_1\overline{p_1}p_3\overline{p_3}\rangle + c\lvert sp_2\overline{p_2}p_3\overline{p_3}\rangle$ $a\lvert \overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle + b\lvert \overline{s}p_1\overline{p_1}p_3\overline{p_3}\rangle + c\lvert \overline{s}p_2\overline{p_2}p_3\overline{p_3}\rangle$	

Table 3.5: Nitrogen $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert s\overline{s}p_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert s\overline{s}p_2\overline{p_2}p_3\overline{p_3}\rangle$ $\lvert p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$	
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{p_1p_2p_3}p_jp_k\rangle - \lvert \overline{s}p_1p_2p_3\overline{p_jp_k}\rangle$	$\lvert s\overline{s}p_1p_2p_3\overline{p_i}\rangle \qquad \lvert s\overline{sp_1p_2p_3}p_i\rangle$ $\lvert sp_1p_2p_3\overline{p_jp_k}\rangle \qquad \lvert \overline{sp_1p_2p_3}p_jp_k\rangle$ $\lvert s\overline{s}p_i\overline{p_i}p_j\overline{p_k}\rangle + \lvert s\overline{s}p_i\overline{p_ip_j}p_k\rangle$ $\lvert s\overline{p_1p_2p_3}p_jp_k\rangle + \lvert \overline{s}p_1p_2p_3\overline{p_jp_k}\rangle$
$\langle \underline{L}^2 \rangle = 6$	$\lvert s\overline{s}p_i\overline{p_i}p_j\overline{p_k}\rangle - \lvert s\overline{s}p_i\overline{p_ip_j}p_j\rangle$ $a\lvert s\overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle + b\lvert s\overline{s}p_1\overline{p_1}p_3\overline{p_3}\rangle + c\lvert s\overline{s}p_2\overline{p_2}p_3\overline{p_3}\rangle$	

Table 3.6: Oxygen $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert sp_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle \qquad \lvert \overline{s}p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{s}p_1p_2p_3\overline{p_ip_j}\rangle \qquad \lvert s\overline{sp_1p_2p_3}p_ip_j\rangle$

Table 3.7: Fluorine $\underline{L}^2$-$\underline{S}^2$ eigenspaces.

the dimension of these eigenspaces can be determined from the dimension of the $\underline{L}^2$-$\underline{S}^2$ eigenspaces. The dimension of an eigenspace with $\underline{L}^2$ eigenvalue $\ell(\ell+1)$ and $\underline{S}^2$ eigenvalue $s(s+1)$ must be a multiple of $d := (2\ell+1)(2s+1)$. This follows from the fact that the minimal number of L_3 eigenfunctions is $2\ell+1$, with eigenvalues $-\ell, \ldots, 0, \ldots, \ell$, and the minimal number of S_3 eigenfunctions is $2s+1$, with eigenvalues $-s, \ldots, 0, \ldots, s$. The dimension of each $\underline{L}^2$-$\underline{S}^2$-L_3-S_3 is then given by the dimension of the $\underline{L}^2$-$\underline{S}^2$ eigenspace divided by d.

The dimension of each $\underline{L}^2$-$\underline{S}^2$ eigenspace and the corresponding value of d are given in Table 3.8.

$\langle \underline{L}^2 \rangle$	$\langle \underline{S}^2 \rangle$	d	Li	Be	B	C	N	O	F
0	0	1		2		2		2	
0	$\frac{3}{4}$	2	2		2		2		2
0	2	3				3			
0	$\frac{15}{4}$	4			4		4		
0	6	5				5			
2	0	3		3		3		3	
2	$\frac{3}{4}$	6	6		18		18		6
2	2	9		18		27		18	
2	$\frac{15}{4}$	12			12		12		
6	0	5		5		15		5	
6	$\frac{3}{4}$	10			20		20		
6	2	15				15			
Total			8	28	56	70	56	28	8

Table 3.8: Dimensions of the $\underline{L}^2$-$\underline{S}^2$ simultaneous eigenspaces for Lithium to Neon. d is the minimal dimension of a given $\underline{L}^2$-$\underline{S}^2$ eigenspace.

3.2.3 Explicit $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ Simultaneous Eigenspaces

The simultaneous $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces are shown in Tables 3.9-3.15, again with the conventions that we do not show the $1s\overline{1s}$ orbitals and that the eigenfunctions are not normalized. The parity subspaces are separated by a line.

Investigating the $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces shows a number of interesting properties. Firstly we note that the maximum dimension is two, which is clearly much more promising than the 70-dimensional space of Carbon.

Further, all Slater determinants in any $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspace given in Tables 3.9-3.15 differ by an even number of one-electron orbitals. We propose that that this is an intrinsic property of the operators and choice of basis. We have

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\vert s \rangle$
$\langle \underline{L}^2 \rangle = 2$	$\vert p_3 \rangle$

Table 3.9: Lithium $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$
$\langle \underline{L}^2 \rangle = 0$	$\vert s\overline{s} \rangle$ $\vert p_1\overline{p_1} \rangle + \vert p_2\overline{p_2} \rangle + \vert p_3\overline{p_3} \rangle$	
$\langle \underline{L}^2 \rangle = 2$	$\vert s\overline{p_3} \rangle - \vert \overline{s}p_3 \rangle$	$\vert sp_3 \rangle$
		$\vert p_1p_2 \rangle$
$\langle \underline{L}^2 \rangle = 6$	$2\vert p_3\overline{p_3} \rangle - \vert p_1\overline{p_1} \rangle - \vert p_2\overline{p_2} \rangle$	

Table 3.10: Beryllium $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$	$\langle \underline{S}^2 \rangle = \frac{15}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\vert sp_1\overline{p_1} \rangle + \vert sp_2\overline{p_2} \rangle + \vert sp_3\overline{p_3} \rangle$	$\vert p_1p_2p_3 \rangle$
$\langle \underline{L}^2 \rangle = 2$	$\vert s\overline{s}p_3 \rangle$ $\vert p_3p_1\overline{p_1} \rangle + \vert p_3p_2\overline{p_2} \rangle$	$\vert sp_1p_2 \rangle$
	$2\vert \overline{s}p_1p_2 \rangle - \vert s\overline{p_1}p_2 \rangle - \vert sp_1\overline{p_2} \rangle$	
$\langle \underline{L}^2 \rangle = 6$	$2\vert sp_3\overline{p_3} \rangle - \vert sp_1\overline{p_1} \rangle - \vert sp_2\overline{p_2} \rangle$	
	$2\vert \overline{p_3}p_1p_2 \rangle - \vert p_3p_1\overline{p_2} \rangle - \vert p_3\overline{p_1}p_2 \rangle$	

Table 3.11: Boron $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$	$\langle \underline{S}^2 \rangle = 6$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\bar{s}p_1\overline{p_1}\rangle + \lvert s\bar{s}p_2\overline{p_2}\rangle + \lvert s\bar{s}p_3\overline{p_3}\rangle$ $\lvert p_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert p_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert p_2\overline{p_2}p_3\overline{p_3}\rangle$	$3\lvert \bar{s}p_3p_1p_2\rangle - \lvert s\overline{p_3}p_1p_2\rangle$ $-\lvert sp_3\overline{p_1}p_2\rangle - \lvert sp_3p_1\overline{p_2}\rangle$	$\lvert sp_1p_2p_3\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{p_3}p_1\overline{p_1}\rangle - \lvert \bar{s}p_3p_1\overline{p_1}\rangle + \lvert s\overline{p_3}p_2\overline{p_2}\rangle - \lvert \bar{s}p_3p_2\overline{p_2}\rangle$	$\lvert s\bar{s}p_1p_2\rangle$ $\lvert p_3\overline{p_3}p_1p_2\rangle$	
		$\lvert sp_3p_1\overline{p_1}\rangle + \lvert sp_3p_2\overline{p_2}\rangle$	
$\langle \underline{L}^2 \rangle = 6$	$2\lvert s\bar{s}p_3\overline{p_3}\rangle - \lvert s\bar{s}p_1\overline{p_1}\rangle - \lvert s\bar{s}p_2\overline{p_2}\rangle$ $2\lvert p_1\overline{p_1}p_2\overline{p_2}\rangle - \lvert p_3\overline{p_3}p_1\overline{p_1}\rangle - \lvert p_3\overline{p_3}p_2\overline{p_2}\rangle$	$2\lvert s\overline{p_3}p_1p_2\rangle - \lvert sp_3p_1\overline{p_2}\rangle - \lvert sp_3\overline{p_1}p_2\rangle$	
	$2\lvert sp_3\overline{p_1}\overline{p_2}\rangle - \lvert s\overline{p_3}p_1\overline{p_2}\rangle - \lvert s\overline{p_3}\overline{p_1}p_2\rangle$ $+2\lvert \bar{s}\overline{p_3}p_1p_2\rangle - \lvert \bar{s}p_3p_1\overline{p_2}\rangle - \lvert \bar{s}p_3\overline{p_1}p_2\rangle$		

Table 3.12: Carbon $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$	$\langle \underline{S}^2 \rangle = \frac{15}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert sp_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert sp_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert sp_2\overline{p_2}p_3\overline{p_3}\rangle$	$\lvert s\overline{s}p_1p_2p_3\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert p_3p_1\overline{p_1}p_2\overline{p_2}\rangle$ $\lvert s\overline{s}p_3p_1\overline{p_1}\rangle + \lvert s\overline{s}p_3p_2\overline{p_2}\rangle$	$\lvert sp_3\overline{p_3}p_1p_2\rangle$
	$2\lvert \overline{s}p_3\overline{p_3}p_1p_2\rangle - \lvert sp_3\overline{p_3}\overline{p_1}p_2\rangle - \lvert sp_3\overline{p_3}p_1\overline{p_2}\rangle$	
$\langle \underline{L}^2 \rangle = 6$	$2\lvert s\overline{s}\overline{p_3}p_1p_2\rangle - \lvert s\overline{s}p_3p_1\overline{p_2}\rangle - \lvert s\overline{s}p_3\overline{p_1}p_2\rangle$	
	$2\lvert sp_1\overline{p_1}p_2\overline{p_2}\rangle - \lvert sp_3\overline{p_3}p_1\overline{p_1}\rangle - \lvert sp_3\overline{p_3}p_2\overline{p_2}\rangle$	

Table 3.13: Nitrogen $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = 0$	$\langle \underline{S}^2 \rangle = 2$
$\langle \underline{L}^2 \rangle = 0$	$\lvert s\overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle + \lvert s\overline{s}p_1\overline{p_1}p_3\overline{p_3}\rangle + \lvert s\overline{s}p_2\overline{p_2}p_3\overline{p_3}\rangle$ $\lvert p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$	
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{p_3}p_1\overline{p_1}p_2\overline{p_2}\rangle - \lvert \overline{s}p_3p_1\overline{p_1}p_2\overline{p_2}\rangle$	$\lvert sp_3p_1\overline{p_1}p_2\overline{p_2}\rangle$
		$\lvert s\overline{s}p_3\overline{p_3}p_1p_2\rangle$
$\langle \underline{L}^2 \rangle = 6$	$2\lvert s\overline{s}p_1\overline{p_1}p_2\overline{p_2}\rangle - \lvert s\overline{s}p_3\overline{p_3}p_1\overline{p_1}\rangle - \lvert s\overline{s}p_3\overline{p_3}p_2\overline{p_2}\rangle$	

Table 3.14: Oxygen $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

	$\langle \underline{S}^2 \rangle = \frac{3}{4}$
$\langle \underline{L}^2 \rangle = 0$	$\lvert sp_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$
$\langle \underline{L}^2 \rangle = 2$	$\lvert s\overline{s}p_3p_1\overline{p_1}p_2\overline{p_2}\rangle$

Table 3.15: Fluorine $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces.

proved this by exhaustive search but this does nothing to aid in understanding why this is true and hopefully a rigorous mathematical proof will be of benefit.

3.3 Excitation In $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ Eigenspaces

Theorem 3.3.1 (Excitation in Simultaneous Eigenspaces). *For every $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ simultaneous eigenspace of linear combinations of Slater determinants containing the pair of one-electron orbitals $1s\overline{1s}$ and N one-electron orbitals from the set $\{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\}$, with L_3 eigenvalue equal to zero, all Slater determinants within the simultaneous eigenspace differ by an even number of orbitals.*

This result is entirely novel, which is unsurprising given that the eigenspaces in the previous section are also new. Correspondingly, the operators of Definition 3.3.5 and the Lemmas used in the proof are also previously unknown.

Before we proceed to prove this result we give a number of definitions for frequently used ideas:

Definition 3.3.2. *Denote by X_N the set of all Slater determinants which contain the one-electron orbitals $1s$ and $\overline{1s}$ along with N distinct one-electron orbitals from the set $\{2s, \overline{2s}, p_1, \overline{p_1}, p_2, \overline{p_2}, p_3, \overline{p_3}\}$. Denote the set of all finite linear combinations of $|\Psi_i\rangle \in X_N$ by Σ_N.*

Definition 3.3.3. *Suppose $|\Psi\rangle \in X_N$ is such that there exist $I \in \mathbb{N}$, $|\Psi_i\rangle \in X_N$ and $c_i \in \mathbb{C}$ such that for $\Phi := |\Psi\rangle + \sum_{i=1}^{I} c_i |\Psi_i\rangle$ we have $L_3\Phi = 0$. Denote the set of all such $|\Psi\rangle$ by X_N^0 and the set of all such Φ by Σ_N^0.*

Definition 3.3.4. *Let A be an operator such that $A : X_N \to Span\{X_N\}$. For each Slater determinant $|\Psi\rangle \in X_N$, we denote the set of all Slater determinants in $A|\Psi\rangle$, along with $|\Psi\rangle$, by $\{A|\Psi\rangle\}$.*

We note here that since X_N^0 contains all Slater determinants which may form L_3 eigenfunctions with eigenvalue zero, any operator A which commutes with L_3 must satisfy $\{A|\Psi\rangle\} \subset X_N^0$ for all $|\Psi\rangle \in X_N^0$.

The aim of this section is to find suitable operators which encode the necessary information to show that for an odd excitation $|\Psi\rangle \mapsto |\Phi\rangle$ the sets of of eigenvalues of $\underline{L}^2$, $\underline{S}^2$, L_3, S_3 and $\hat{R}$ for eigenfunctions containing $|\Psi\rangle$ and $|\Phi\rangle$ are disjoint. An obvious property to look for in these operators is that they commute with $\underline{L}^2$, $\underline{S}^2$, L_3 S_3 and $\hat{R}$,

Definition 3.3.5. *We define the following four operators that can be thought of as counting the number of p-orbital spin pairs, spin up p-orbitals, spin down p-orbitals and total number of p-orbitals respectively:*

$$\hat{P} := \sum_{i=1}^{3} a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i) \tag{3.14}$$

$$\hat{Q}_\uparrow := \sum_{i=1}^{3} a^\dagger(p_i)a(p_i) \tag{3.15}$$

$$\hat{Q}_\downarrow := \sum_{i=1}^{3} a^\dagger(\overline{p_i})a(\overline{p_i}) \tag{3.16}$$

$$\hat{Q} := \hat{Q}_\uparrow + \hat{Q}_\downarrow \tag{3.17}$$

The expectation that these operators commute with $\underline{L}^2$ is motivated both by the eigenspace tables and by the second quantized form of $\underline{L}^2$ restricted to Σ_N:

$$\begin{aligned}\underline{L}^2 = {} & 2\sum_{i=1}^{3} a^\dagger(p_i)a(p_i) + a^\dagger(\overline{p_i})a(\overline{p_i}) \\ & - \sum_{\substack{i,j=1\\ i\neq j}}^{3} \big[2a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i) + 2a^\dagger(\overline{p_i})a^\dagger(p_j)a(\overline{p_j})a(p_i) \\ & \qquad + a^\dagger(p_i)a^\dagger(p_j)a(p_j)a(p_i) + a^\dagger(\overline{p_i})a^\dagger(\overline{p_j})a(\overline{p_j})a(\overline{p_i})\big].\end{aligned} \tag{3.18}$$

We first note that the $i \neq j$ restriction on the second sum is not necessary, the $i = j$ terms cancel, however, this form simplifies the following arguments.

It is clear that the first two terms of (3.18) are simply $2\hat{Q}_\uparrow + 2\hat{Q}_\downarrow$. Now consider $\hat{Q}_\uparrow^2$:

$$\begin{aligned}\hat{Q}_\uparrow^2 &= \sum_{i=1}^{3} a^\dagger(p_i)a(p_i)\sum_{j=1}^{3} a^\dagger(p_j)a(p_j) = \sum_{i=1}^{3} a^\dagger(p_i)a(p_i) + \sum_{\substack{i,j=1\\ i\neq j}}^{3} a^\dagger(p_i)a(p_i)a^\dagger(p_j)a(p_j) \\ &= \hat{Q}_\uparrow + \sum_{\substack{i,j=1\\ i\neq j}}^{3} a^\dagger(p_i)a(p_i)a^\dagger(p_j)a(p_j).\end{aligned}$$

The formula for $\hat{Q}_\downarrow^2$ is analogous and hence

$$\underline{L}^2 = 3(\hat{Q}_\uparrow + \hat{Q}_\downarrow) - \hat{Q}_\uparrow^2 - \hat{Q}_\downarrow^2 \tag{3.19}$$

$$- \sum_{\substack{i,j=1 \\ i\neq j}}^{3} \left[2a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i) + 2a^\dagger(\overline{p_i})a^\dagger(p_j)a(\overline{p_j})a(p_i)\right] \tag{3.20}$$

The eigenvalue associated with (3.19) is the expected value of $\underline{L}^2$ on a Slater determinant.

We next want to show that the operators defined in Definition 3.3.5 commute with themselves and with $\underline{L}^2$, we then aim to use these operators to classify the eigenvalues of $\underline{L}^2$ and prove Theorem 3.3.1. In the following, we use the second-quantization identities (e.g. [SO96])

$$\begin{aligned} a^\dagger(\chi_i)a(\chi_j) &= \delta_{i,j} - a(\chi_j)a^\dagger(\chi_i) \\ a^\dagger(\chi_i)a^\dagger(\chi_j) &= (\delta_{i,j} - 1)a^\dagger(\chi_j)a^\dagger(\chi_i) \\ a(\chi_i)a(\chi_j) &= (\delta_{i,j} - 1)a(\chi_j)a(\chi_i), \end{aligned}$$

where δ tests both spatial and spin parts.

Recall the the parity operator $\hat{R}$ as given in (3.3), which for $|\Psi\rangle \in X_N$ is equivalent to the oddness or evenness of $\langle\hat{Q}\rangle_{|\Psi\rangle}$, i.e. $\langle\hat{R}\rangle_{|\Psi\rangle} = \langle\hat{Q}\rangle_{|\Psi\rangle} \mod 2$. This is because s-orbitals are unchanged by inversion whilst p-orbitals change sign under inversion. Clearly the parity is not well-defined (i.e. not ±1) for a general wavefunction. However, as shown in Section 3.1.3, it does commute with $\underline{L}^2$, $\underline{S}^2$, L_3 and S_3. Further it commutes with the operators defined above:

Lemma 3.3.6. *The operators $\hat{P}$, $\hat{Q}_\uparrow$, $\hat{Q}_\downarrow$ and $\hat{R}$ commute with each other on X_N.*

Proof The proof is trivial since each $|\Psi\rangle \in X_N$ is an eigenfunction of all three operators. □

We next consider how our new operators interact with the angular momentum and spin operators:

Lemma 3.3.7. *The operators $\hat{P}$, $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ commute with $\underline{L}^2$ on X_N.*

Proof Using the form of $\underline{L}^2$ given in (3.19) and (3.20), and since, by Lemma 3.3.6, $\hat{Q}_\uparrow$, $\hat{Q}_\downarrow$ and $\hat{P}$ commute, we need only compute their commutators with the two terms in (3.20):

$$\sum_{\substack{j,k=1 \\ j\neq k}}^{3} a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_k})a(p_k) \tag{3.21}$$

and

$$\sum_{\substack{j,k=1\\ j\neq k}}^{3} a^\dagger(\overline{p_j})a^\dagger(p_k)a(\overline{p_k})a(p_j), \tag{3.22}$$

where ignoring the coefficients has no effect on the commutation relations.

Consider one element of $\hat{Q}_\uparrow$, namely $a^\dagger(p_i)a(p_i)$. It is clear that each term of $[a^\dagger(p_i)a(p_i), (3.21)]$ is zero unless j or k is equal to i. If $j = i$ then we have

$$a^\dagger(p_i)a(p_i)\cdot a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k) - a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k)\cdot a^\dagger(p_i)a(p_i),$$

which, since $k \neq i$, simplifies to $a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k)$.

Similarly, if $k = i$ then the term is

$$a^\dagger(p_i)a(p_i)\cdot a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i) - a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i)\cdot a^\dagger(p_i)a(p_i)$$

and since $j \neq i$ this is equivalent to $-a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i)$.

It follows that

$$[\hat{Q}_\uparrow, (3.21)] = \sum_{i=1}^{3}\sum_{\substack{j=1\\ j\neq i}}^{3} a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_j})a(p_j) - a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i).$$

Noting that if $i = j$ the summand gives zero, including this term and interchanging the dummy variables in the second term of the summand gives

$$[\hat{Q}_\uparrow, (3.21)] = \sum_{i,j=1}^{3} a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_j})a(p_j) - a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_j})a(p_j) = 0$$

The proofs that $[\hat{Q}_\uparrow, (3.22)]$, $[\hat{Q}_\downarrow, (3.21)]$ and $[\hat{Q}_\downarrow, (3.22)]$ are all zero are completely analogous. It follows that $[\hat{Q}_\uparrow, \underline{L}^2] = [\hat{Q}_\downarrow, \underline{L}^2] = 0$, and it remains to show that $[\hat{P}, \underline{L}^2] = 0$.

We follow a similar method to the above proof, starting with $a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i)$ and considering the commutator with (3.21). Again the only non-trivial cases are when $j = i$ or $k = i$. If $j = i$ then we have

$$\begin{aligned}&a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i)\cdot a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k)\\ &\qquad\qquad -a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k)\cdot a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i)\end{aligned}$$

and since $k \neq i$ this reduces to $a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_k})a(p_k)$. Similarly, if $k = i$ (and so $j \neq i$), we obtain $-a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i)$.

Hence

$$[\hat{P},(3.21)] = \sum_{i=1}^{3}\sum_{\substack{j=1\\ j\neq i}}^{3} a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_j})a(p_j) - a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i),$$

which is, as before, equal to zero.

The proof that $[\hat{P},(3.22)] = 0$ is again analogous, this time noting that when $j = i$ we obtain, with $k \neq i$,

$$\begin{aligned} &a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i)\cdot a^\dagger(\overline{p_i})a^\dagger(p_k)a(\overline{p_k})a(p_i) \\ &\qquad -a^\dagger(\overline{p_i})a^\dagger(p_k)a(\overline{p_k})a(p_i)\cdot a^\dagger(p_i)a^\dagger(\overline{p_i})a(\overline{p_i})a(p_i) = 0 \end{aligned}$$

and the term is similarly zero when $k = i$, giving that the commutator is trivially also zero and the result follows. ☐

Lemma 3.3.8. *The operators $\hat{P}$, $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ commute with S_3 on X_N.*

Proof This is trivial since all Slater determinants are eigenfunctions of all four operators. ☐

Lemma 3.3.9. *The operators $\hat{P}$ and $\hat{Q}$ commute with $\underline{S}^2$ on X_N.*

Proof Since $\hat{Q}$ is independent of spin it is clear that it commutes with $\underline{S}^2$. For $\hat{P}$ we begin by giving the second quantized form of $\underline{S}^2$ restricted to X_N:

$$\begin{aligned} \underline{S}^2 = \tfrac{3}{4}\sum_{i=1}^{5} &\big[a^\dagger(\psi_i)a(\psi_i) + a^\dagger(\overline{\psi_i})a(\overline{\psi_i})\big] \\ +\tfrac{1}{2}\sum_{1\le i<j\le 5} &\big[a^\dagger(\overline{\psi_i})a^\dagger(\psi_j)a(\overline{\psi_j})a(\psi_i) + a^\dagger(\overline{\psi_j})a^\dagger(\psi_i)a(\overline{\psi_i})a(\psi_j) \\ &- a^\dagger(\psi_i)a^\dagger(\overline{\psi_j})a(\overline{\psi_j})a(\psi_i) - a^\dagger(\psi_j)a^\dagger(\overline{\psi_i})a(\overline{\psi_i})a(\psi_j) \\ &+ a^\dagger(\psi_i)a^\dagger(\psi_j)a(\psi_j)a(\psi_i) + a^\dagger(\overline{\psi_i})a^\dagger(\overline{\psi_j})a(\overline{\psi_j})a(\overline{\psi_i})\big], \end{aligned} \tag{3.23}$$

where the sums run over the five spatial orbitals. It is clear that, when acting on a single Slater determinant, all terms in (3.23) give a multiple of the identity, except for the first two terms in the double sum, which swap the spins of two different spatial orbitals. Hence, with the exception of these two terms, $\hat{P}$ commutes with $\underline{S}^2$.

It is clear that all terms in $\hat{P}$ will commute with $a^\dagger(\overline{\psi_i})a^\dagger(\psi_j)a(\overline{\psi_j})a(\psi_i)$ and $a^\dagger(\overline{\psi_j})a^\dagger(\psi_i)a(\overline{\psi_i})a(\psi_j)$ unless $\psi_i = p_k$ or $\psi_j = p_k$, for some $k \in \{1,2,3\}$, in which case we are interested in only the $a^\dagger(p_k)a(p_k)a^\dagger(\overline{p_k})a(\overline{p_k})$ term from $\hat{P}$. It is trivial to see that, for $\psi \neq p_k$,

$$\begin{aligned}
&a^\dagger(p_k)\underline{a(p_k)}a^\dagger(\overline{p_k})a(\overline{p_k}) \cdot a^\dagger(\overline{p_k})a^\dagger(\psi)a(\overline{\psi})\underline{a(p_k)} = 0, \\
&\underline{a^\dagger(\overline{p_k})}a^\dagger(\psi)a(\overline{\psi})a(p_k) \cdot a^\dagger(p_k)a(p_k)\underline{a^\dagger(\overline{p_k})}a(\overline{p_k}) = 0, \\
&a^\dagger(p_k)a(p_k)a^\dagger(\overline{p_k})\underline{a(\overline{p_k})} \cdot a^\dagger(\overline{\psi})a^\dagger(p_k)\underline{a(\overline{p_k})}a(\psi) = 0, \text{ and} \\
&a^\dagger(\overline{\psi})\underline{a^\dagger(p_k)}a(\overline{p_k})a(\psi) \cdot \underline{a^\dagger(p_k)}a(p_k)a^\dagger(\overline{p_k})a(\overline{p_k}) = 0,
\end{aligned}$$

where we have underlined the operators which should be anticommuted next to each other to give zero. Hence the commutators are also zero and the result follows.

This can also be seen intuitively as the spin swap must act on unpaired spin orbitals or the resultant Slater determinant will be zero and thus $\underline{S}^2$ has no effect on spin pairs. ☐

Note that the stronger result that $\underline{S}^2$ commutes with $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ separately is not true. This is due to spin exchange between s- and p-orbitals given by the first two terms in the double sum in (3.23).

Lemma 3.3.10. *The operators $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ commute with L_3 on X_N*

Proof We begin by showing that $[\hat{Q}_\uparrow, L_3] = 0$. Writing L_3 restricted to X_N in second quantized form gives

$$L_3 = i\left[a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2) + a^\dagger(\overline{p_2})a(\overline{p_1}) - a^\dagger(\overline{p_1})a(\overline{p_2})\right].$$

For ease of calculation we drop the factor i and note that any commutator relation still holds. We then have

$$\hat{Q}_\uparrow L_3 = \sum_{i=1}^{3} a^\dagger(p_i)a(p_i)[a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2) + a^\dagger(\overline{p_2})a(\overline{p_1}) - a^\dagger(\overline{p_1})a(\overline{p_2})],$$

and it is clear that the last two terms of L_3 commute for all i and any terms with $i = 3$ also commute. Since the same applies for $L_3\hat{Q}_\uparrow$, this gives

$$\begin{aligned}
\hat{Q}_\uparrow L_3 - L_3\hat{Q}_\uparrow = \sum_{i=1}^{2} & a^\dagger(p_i)a(p_i)[a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2)] \\
& - [a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2)]a^\dagger(p_i)a(p_i) \\
= & -a^\dagger(p_1)a(p_1)a^\dagger(p_1)a(p_2) - a^\dagger(p_2)a(p_1)a^\dagger(p_1)a(p_1)
\end{aligned}$$

$$
\begin{aligned}
&+ a^\dagger(p_2)a(p_2)a^\dagger(p_2)a(p_1) + a^\dagger(p_1)a(p_2)a^\dagger(p_2)a(p_2)\\
&= -a^\dagger(p_1)a(p_2) - a^\dagger(p_2)a(p_1) + a^\dagger(p_1)a(p_2) + a^\dagger(p_2)a(p_1)\\
&= 0.
\end{aligned}
$$

The case for $\hat{Q}_\downarrow$ is analogous, simply taking the spin down rather than spin up terms from L_3. □

We note that $\hat{P}$ clearly does not commute with L_3 in general, for example

$$L_3\hat{P}|p_1\overline{p_1}\rangle = L_3|p_1\overline{p_1}\rangle = i|p_2\overline{p_1}\rangle + i|p_1\overline{p_2}\rangle$$

whereas

$$\hat{P}L_3|p_1\overline{p_1}\rangle = \hat{P}(i|p_2\overline{p_1}\rangle + i|p_1\overline{p_2}\rangle) = 0.$$

We require one further result which differentiates the $L_3\Psi = 0$ eigenspace from the other L_3 eigenspaces:

Lemma 3.3.11. *Let $|\Psi\rangle \in X_N^0$, then the total number of $p_1, \overline{p_1}, p_2$ and $\overline{p_2}$ orbitals in $|\Psi\rangle$ is even.*

Proof Suppose that the total number of p_1 and p_2 orbitals in a given Slater determinant $|\Psi\rangle$ is odd. If both p_1 and p_2 of the same spin are present, L_3 gives zero on these orbitals. Hence, without loss of generality, we can consider the action of L_3 on only $|p_1\rangle$ and $|p_2\rangle$ for which we have $L_3|p_1\rangle = i|p_2\rangle$ and $L_3|p_2\rangle = -i|p_1\rangle$, from which it is clear that the L_3 eigenvalues are ±1 and the result follows. □

We now move on to investigate the $\underline{L}^2$ matrix in terms of eigenvalues of these operators.

Lemma 3.3.12 ($\underline{L}^2$ Matrix in P, $Q_\uparrow$ and $Q_\downarrow$)**.** *Let $|\Psi_1\rangle \in X_N$ satisfy*

$$\hat{P}|\Psi_1\rangle = P|\Psi_1\rangle, \quad \hat{Q}_\uparrow|\Psi_1\rangle = Q_\uparrow|\Psi_1\rangle, \quad \hat{Q}_\downarrow|\Psi_1\rangle = Q_\downarrow|\Psi_1\rangle,$$

and consider the minimal set of Slater determinants $A := \{|\Psi_1\rangle, \dots, |\Psi_D\rangle\}$ such that $\underline{L}^2|\Psi_i\rangle \in$ Span A for $i = 1, \dots, D$. Then $A = \{\underline{L}^2|\Psi_1\rangle\}$, the dimension of which is given by

$$D := 1 + (3 + P - Q_\uparrow - Q_\downarrow)P + (Q_\uparrow - P)(Q_\downarrow - P) \in \{1, 2, 3\}.$$

Furthermore, the diagonal entries of the $D \times D$ matrix $\langle\Psi_i|\underline{L}^2|\Psi_j\rangle$ for A are all *equal to F where*

$$F := Q_\uparrow(3 - Q_\uparrow) + Q_\downarrow(3 - Q_\downarrow) \in \{0, 2, 4\}.$$

Proof The diagonal elements of the matrix are given by $\langle \Psi_i | \underline{L}^2 | \Psi_i \rangle$ and from (3.19) and (3.20) we see that this is given by $F := 3(Q_\uparrow + Q_\downarrow) - Q_\uparrow^2 - Q_\downarrow^2 = Q_\uparrow(3-Q_\uparrow)+Q_\downarrow(3-Q_\downarrow)$ (which is conserved under $\underline{L}^2$). Clearly, $Q_\uparrow, Q_\downarrow \in \{0, \dots, 3\}$ and are independent, which gives that $Q_\alpha(3 - Q_\alpha) \in \{0, 2\}$ and the range of F follows.

We now need to consider the dimension of the matrix, which is equivalent to the dimension of the minimal $\underline{L}^2$-invariant space containing $|\Psi_1\rangle$. We begin by noting that one application of $\underline{L}^2$ generates the entire invariant space of a given Slater determinant and hence the necessary space is $\{\underline{L}^2 |\Psi_1\rangle\}$.

If not then consider a ‘missing’ Slater determinant generated by a second application of $\underline{L}^2$, then by (3.20) this can only be generated by swapping a p-orbital pair as well as swapping the spins of two unpaired p-orbitals. However, for both of these operations to give non-zero results we must have at least one spin pair of p-electrons and one each of the other two p-orbitals with opposite spin. This then gives only the spin swap as we cannot change the paired p-orbitals for a different pair.

Again from (3.20) we see that the set of Slater determinants generated by the first and second terms in the sum are disjoint (the first only operates on spin pairs and the second on unpaired p-orbitals). For the first term we create the number of Slater determinants equal to the number of pairs present, P, multiplied by the number of ‘empty’ pairs, i.e. where neither p_i nor $\overline{p_i}$ are present. The number of such pairs is given by $3 - P - (Q_\uparrow + Q_\downarrow - 2P)$ where 3 is the number of possible pairs, P is the number of pairs already present and $(Q_\uparrow + Q_\downarrow - 2P)$ is the number of unpaired p-orbitals. Hence the first term contributes $P(3 + P - Q_\uparrow - Q_\downarrow)$ Slater determinants.

Similarly, the second term contributes a number of terms equal to the number of opposite spin pairs in different spatial orbitals, which is given by the number of unpaired spin up p-orbitals multiplied by the number of unpaired spin down p-orbitals, or $(Q_\uparrow - P)(Q_\downarrow - P)$, and hence the dimension is given by

$$D := 1 + (3 + P - Q_\uparrow - Q_\downarrow)P + (Q_\uparrow - P)(Q_\downarrow - P).$$

The range of D is once again given trivially by inspection of the possible combinations of $Q_\uparrow$, $Q_\downarrow$ and P. □

The sum of the eigenvalues of the $\underline{L}^2$ matrix is equal to the trace, which in turn is equal to FD, which depends only on P, $Q_\uparrow$ and $Q_\downarrow$. From abstract Lie Algebra theory, we know that the possible eigenvalues of $\underline{L}^2$ are 0, 2 and 6 and we wish to uniquely determine the possible eigenvalues of eigenfunctions containing a given Slater determinant using only $Q_\uparrow$, $Q_\downarrow$ and P.

Lemma 3.3.13. *Let* $|\Psi\rangle \in X_N$ *with* $\{\underline{L}^2|\Psi\rangle\} = \{|\Psi_1\rangle, \ldots, |\Psi_D\rangle\}$. *Then the* $D\times D$ *matrix with entries* $\langle\Psi_i|\underline{L}^2|\Psi_j\rangle$ *and diagonal elements all equal to* F *has possible trace and eigenvalue combinations given by*

Trace	*F*	*D*	*Evals*
0	*0*	*1*	*0*
2	*2*	*1*	*2*
8	*4*	*2*	*6,2*
12	*4*	*3*	*6,6,0*

Proof We begin with $F = 0$ which gives that $Q_\uparrow, Q_\downarrow \in \{0, 3\}$ and hence $P = \min(Q_\uparrow, Q_\downarrow)$ and so $(3 + P - Q_\uparrow - Q_\downarrow)P = 0$ and $(Q_\uparrow - P)(Q_\downarrow - P) = 0$, giving $D = 1$. Similarly, if $F = 2$ we have wlog $Q_\uparrow \in \{1, 2\}$ and $Q_\downarrow \in \{0, 3\}$, which again gives that $P = \min(Q_\uparrow, Q_\downarrow)$ and hence $D = 1$. Finally note that the possible eigenvalues of $\underline{L}^2$ are $0, 2$ and 6 and hence the case $F = 4$, $D = 1$ is not possible. □

We wish to further eliminate the case $F = 4$, $D = 2$, Eval=6, as this means that we need only eliminate odd excitations that change between traces 0 and 12, and 2 and 8.

Lemma 3.3.14. *Let* $|\Psi\rangle \in X_N^0$ *with* $\{\underline{L}^2|\Psi\rangle\} = \{|\Psi_1\rangle, \ldots, |\Psi_D\rangle\}$. *Then the* $D\times D$ *matrix with entries* $\langle\Psi_i|\underline{L}^2|\Psi_j\rangle$ *and diagonal elements all equal to* F *has possible trace and eigenvalue combinations given by*

Trace	*F*	*D*	*Evals*
0	*0*	*1*	*0*
2	*2*	*1*	*2*
8	*4*	*2*	*2*
12	*4*	*3*	*6,6,0*

Proof By Lemma 3.3.13, it suffices to show that if $F = 4$, $D = 2$, $\underline{L}^2\Psi = 6$ then $L_3\Psi \neq 0$. Suppose the eigenfunction Ψ satisfies $L_3\Psi = 0$ then by Lemma 3.3.11 we see that the number of p_1 and p_2 orbitals must be even.

Since $D = 2$, exactly one term in (3.20) must be non-zero (since the original Slater determinant is always present). For one of the first type to be non-zero we require $P(3+P-Q_\uparrow-Q_\downarrow) = 1$ which is only satisfied for $P = 1$, $\{Q_\uparrow, Q_\downarrow\} = \{1, 2\}$, which in turn gives that $(Q_\uparrow - P)(Q_\downarrow - P) = 0$ so there are no contributions of the second type.

Conversely, for $(Q_\uparrow - P)(Q_\downarrow - P) = 1$ we require $P = 0$, $Q_\uparrow = Q_\downarrow = 1$ or $P = 1$, $Q_\uparrow = Q_\downarrow = 2$, both of which lead to $P(3 + P - Q_\uparrow - Q_\downarrow) = 0$ and hence there is indeed only one extra term. In all of these cases we see that $F = Q_\uparrow(3 - Q_\uparrow) + Q_\downarrow(3 - Q_\downarrow) = 4$ and the $\underline{L}^2$ matrix is given by

$$\begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix},$$

and hence has eigenfunctions $|\Psi\rangle + |\Phi\rangle$ (with eigenvalue 2) and $|\Psi\rangle - |\Phi\rangle$ (with eigenvalue 6).

Now, with the restriction that the number of p_1 and p_2 orbitals must be even we see that the combinations of p-orbitals are

$Q_\uparrow$	$Q_\downarrow$	P	orbitals
1	1	0	$p_i\overline{p_j}$
1	2	1	$p_i\overline{p_i}\overline{p_3}$
2	1	1	$p_i\overline{p_i}p_3$
2	2	1	$p_i\overline{p_j}p_3\overline{p_3}$

where $i, j \in \{1, 2\}, i \neq j$. It is therefore clear that the non-zero term in $\underline{L}^2$ is $a^\dagger(p_j)a^\dagger(\overline{p_j})a(\overline{p_i})a(p_i)$ or $a^\dagger(\overline{p_i})a^\dagger(p_j)a(\overline{p_j})a(p_i)$.

Without loss of generality suppose

$$|\Phi\rangle := a^\dagger(p_2)a^\dagger(\overline{p_2})a(\overline{p_1})a(p_1)|\Psi\rangle \neq 0.$$

Clearly,

$$L_3|\Psi\rangle = \left[a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2) + a^\dagger(\overline{p_2})a(\overline{p_1}) - a^\dagger(\overline{p_1})a(\overline{p_2})\right]|\Psi\rangle$$

and we see that the second and fourth terms are zero, and hence

$$L_3|\Psi\rangle = \left[a^\dagger(p_2)a(p_1) + a^\dagger(\overline{p_2})a(\overline{p_1})\right]|\Psi\rangle \neq 0.$$

The fact that this is non-zero follows the fact that $|\Phi\rangle \neq 0$. Now,

$$\begin{aligned} L_3|\Phi\rangle = \big[&a^\dagger(p_2)a(p_1) - a^\dagger(p_1)a(p_2) + a^\dagger(\overline{p_2})a(\overline{p_1}) - a^\dagger(\overline{p_1})a(\overline{p_2})\big] \\ &\times a^\dagger(p_2)a^\dagger(\overline{p_2})a(\overline{p_1})a(p_1)|\Psi\rangle \end{aligned}$$

and this time it is clear that the first and third terms are zero. Simplifying the remaining two terms we have

$$L_3|\Phi\rangle = [-a^\dagger(\overline{p_2})a(p_1) - a^\dagger(p_2)a(p_1)]|\Psi\rangle = -L_3|\Psi\rangle.$$

By the same method we see that, if

$$|\Phi\rangle := a^\dagger(p_1)a^\dagger(\overline{p_2})a(\overline{p_2})a(p_1)|\Psi\rangle \neq 0$$

then $L_3|\Phi\rangle = -L_3|\Psi\rangle \neq 0$.

Recalling that the eigenfunction with $\underline{L}^2$ eigenvalue 6 had the form $|\Psi\rangle - |\Phi\rangle$ we see that $L_3(|\Psi\rangle - |\Phi\rangle) = 2L_3|\Psi\rangle \neq 0$ and the result follows. ☐

From the result of Lemma 3.3.14 we see that the space of $\underline{L}^2$ matrices on X_N^0 can be partitioned into two sets with disjoint eigenvalues, in particular we have $V_{0,6}$ and V_2 where the restrictions can be formulated as

$$\begin{aligned} V_{0,6} :&= \big\{|\Psi\rangle \in X_N^0 \mid V_{|\Psi\rangle} := \langle\Psi_i|\underline{L}^2|\Psi_j\rangle, \Psi_i \in \{\underline{L}^2|\Psi\rangle\}, \text{has } \mathrm{Tr}(V_{|\Psi\rangle}) \in \{0,12\}\big\}, \\ V_2 :&= \big\{|\Psi\rangle \in X_N^0 \mid V_{|\Psi\rangle} := \langle\Psi_i|\underline{L}^2|\Psi_j\rangle, \Psi_i \in \{\underline{L}^2|\Psi\rangle\}, \text{has } \mathrm{Tr}(V_{|\Psi\rangle}) \in \{2,8\}\big\}. \end{aligned}$$

Further, these conditions can be rewritten in terms of the eigenvalues P, $Q_\uparrow$ and $Q_\downarrow$ of the operators $\hat{P}$, $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ which commute with $\underline{L}^2$ and are therefore well-defined for all $|\Psi\rangle \in \{\underline{L}^2|\Psi_i\rangle\}$. We note that F and D are both symmetric in $Q_\uparrow$ and $Q_\downarrow$ and so may replace these values by

$$Q_+ := \max\{Q_\uparrow, Q_\downarrow\}, \quad Q_- := \min\{Q_\uparrow, Q_\downarrow\}.$$

Denoting the triple of eigenvalues by $(Q_+, Q_-, P)_{|\Psi\rangle}$, noting (by Lemma 3.3.14) that there exists a bijection between trace values and the pairs (F, D), and using the definitions of F and D in Lemma 3.3.12 we see that the sets are equivalent to

$$\begin{aligned} V_{0,6} := \big\{|\Psi\rangle \in X_N^0 \mid (Q_+, Q_-, P)_{|\Psi\rangle} \in \{&(0,0,0), (3,0,0), (3,3,3), (1,1,1), \\ &(2,1,0), (2,2,2)\}\big\} \end{aligned}$$

$$V_2 := \big\{|\Psi\rangle \in X_N^0 \mid (Q_+, Q_-, P)_{|\Psi\rangle} \in \{(1,0,0),(2,0,0),(3,1,1),(3,2,2),\\ (1,1,0),(2,1,1),(2,2,1)\}\big\},$$

where the two sets of conditions on $(Q_+, Q_-, P)_{|\Psi\rangle}$ are once again disjoint.

Corollary 3.3.15. *Let* $|\Psi\rangle, |\Phi\rangle \in V$ *where* $V \in \{V_{0,6}, V_2\}$ *with* $\langle\hat{Q}\rangle_{|\Psi\rangle} = \langle\hat{Q}\rangle_{|\Phi\rangle}$. *Then* $\langle\hat{P}\rangle_{|\Psi\rangle} = \langle\hat{P}\rangle_{|\Phi\rangle}$, *i.e. any* $\underline{L}^2$-L_3-$\hat{Q}$ *eigenfunction with* L_3 *eigenvalue zero is also an eigenfunction of* $\hat{P}$.

Proof. This is clear from the triplet forms of the eigenvalues allowed in each of $V_{0,6}$ and V_2 above. □

It remains to prove

Lemma 3.3.16. *Suppose* $|\Psi\rangle \in X_N^0$ *s.t.* $S_3|\Psi\rangle = s|\Psi\rangle$, $\hat{Q}_\uparrow|\Psi\rangle = Q_\uparrow|\Psi\rangle$, $\hat{Q}_\downarrow|\Psi\rangle = Q_\downarrow|\Psi\rangle$, $\hat{P}|\Psi\rangle = P|\Psi\rangle$, *and* $\hat{R}|\Psi\rangle = Q_2|\Psi\rangle$. *Let* $Q_+ := \max\{Q_\uparrow, Q_\downarrow\}$ *and* $Q_- := \min\{Q_\uparrow, Q_\downarrow\}$, *denote the triple of eigenvalues of* $|\Psi\rangle$ *by* $(Q_+, Q_-, P)_{|\Psi\rangle}$, *and the two sets*

$$W_{0,6} := \{(0,0,0),(3,0,0),(3,3,3),(1,1,1),(2,1,0),(2,2,2)\},$$
$$W_2 := \{(1,0,0),(2,0,0),(3,1,1),(3,2,2),(1,1,0),(2,1,1),(2,2,1)\}.$$

Then for any odd excitation $|\Psi\rangle \mapsto |\Phi\rangle$ *such that* $|\Phi\rangle \in X_N^0$ *with* $S_3|\Phi\rangle = s|\Phi\rangle$ *and* $\hat{R}|\Phi\rangle = Q_2|\Phi\rangle$, *we have either* $(Q_+, Q_-, P)_{|\Psi\rangle} \in W_{0,6}$, $(Q_+, Q_-, P)_{|\Phi\rangle} \in W_2$ *or* $(Q_+, Q_-, P)_{|\Psi\rangle} \in W_2$, $(Q_+, Q_-, P)_{|\Phi\rangle} \in W_6$.

Proof We begin by noting that the possible odd excitations change either one or three orbitals. Further, by the requirement that $|\Psi\rangle$ and $|\Phi\rangle$ may form L_3-eigenfunctions with eigenvalue zero and Lemma 3.3.11, we see that the number of p_1 and p_2 orbitals in both $|\Psi\rangle$ and $|\Phi\rangle$ must be even.

Hence, for a one-orbital change that conserves s and Q_2, there are only two possibilities: $p_1 \leftrightarrow p_2$ and $\overline{p_1} \leftrightarrow \overline{p_2}$, where by symmetry we need only consider the spin up case. It is clear that both $|\Psi\rangle$ and $|\Phi\rangle$ must contain exactly two p_1/p_2 orbitals, one spin up and the other spin down, leading to the possibilities in Table 3.16, all of which show a change between $W_{0,6}$ and W_2.

We have therefore shown that all possible single excitations of a Slater determinant $|\Psi\rangle$ must change the $\underline{L}^2$ eigenvalues of eigenfunctions which contain $|\Psi\rangle$.

We now move on to considering triple excitations. We note that if all three orbitals to be excited have the same spin then there is no possible excitation so,

$\vert\Psi\rangle$	$(Q_+, Q_-, P)_{\vert\Psi\rangle}$			$(Q_+, Q_-, P)_{\vert\Phi\rangle}$	$\vert\Phi\rangle$
$p_1\overline{p_2}$	$(1,1,0)$	W_2	$W_{0,6}$	$(1,1,1)$	$p_2\overline{p_2}$
$p_1\overline{p_2}p_3$	$(2,1,0)$	$W_{0,6}$	W_2	$(2,1,1)$	$p_2\overline{p_2}p_3$
$p_1\overline{p_2}\overline{p_3}$	$(2,1,0)$	$W_{0,6}$	W_2	$(2,1,1)$	$p_2\overline{p_2}\overline{p_3}$
$p_1\overline{p_2}p_3\overline{p_3}$	$(2,2,1)$	W_2	$W_{0,6}$	$(2,2,2)$	$p_2\overline{p_2}p_3\overline{p_3}$

Table 3.16: The effect of the one-orbital change $p_1 \leftrightarrow p_2$ taking $|\Psi\rangle \leftrightarrow |\Phi\rangle$ on the triplet of eigenvalues of the operators $\hat{Q}_\uparrow$, $\hat{Q}_\downarrow$ and $\hat{P}$. Also shown is the space in which these triplets lie.

without loss of generality, we need consider only triplets of the form $\psi_1\overline{\psi_2}\psi_3$ of which there are 24. In fact, this space is only 12 dimensional since the excitations between $s\overline{p_i}p_3 \leftrightarrow p_i\overline{p_j}p_j$ $(i \neq j \in \{1,2\})$ must be between two Slater determinants with an odd number of p_1 and p_2 orbitals, and those of the form $s\overline{p_i}p_j \leftrightarrow p_3\overline{p_j}p_i$, $(i, j \in \{1,2\})$ must change Q_2. Ignoring extra $2s$ orbitals that do not contribute to Q_+, Q_- and P we have the possibilities in Table 3.17, all of which lead to a change $W_{0,6} \leftrightarrow W_2$. This clearly shows that all possible triple excitations must change the $\underline{L}^2$ eigenvalue and this proves the lemma. □

Ψ	$(Q_+, Q_-, P)_{\vert\Psi\rangle}$		χ		$(Q_+, Q_-, P)_{\vert\Phi\rangle}$	Φ
$s\overline{s}p_3$	$(1,0,0)$ $(1,2,0)$	W_2 $W_{0,6}$	$\overline{p_1p_2}$	$W_{0,6}$ W_2	$(2,1,0)$ $(3,2,2)$	$p_i\overline{p_3}p_j$
$s\overline{s}p_i$	$(1,1,0)$ $(1,1,1)$	W_2 $W_{0,6}$	$\overline{p_j}$ $\overline{p_i}$	$W_{0,6}$ W_2	$(2,2,2)$ $(2,2,1)$	$p_3\overline{p_3}p_j$
$s\overline{p_3}p_3$	$(1,1,1)$ $(3,1,1)$	$W_{0,6}$ W_2	$\overline{p_ip_j}$	W_2 $W_{0,6}$	$(2,0,0)$ $(2,2,2)$	$p_i\overline{s}p_j$
$s\overline{p_3}p_i$	$(2,1,0)$ $(2,1,1)$	$W_{0,6}$ W_2	$\overline{p_j}$ $\overline{p_i}$	W_2 $W_{0,6}$	$(2,1,1)$ $(2,1,0)$	$p_3\overline{s}p_j$

Table 3.17: The effect of the triple excitation $\Psi \leftrightarrow \Phi$ of a Slater determinant containing the orbitals $\Psi \cup \chi \leftrightarrow \Phi \cup \chi$ on the triple of eigenvalues of $\hat{Q}_\uparrow$, $\hat{Q}_\downarrow$ and $\hat{P}$. Also shown is the space in which these triplets lie.

Proof of Theorem 3.3.1 The proof of the theorem follows directly from the result of Lemma 3.3.16 and the fact that different $\underline{L}^2$-L_3-S_3-$\hat{R}$ eigenspaces are not coupled by H (Lemma 3.1.4). In fact, Lemma 3.3.16 proves a more general result, namely that for a given $\underline{L}^2$-S_3-$\hat{R}$ eigenspace, choosing the subspace with L_3 eigenvalue zero gives a basis of evenly-excited Slater determinants. Noting that

$\underline{S}^2$ commutes with $\underline{L}^2$, L_3, S_3 and $\hat{R}$ gives the same result for the simultaneous $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspaces with L_3 eigenvalue zero, proving the theorem. □

The more general result without the $\underline{S}^2$ dependence is not immediately obvious from Tables 3.9-3.15 as, for ease of notation, we have taken different S_3 eigenfunctions in each space.

It is worth noting that the even excitation result of Theorem 3.3.1 shows the H_0 matrix on each $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenspace is diagonal, with elements given by the sum of $\langle \psi_i | H_0 | \psi_i \rangle$ over the one-electron orbitals in any Slater determinant of the spin-angular momentum eigenfunction with fixed value of Q.

One interesting question that follows from this result is whether it is true in more general cases. It is trivial (since we have not used the radial parts of the orbitals in the above) that it can be extended to any set of wavefunctions each of which have the same filled core shells and only s and p orbitals of the same principal quantum number n in the valence shell. However, it is clearly not true if we include Slater determinants with different core shells, e.g. $\Psi := |1s2s\overline{2s}\rangle$ for Lithium, which is in the same joint eigenspace as $|1s\overline{1s}2s\rangle$. However, these cases should have high energies (due to the H_0 terms) and make a small contribution to the energies of low-lying states.

The extension which is most needed is that to the next principal shell (i.e. $n = 3$), which contains d-orbitals, and hence we would need to calculate the effect of angular momentum on these orbitals and hope for similar results as for the p-shell. As a further complication, the conventional chemistry shell ordering scheme gives some cases where $4s$ is preferred to $3d$, and so any method should include this possiblity, at least in the hope of justifying these empirical rules mathematically. This will lead to terms analogous to those in the Lithium case described above and hence break the double excitation rule.

3.4 The Generality of the Preceding Discussion

It is worth taking a moment to comment on the generality of the previous results. It should be clear that all results thus far, with the obvious exception the final integral values of Section 2.5, apply to a more general basis than that discussed in Section 2.4. The minimal requirements are that we have five spatial orbitals denoted $\psi_{1s}, \psi_{2s}, \psi_{2p_1}, \psi_{2p_2}$ and ψ_{2p_3}, which are orthogonal and for which we allow a choice of two orthogonal spin states, with the additional restriction that $L_i(\psi_{1s}) = L_i(\psi_{2s}) = 0$ and the conditions on the action of the L_i on the ψ_{2p_j} given by (3.6).

In particular, these restrictions are thus independent of choice of the radial parts (assuming orthogonality is maintained and the radial part is the same for each of the p-orbitals). This will be important in later sections where we wish to choose more flexible basis sets with the same general structure but will still be able to use important results such as the eigenspace tables in Sections 3.2.2 and 3.2.3.

However, a large part of the motivation for using this method is the (relative) simplicity of the calculations of Sections 2.5.1 and 2.5.2. If, for example, the resulting integrals had turned out to give non-analytic functions, as compared to surprisingly simple rational functions, then the method would not be anywhere near as attractive. In such a case we would need to use numerical approximations for the integrals, introducing computational errors unconnected to the choice of orbitals. Hence, the specific choice of orbitals in Section 2.4 helps to justify our method and the real test will follow in the next section, where we compute the energies of our wavefunctions.

Chapter 4

Atomic Energy Levels

4.1 Atomic Spectra

In this section we make use of the calculated eigenspace tables of Section 3.2.3 and the explicit integrals given in Table 2.3 to calculate the energies of these eigenfunctions. Along with obtaining the ground state energies (without having to specify the angular momentum and spin eigenstates through experimental data), this will allow us to investigate the spectra of the second row atoms. We will then compare these values to experimental data, which is undeniably the best test of our theory.

4.1.1 Calculating the Energies

We calculate the energies using Slater's rules from Corollary 2.3.2 and the values of the integrals from Table 2.3. It is worth recalling that the wavefunctions in Tables 3.9-3.15 are not normalized, but this is trivial to correct. Unless otherwise stated, all experimental data in this section are taken from the NIST atomic database [NIS06] (for excited states) or [Huh93] (for ground state energies, obtained by summing the ionization energies).

We begin with the example of Lithium. As can be seen from Table 3.9, there is only one wavefunction in each eigenspace. This makes the calculation easier as we do not need to calculate inter-wavefunction terms nor perform any diagonalization. The two expected values are given by

$$
\begin{aligned}
\langle V_{ee}\rangle_{|1\bar{1}2\rangle} &= (11|11) + 2(11|22) - (12|21) = \frac{5965}{5832}Z \\
\langle V_{ee}\rangle_{|1\bar{1}3\rangle} &= (11|11) + 2(11|33) - (13|31) = \frac{57397}{52488}Z.
\end{aligned}
$$

As noted previously, the contribution to the energy from H_0 is the same for both wavefunctions. Hence, in order to determine the ordering of the spectrum, it remains to determine which of the two expected value is the lowest. We have that

$$\frac{5965}{5832}Z - \frac{57397}{52488}Z = -\frac{464}{6561}Z,$$

and hence the ground state has $\langle \underline{L}^2 \rangle = 0$ and $\langle \underline{S}^2 \rangle = \frac{3}{4}$, which agrees with experimental data. By Theorem 2.2.2 we have the total energy to first order:

$$E_{|\phi_1 \ldots \phi_N\rangle} = Z^2 \sum_{i=1}^{N} E_i + \langle V_{ee} \rangle_{|\phi_1 \ldots \phi_N\rangle}$$

where E_i are the non-interacting energies for the corresponding ϕ_i.

In fact, from our choice of basis it is easy to see that, for $2 \leq N \leq 10$,

$$Z^2 \sum_{i=1}^{N} E_i = Z^2 \left(-1 - \frac{(n-2)}{8} \right).$$

The ground state energy is therefore given by

$$E_0^{Li} = 3^2 \left(\frac{-9}{8} \right) + 3\frac{5965}{5832} = -\frac{6859}{972} \sim -7.0566.$$

Similarly, the energy of the first excited state is given by

$$E_1^{Li} = 3^2 \left(\frac{-9}{8} \right) + 3\frac{57397}{52488} = -\frac{59875}{8748} \sim -6.8444.$$

Comparing our ground state energy with the experimental value of -7.4779, we find that we capture 94.3% of the experimental value.

The spectral gap can also be compared to experimental data and we find that the $\underline{L}^2$-$\underline{S}^2$ ordering is correct. Whilst we have shown that the $\langle \underline{L}^2 \rangle = 2$ case has higher energy than the ground state, it is not trivial that there are no other states between the two states we have investigated. For example, *a priori* the $\langle \underline{L}^2 \rangle = 0$ case formed by swapping the $2s$ orbital for a $3s$ orbital could have been the next experimental spectral line.

The experimental spectral gap is 0.0679 whilst our spectral gap is 0.2122, which is more than three times the experimental value. It seems a little odd that the ground state energy is so close to the experimental value but the spectral gap is so far off, especially since taking differences normally reduces systematic errors. We will address this idea later, but for now we continue with our calculation of the ground and excited state energies for the remaining atoms.

The $\langle V_{ee} \rangle$ matrix element expressions for Beryllium-Neon are given in Tables 4.3 - 4.5. It is clear that the highest dimension of matrix we need to diagonalize is two, and for a general real, symmetric, 2×2 matrix

$$\begin{pmatrix} \langle V_{ee} \rangle_{\Psi_1} & \langle \Psi_1 | V_{ee} | \Psi_2 \rangle \\ \langle \Psi_1 | V_{ee} | \Psi_2 \rangle & \langle V_{ee} \rangle_{\Psi_2} \end{pmatrix} := \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad b \neq 0,$$

the eigenvalues and eigenvectors are given by $\lambda_\pm = \frac{a+c}{2} \pm \sqrt{\left(\frac{a-c}{2}\right)^2 + b^2}$ and $\Psi_\pm = \Psi_1 + (\frac{\lambda_\pm - a}{b})\Psi_2$, which are not normalized.

The resulting eigenfunctions and energies are given in Tables 4.6 and 4.7. We note that all the energies in the the cases where no matrix diagonalization is required are rational multiples of Z; we also give an approximate decimal, allowing easier comparison of the energies. Those corresponding to diagonalized terms are given purely as a decimal as the analytic forms are too complicated to be useful in comparing energies.

The non-correlated ground state energies are not a new result, this is a consequence of the fact that these wavefunctions can be represented by a single Slater determinant, and are thus in fact Hartree-Fock wavefunctions. In particular, we see that the values for the ground state match with those given for the first order perturbation in Table 2 of [Wil71]. Although given as decimals in that table, the references for the integrals [Sha62, SW68a, SW68b, SW66] are in rational form and are the same as those given in Table 2.3. However, the papers of Sharma and Wilson do not explicitly state how the values of these integrals were arrived at, and it is unclear how it extends to other basis sets. This contrasts with our method of calculation and, as will be demonstrated later, the ability to calculate the integrals for more complicated choices of basis is very important.

For the excited states, since by Theorem 2.2.2 the V_{ee} term is the order Z perturbation in the energy expansion, our results agree with the numerical results of [IS75] after combining their order Z corrections from Tables 2 and 4. However, our results are more accurate due to them being given explicitly as rational (or square roots of rational) values. The same numerical results are given in [LS60], and although given as decimals, the rational integrals necessary to compute these are given in the same paper. The same paper contains the integral expansions of Tables 4.3-4.5, although using purely radial integrals, although it is unclear how these were derived without forming the eigenfunctions. However, their aim was to investigate the correlation energy rather and do not give explicit wavefunctions nor compare their results to experiment, which we will now do.

4.1.2 Comparison with Experiment

We now move on to the most important test of our results - how well they agree with the experimental data. Our energies are shown along with the experimental values in Tables 4.8 and 4.9. These data are all taken from the online NIST atomic database [NIS06], with the exception of the Carbon $2s2p^3$ ^{1}D state, which is taken from [Moo70]. The state (or term) symbols are standard chemistry notation for the spin and angular momentum eigenspaces. The letter denotes the angular momentum eigenvalue, $S = 0$, $P = 2$, $D = 6$, whilst the number denotes the spin multiplicity, 1 corresponding to a spin eigenvalue of 0, 2 to 3/4, 3 to 1 and so on.

The NIST data have been adjusted to account for total angular momentum of non-singlet states. In the full Schrödinger equation there are terms that couple the spatial and spin parts of the wavefunction (so called spin-orbit coupling), leading to a range of different total angular momenta for each $\underline{L}^2$-$\underline{S}^2$ eigenspace, e.g [Sch01]. Clearly these terms are not present in our approximation and so the idea of the spin and spatial angular momentum coupling in this way does not apply. For this reason we take the mean of the different values for each $\underline{L}^2$-$\underline{S}^2$ eigenspace on NIST, although the effect on significant digits is small.

From Tables 4.8 and 4.9 we see that, as with the case of Lithium, there is good agreement between our calculated total ground state energy and that of experiment, the errors in which are shown in Table 4.1. It is clear that the relative error increases as the nuclear charge increases, one obvious reason for this would be if the basis set was a poorer approximation of the real wavefunction for higher nuclear charge.

Atom	Li	Be	B	C	N	O	F	Ne
Error	5.6%	6.2%	7.8%	9.0%	10.0%	11.2%	12.2%	13.0%

Table 4.1: Percentage error in the ground state energy.

This suggests that a better one-electron basis should have a less trivial dependence on Z than the eigenfunctions of H_0. For our method, this change is preferable to increasing the size of the basis set, mainly due to the rapid increase in the dimension of the basis set of Slater determinants as n is increased. For example, including the $n = 3$ orbitals in the case of Carbon, there would be $\binom{26}{4} = 14950$ (18 orbitals from the 3 shell added to the current 8, choose 4) possible Slater determinants which is obviously much larger than the current case of 70. This will be investigated in the next section.

Another measure of how well the theoretical ground states approximate the physical ground states uses the virial theorem. For a Coulomb potential, this says that $-\langle V\rangle_\Psi/\langle K\rangle_\Psi = 2$ where K is the kinetic energy and V the potential energy, which is a necessary but not sufficient constraint for the wavefunction to be a true minimizer [Sch01]. These values are shown in Table 4.2 and are not particularly close to two, which indicates that the wavefunctions are relatively poor approximations to the true ground states. As before, the agreement becomes poorer as we move across the period, this is not really surprising as we would expect that poorer energies (see Table 4.1) come from poorer wavefunctions.

Atom	Li	Be	B	C	N	O	F	Ne
Ratio	1.6969	1.6881	1.6615	1.6379	1.6173	1.5956	1.5778	1.5615

Table 4.2: The ratio $-\langle V\rangle_\Psi/\langle K\rangle_\Psi$ for the ground state wavefunctions of Li-Ne. A physically accurate wavefunction should have a value of 2 due to the virial theorem.

Also similar to the Lithium case is that the spectral gaps for the other atoms are in fairly poor agreement with experiment, being around $150-200\%$ out from the experimental values. This suggests that there are errors present in the excited state wavefunctions which are not present in that of the ground state, some possible explanations of this will be discussed in the next section.

One thing of note is that, despite the poor numerical comparisons of the spectral gaps, the ordering of the levels is excellent. If we ignore the experimental states attributed to wavefunctions containing orbitals with $n > 2$ (which unlike in the Lithium case lie in the range of our spectrum), the results differ only by the interchange of two levels (^1D and ^3P) of Beryllium (which in fact are in violation of Hund's rules - see below). This is quite a remarkable result and is one reason to believe that, despite its simplicity, our method captures important qualities of the wavefunctions.

Theorem 2.2.2 suggests that, to first order and for small λ, the ground state energy should be well approximated by our results. However, the physical value of $\lambda = 1$ is not 'small' and thus the agreement with the physical ordering is by no means guaranteed by this result. We will later see that the spectral ordering is, to some extent, pre-determined by the properties of the one-electron basis. This is analogous to a rigorous version of Hund's rules for the ordering of the spectrum and is discussed in Section 4.3.

Hund's rules state that for the lowest energy state, the spins of the electrons should be aligned as far as possible. This is a rigorous mathematical result for single Slater determinants and can be seen by noting that the exchange energy is lowest when $\sum_{i\neq j}|s_i\cdot s_j|$ is maximized (see [Fri0X]).

Further, for states of the same total spin, Hund's rules state that the highest angular momentum state has the lowest energy. It is not possible to form multiple angular momentum states with a single Slater determinant so there is no analogous mathematical proof of this statement.

The other commonly used rule is a 'shell ordering' rule that states that the one-electron orbitals should be filled in the order $1s$, $2s$, $2p$, $3s$, $3p$, and so on. We note that there are a number of exceptions to this simple ordering but they do not affect the $n=2$ shell. Again, there does not seem to be a rigorous mathematical proof that this should be the case.

We note that Hund's rules are designed to be applied only to the ground state of an atom. However, they are often used to predict the ordering of the spectra of atoms. Whilst this is a misapplication, it often leads to accurate predictions.

4.1.3 Spectral Gap Errors

One possible source of errors in the energies of excited states is the restriction that they be orthogonal to the calculated ground state. This means that any errors in the calculated ground state wavefunction cause incorrect restrictions on the excited states. In other words, a poor choice of ground state wavefunction, even if it leads to a good energy prediction, will almost inevitably lead to poor energies for the excited states.

Another related idea concerns degenerate eigenspaces, i.e. ones where there are multiple orthogonal wavefunctions with the same $\underline{L}^2$-$\underline{S}^2$-L_3-S_3-$\hat{R}$ eigenvalues. In these cases correlation effects can be very important in determining the ordering of the energies. From the form of $\lambda_\pm$ above, we see that when two wavefunctions are correlated, the lower energy is shifted down and the higher energy shifted up.

Consider a simultaneous eigenspace in which there is a wavefunction containing orbitals with $n\geq 3$ of lower energy than a wavefunction we have included. Then the higher energy levels are not rigorous upper bounds to the true energy levels, since correlation with the missing wavefunction would cause the energy to be shifted upwards. The opposite is true of missing correlated wavefunctions that have higher energy, these will cause the calculated energy to be too high.

For example, there are a large number of experimental ${}^1S$ states of Beryllium between the ground state and the state attributed to $\Psi_{2p}:=|1s\overline{1s}2p_1\overline{2p_1}\rangle+|1s\overline{1s}2p_2\overline{2p_2}\rangle+|1s\overline{1s}2p_3\overline{2p_3}\rangle$ (from $2s3s$ up to at least $2s11s$ [NIS06]), and hence the energy level could be missing a reasonable amount of (positive) correlation energy. In contrast, it should also be missing negative correlation energy from higher ^1S states, but the experimental values for these states are not available and hence it is difficult to make a quantitative estimate of the relative effects.

It is clear from Tables 4.8 and 4.9 that this effect is not large, since all of our results are upper bounds to the true energies. However, it is worth remembering this possible correlation effect when using basis sets and methods that do not include all correlations within an eigenspace.

	$\langle \underline{L}^2 \rangle$	$\langle \underline{S}^2 \rangle$	$\langle \hat{R} \rangle$	Ψ	$\langle V_{ee} \rangle$
Li	0	$\frac{3}{4}$	1	$\vert 1\bar{1}2\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21)$
	2	$\frac{3}{4}$	-1	$\vert 1\bar{1}3\rangle$	$(11\vert 11) + 2(11\vert 33) - (13\vert 31)$
Be	0	0	1	$\vert 1\bar{1}2\bar{2}\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + (22\vert 22)$
				$\frac{1}{\sqrt{3}}\left(\vert 1\bar{1}3\bar{3}\rangle + \vert 1\bar{1}4\bar{4}\rangle + \vert 1\bar{1}5\bar{5}\rangle\right)$	$(11\vert 11) + 4(11\vert 33) - 2(13\vert 31) + (33\vert 33) + 2(34\vert 43)$
				cross	$\sqrt{3}(23\vert 32)$
	2	0	-1	$\frac{1}{\sqrt{2}}\left(\vert 1\bar{1}2\bar{3}\rangle - \vert 1\bar{1}\bar{2}3\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 2(11\vert 33) - (13\vert 31)$ $+(22\vert 33) + (23\vert 32)$
	2	2	-1	$\vert 1\bar{1}23\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 2(11\vert 33) - (13\vert 31)$ $+(22\vert 33) - (23\vert 32)$
			1	$\vert 1\bar{1}45\rangle$	$(11\vert 11) + 4(11\vert 33) - 2(13\vert 31) + (33\vert 44) - (34\vert 43)$
	6	0	1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}3\bar{3}\rangle - \vert 1\bar{1}4\bar{4}\rangle - \vert 1\bar{1}5\bar{5}\rangle\right)$	$(11\vert 11) + 4(11\vert 33) - 2(13\vert 31) + (33\vert 33) - (34\vert 43)$
B	0	$\frac{3}{4}$	1	$\frac{1}{\sqrt{3}}\left(\vert 1\bar{1}23\bar{3}\rangle + \vert 1\bar{1}24\bar{4}\rangle + \vert 1\bar{1}25\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 4(11\vert 33) - 2(13\vert 31)$ $+2(22\vert 33) - (23\vert 32) + (33\vert 33) + 2(34\vert 43)$
	0	$\frac{15}{4}$	-1	$\vert 1\bar{1}345\rangle$	$(11\vert 11) + 6(11\vert 33) - 3(13\vert 31) + 3(33\vert 44) - 3(34\vert 43)$
	2	$\frac{3}{4}$	-1	$\vert 1\bar{1}2\bar{2}3\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 2(11\vert 33) - (13\vert 31)$ $+(22\vert 22) + 2(22\vert 33) - (23\vert 32)$
				$\frac{1}{\sqrt{2}}\left(\vert 1\bar{1}34\bar{4}\rangle + \vert 1\bar{1}35\bar{5}\rangle\right)$	$(11\vert 11) + 6(11\vert 33) - 3(13\vert 31) + (33\vert 33) + 2(33\vert 44)$
				cross	$\sqrt{2}(23\vert 32)$
			1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}\bar{2}45\rangle - \vert 1\bar{1}2\bar{4}5\rangle - \vert 1\bar{1}24\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 4(11\vert 33) - 2(13\vert 31)$ $+2(22\vert 33) + (23\vert 32) + (33\vert 44) - (34\vert 43)$
	2	$\frac{15}{4}$	1	$\vert 1\bar{1}245\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 4(11\vert 33) - 2(13\vert 31)$ $+2(22\vert 33) - 2(23\vert 32) + (33\vert 44) - (34\vert 43)$
	6	$\frac{3}{4}$	1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}23\bar{3}\rangle - \vert 1\bar{1}24\bar{4}\rangle - \vert 1\bar{1}25\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 4(11\vert 33) - 2(13\vert 31)$ $+2(22\vert 33) - (23\vert 32) + (33\vert 33) - (34\vert 43)$
			-1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}\bar{3}45\rangle - \vert 1\bar{1}3\bar{4}5\rangle - \vert 1\bar{1}34\bar{5}\rangle\right)$	$(11\vert 11) + 6(11\vert 33) - 3(13\vert 31) + 3(33\vert 44)$

Table 4.3: V_{ee} matrix element expressions for Li-B, eigenfunctions from Tables 3.9-3.11 with orbital notation from (2.7), integral notation as in (2.4). 'cross' denotes the diagonal term in the 2×2 matrix.

	$\langle \underline{L}^2 \rangle$	$\langle \underline{S}^2 \rangle$	$\langle \hat{R} \rangle$	Ψ	$\langle V_{ee} \rangle$
C	0	0	1	$\frac{1}{\sqrt{3}}\left(\vert 1\bar{1}2\bar{2}3\bar{3}\rangle + \vert 1\bar{1}2\bar{2}4\bar{4}\rangle + \vert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 4(11\vert 33) - 2(13\vert 31) + (22\vert 22) + 4(22\vert 33) - 2(23\vert 32) + (33\vert 33) + 2(34\vert 43)$
				$\frac{1}{\sqrt{3}}\left(\vert 1\bar{1}3\bar{3}4\bar{4}\rangle + \vert 1\bar{1}3\bar{3}5\bar{5}\rangle + \vert 1\bar{1}4\bar{4}5\bar{5}\rangle\right)$	$(11\vert 11) + 8(11\vert 33) - 4(13\vert 31) + 2(33\vert 33) + 4(33\vert 44)$
				cross	$2(23\vert 32)$
	0	2	-1	$\frac{1}{\sqrt{12}}\left(3\vert 1\bar{1}\bar{2}345\rangle - \vert 1\bar{1}2\bar{3}45\rangle - \vert 1\bar{1}23\bar{4}5\rangle - \vert 1\bar{1}234\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) + (23\vert 32) + 3(33\vert 44) - 3(34\vert 43)$
	0	6	-1	$\vert 1\bar{1}2345\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) - 3(23\vert 32) + 3(33\vert 44) - 3(34\vert 43)$
	2	0	-1	$\frac{1}{2}\left(\vert 1\bar{1}2\bar{3}4\bar{4}\rangle - \vert 1\bar{1}\bar{2}34\bar{4}\rangle + \vert 1\bar{1}2\bar{3}5\bar{5}\rangle - \vert 1\bar{1}\bar{2}35\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) + (33\vert 33) + 2(33\vert 44)$
	2	2	1	$\vert 1\bar{1}2\bar{2}45\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 4(11\vert 33) - 2(13\vert 31) + (22\vert 22) + 4(22\vert 33) - 2(23\vert 32) + (33\vert 44) - (34\vert 43)$
				$\vert 1\bar{1}3\bar{3}45\rangle$	$(11\vert 11) + 8(11\vert 33) - 4(13\vert 31) + (33\vert 33) + 5(33\vert 44) - 3(34\vert 43)$
				cross	$(23\vert 32)$
			-1	$\frac{1}{\sqrt{2}}\left(\vert 1\bar{1}234\bar{4}\rangle + \vert 1\bar{1}235\bar{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) - 2(23\vert 32) + (33\vert 33) + 2(33\vert 44)$
	6	0	1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}2\bar{2}3\bar{3}\rangle - \vert 1\bar{1}2\bar{2}4\bar{4}\rangle - \vert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 4(11\vert 33) - 2(13\vert 31) + (22\vert 22) + 4(22\vert 33) - 2(23\vert 32) + (33\vert 33) - (34\vert 43)$
				$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}4\bar{4}5\bar{5}\rangle - \vert 1\bar{1}3\bar{3}4\bar{4}\rangle - \vert 1\bar{1}3\bar{3}5\bar{5}\rangle\right)$	$(11\vert 11) + 8(11\vert 33) - 4(13\vert 31) + 2(33\vert 33) + 4(33\vert 44) - 3(34\vert 43)$
				cross	$-(23\vert 32)$
			-1	$\frac{1}{\sqrt{12}}\left(2\vert 1\bar{1}23\bar{4}\bar{5}\rangle - \vert 1\bar{1}2\bar{3}4\bar{5}\rangle - \vert 1\bar{1}2\bar{3}\bar{4}5\rangle + 2\vert 1\bar{1}\bar{2}\bar{3}45\rangle - \vert 1\bar{1}\bar{2}34\bar{5}\rangle - \vert 1\bar{1}\bar{2}3\bar{4}5\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) + 3(33\vert 44)$
	6	2	-1	$\frac{1}{\sqrt{6}}\left(2\vert 1\bar{1}2\bar{3}45\rangle - \vert 1\bar{1}234\bar{5}\rangle - \vert 1\bar{1}23\bar{4}5\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 6(11\vert 33) - 3(13\vert 31) + 3(22\vert 33) - 2(23\vert 32) + 3(33\vert 44)$

Table 4.4: V_{ee} matrix element expressions for C, eigenfunctions from Table 3.12 with orbital notation from (2.7), integral notation as in (2.4). ‘cross’ denotes the diagonal term in the 2 × 2 matrix.

	$\langle \underline{L}^2 \rangle$	$\langle \underline{S}^2 \rangle$	$\langle \hat{R} \rangle$	Ψ	$\langle V_{ee} \rangle$
N	0	$\frac{3}{4}$	1	$\frac{1}{\sqrt{3}}\big(\vert 1\overline{1}23\overline{3}4\overline{4}\rangle + \vert 1\overline{1}23\overline{3}5\overline{5}\rangle$ $+\vert 1\overline{1}24\overline{4}5\overline{5}\rangle\big)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+4(22\vert 33) - 2(23\vert 32) + 2(33\vert 33) + 4(33\vert 44)$
	0	$\frac{15}{4}$	-1	$\vert 1\overline{1}2\overline{2}345\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 6(11\vert 33) - 3(13\vert 31)$ $+(22\vert 22) + 6(22\vert 33) - 3(23\vert 32) + 3(33\vert 44) - 3(34\vert 43)$
	2	$\frac{3}{4}$	-1	$\vert 1\overline{1}34\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 10(11\vert 33) - 5(13\vert 31) + 2(33\vert 33) + 8(33\vert 44)$ $-4(34\vert 43)$
				$\frac{1}{\sqrt{2}}\left(\vert 1\overline{1}2\overline{2}34\overline{4}\rangle + \vert 1\overline{1}2\overline{2}35\overline{5}\rangle\right)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 6(11\vert 33) - 3(13\vert 31)$ $+(22\vert 22) + 6(22\vert 33) - 3(23\vert 32) + (33\vert 33) + 2(33\vert 44)$
				cross	$\sqrt{2}(23\vert 32)$
			1	$\frac{1}{\sqrt{6}}\big(2\vert 1\overline{1}\overline{2}3\overline{3}45\rangle - \vert 1\overline{1}23\overline{3}\overline{4}5\rangle$ $-\vert 1\overline{1}23\overline{3}4\overline{5}\rangle\big)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+4(22\vert 33) + (33\vert 33) + 5(33\vert 44) - 3(34\vert 43)$
	2	$\frac{15}{4}$	1	$\vert 1\overline{1}23\overline{3}45\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+4(22\vert 33) - 3(23\vert 32) + (33\vert 33) + 5(33\vert 44) - 3(34\vert 43)$
	6	$\frac{3}{4}$	-1	$\frac{1}{\sqrt{6}}\big(2\vert 1\overline{1}2\overline{2}\overline{3}45\rangle - \vert 1\overline{1}2\overline{2}34\overline{5}\rangle$ $-\vert 1\overline{1}2\overline{2}3\overline{4}5\rangle\big)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 6(11\vert 33) - 3(13\vert 31)$ $+(22\vert 22) + 6(22\vert 33) - 3(23\vert 32) + 3(33\vert 44)$
			1	$\frac{1}{\sqrt{6}}\big(2\vert 1\overline{1}24\overline{4}5\overline{5}\rangle - \vert 1\overline{1}23\overline{3}4\overline{4}\rangle$ $-\vert 1\overline{1}23\overline{3}5\overline{5}\rangle\big)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+4(22\vert 33) - 2(23\vert 32) + 2(33\vert 33) + 4(33\vert 44) - 3(34\vert 43)$
O	0	0	1	$\frac{1}{\sqrt{3}}\big(\vert 1\overline{1}2\overline{2}3\overline{3}4\overline{4}\rangle + \vert 1\overline{1}2\overline{2}3\overline{3}5\overline{5}\rangle$ $+\vert 1\overline{1}2\overline{2}4\overline{4}5\overline{5}\rangle\big)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+(22\vert 22) + 8(22\vert 33) - 4(23\vert 32) + 2(33\vert 33) + 4(33\vert 44)$
				$\vert 1\overline{1}3\overline{3}4\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 12(11\vert 33) - 6(13\vert 31) + 3(33\vert 33) + 12(33\vert 44)$ $-6(34\vert 43)$
				cross	$\sqrt{3}(23\vert 32)$
	2	0	-1	$\frac{1}{\sqrt{2}}\left(\vert 1\overline{1}2\overline{3}4\overline{4}5\overline{5}\rangle - \vert 1\overline{1}\overline{2}34\overline{4}5\overline{5}\rangle\right)$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 10(11\vert 33) - 5(13\vert 31)$ $+5(22\vert 33) - (23\vert 32) + 2(33\vert 33) + 8(33\vert 44) - 4(34\vert 43)$
	2	2	-1	$\vert 1\overline{1}234\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 10(11\vert 33) - 5(13\vert 31)$ $+5(22\vert 33) - 3(23\vert 32) + 2(33\vert 33) + 8(33\vert 44) - 4(34\vert 43)$
			1	$\vert 1\overline{1}2\overline{2}3\overline{3}45\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+(22\vert 22) + 8(22\vert 33) - 4(23\vert 32) + (33\vert 33) + 5(33\vert 44)$ $-3(34\vert 43)$
	6	0	1	$\frac{1}{\sqrt{6}}\big(2\vert 1\overline{1}2\overline{2}4\overline{4}5\overline{5}\rangle - \vert 1\overline{1}2\overline{2}3\overline{3}4\overline{4}\rangle$ $-\vert 1\overline{1}2\overline{2}3\overline{3}5\overline{5}\rangle\big)$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 8(11\vert 33) - 4(13\vert 31)$ $+(22\vert 22) + 8(22\vert 33) - 4(23\vert 32) + 2(33\vert 33) + 4(33\vert 44)$ $-3(34\vert 43)$
F	0	$\frac{3}{4}$	1	$\vert 1\overline{1}23\overline{3}4\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 2(11\vert 22) - (12\vert 21) + 12(11\vert 33) - 6(13\vert 31)$ $+6(22\vert 33) - 3(23\vert 32) + 3(33\vert 33) + 12(33\vert 44) - 6(34\vert 43)$
	2	$\frac{3}{4}$	-1	$\vert 1\overline{1}2\overline{2}34\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 10(11\vert 33) - 5(13\vert 31)$ $+(22\vert 22) + 10(22\vert 33) - 5(23\vert 32) + 2(33\vert 33) + 8(33\vert 44)$ $-4(34\vert 43)$
Ne	0	0	1	$\vert 1\overline{1}2\overline{2}3\overline{3}4\overline{4}5\overline{5}\rangle$	$(11\vert 11) + 4(11\vert 22) - 2(12\vert 21) + 12(11\vert 33) - 6(13\vert 31)$ $+(22\vert 22) + 12(22\vert 33) - 6(23\vert 32) + 3(33\vert 33) + 12(33\vert 44)$ $-6(34\vert 43)$

Table 4.5: V_{ee} matrix element expressions for N-Ne, eigenfunctions from Tables 3.13-3.15 with orbital notation from (2.7), integral notation as in (2.4). ‘cross’ denotes the diagonal term in the 2×2 matrix.

	$\langle \underline{L}^2\rangle$	$\langle \underline{S}^2\rangle$	$\langle \hat{R}\rangle$	Ψ	$\langle V_{ee}\rangle$	
Li	0	$\frac{3}{4}$	1	$\lvert 1\bar{1}2\rangle$	$\frac{5965}{5832}Z$	1.0228Z
	2	$\frac{3}{4}$	-1	$\lvert 1\bar{1}3\rangle$	$\frac{57397}{52488}Z$	1.0935Z
Be	0	0	1	$0.97432\lvert 1\bar{1}2\bar{2}\rangle - 0.22517(\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}3\bar{3}\rangle + \lvert 1\bar{1}4\bar{4}\rangle + \lvert 1\bar{1}5\bar{5}\rangle\right)$		$1.5593Z$
				$0.22517\lvert 1\bar{1}2\bar{2}\rangle + 0.97432(\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}3\bar{3}\rangle + \lvert 1\bar{1}4\bar{4}\rangle + \lvert 1\bar{1}5\bar{5}\rangle\right)$		$1.7906Z$
	2	0	-1	$\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}2\bar{3}\rangle - \lvert 1\bar{1}\bar{2}3\rangle\right)$	$\frac{2826353}{1679616}Z$	$1.6827Z$
	2	2	-1	$\lvert 1\bar{1}23\rangle$	$\frac{1363969}{839808}Z$	$1.6241Z$
			1	$\lvert 1\bar{1}45\rangle$	$\frac{1449605}{839808}Z$	$1.7261Z$
	6	0	1	$\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}3\bar{3}\rangle - \lvert 1\bar{1}4\bar{4}\rangle - \lvert 1\bar{1}5\bar{5}\rangle\right)$	$\frac{14673197}{8398080}Z$	$1.7472Z$
B	0	$\frac{3}{4}$	1	$\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}23\bar{3}\rangle + \lvert 1\bar{1}24\bar{4}\rangle + \lvert 1\bar{1}25\bar{5}\rangle\right)$	$\frac{4151299}{1679616}Z$	$2.4716Z$
	0	$\frac{15}{4}$	-1	$\lvert 1\bar{1}345\rangle$	$\frac{706213}{279936}Z$	$2.5228Z$
	2	$\frac{3}{4}$	-1	$0.98633\lvert 1\bar{1}2\bar{2}3\rangle - 0.16480\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}34\bar{4}\rangle + \lvert 1\bar{1}35\bar{5}\rangle\right)$		$2.3275Z$
				$0.16480\lvert 1\bar{1}2\bar{2}3\rangle + 0.98633\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}34\bar{4}\rangle + \lvert 1\bar{1}35\bar{5}\rangle\right)$		$2.5824Z$
			1	$\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}\bar{2}45\rangle - \lvert 1\bar{1}2\bar{4}5\rangle - \lvert 1\bar{1}24\bar{5}\rangle\right)$	$\frac{8322281}{3359232}Z$	$2.4474Z$
	2	$\frac{15}{4}$	1	$\lvert 1\bar{1}245\rangle$	$\frac{2006759}{839808}Z$	$2.3895Z$
	6	$\frac{3}{4}$	1	$\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}\bar{2}3\bar{3}\rangle - \lvert 1\bar{1}24\bar{4}\rangle - \lvert 1\bar{1}5\bar{5}\rangle\right)$	$\frac{40981549}{16796160}Z$	$2.4399Z$
			-1	$\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}\bar{3}45\rangle - \lvert 1\bar{1}3\bar{4}5\rangle - \lvert 1\bar{1}34\bar{5}\rangle\right)$	$\frac{14301407}{5598720}Z$	$2.5544Z$
C	0	0	1	$0.9797\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}2\bar{2}3\bar{3}\rangle + \lvert 1\bar{1}2\bar{2}4\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$ $-0.2005\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}3\bar{3}4\bar{4}\rangle + \lvert 1\bar{1}3\bar{3}5\bar{5}\rangle + \lvert 1\bar{1}4\bar{2}4\bar{5}\rangle\right)$		$3.3027Z$
				$0.2005\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}2\bar{2}3\bar{3}\rangle + \lvert 1\bar{1}2\bar{2}4\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$ $+0.9797\frac{1}{\sqrt{3}}\left(\lvert 1\bar{1}3\bar{3}4\bar{4}\rangle + \lvert 1\bar{1}3\bar{3}5\bar{5}\rangle + \lvert 1\bar{1}4\bar{2}4\bar{5}\rangle\right)$		$3.6009Z$
	0	2	-1	$\frac{1}{\sqrt{12}}\left(3\lvert 1\bar{1}\bar{2}345\rangle - \lvert 1\bar{1}2\bar{3}45\rangle - \lvert 1\bar{1}23\bar{4}5\rangle - \lvert 1\bar{1}234\bar{5}\rangle\right)$	$\frac{961915}{279936}Z$	$3.4362Z$
	0	6	-1	$\lvert 1\bar{1}2345\rangle$	$\frac{464555}{139968}Z$	$3.3190Z$
	2	0	-1	$\frac{1}{2}\left(\lvert 1\bar{1}2\bar{3}4\bar{4}\rangle - \lvert 1\bar{1}\bar{2}34\bar{4}\rangle + \lvert 1\bar{1}2\bar{3}5\bar{5}\rangle - \lvert 1\bar{1}\bar{2}35\bar{5}\rangle\right)$	$\frac{242119}{69984}Z$	$3.4596Z$
	2	2	1	$0.9945\lvert 1\bar{1}2\bar{2}45\rangle - 0.1050\lvert 1\bar{1}3\bar{3}45\rangle$		$3.2589Z$
				$0.1050\lvert 1\bar{1}2\bar{2}45\rangle + 0.9945\lvert 1\bar{1}3\bar{3}45\rangle$		$3.5393Z$
			-1	$\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}234\bar{4}\rangle + \lvert 1\bar{1}235\bar{5}\rangle\right)$	$\frac{1904147}{559872}Z$	$3.4010Z$
	6	0	1	$0.9838\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}2\bar{2}3\bar{3}\rangle - \lvert 1\bar{1}2\bar{2}4\bar{4}\rangle - \lvert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$ $-0.1792\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}4\bar{4}5\bar{5}\rangle - \lvert 1\bar{1}3\bar{3}4\bar{4}\rangle - \lvert 1\bar{1}3\bar{3}5\bar{5}\rangle\right)$		$3.2777Z$
				$0.1792\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}2\bar{2}3\bar{3}\rangle - \lvert 1\bar{1}2\bar{2}4\bar{4}\rangle - \lvert 1\bar{1}2\bar{2}5\bar{5}\rangle\right)$ $+0.9838\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}4\bar{4}5\bar{5}\rangle - \lvert 1\bar{1}3\bar{3}4\bar{4}\rangle - \lvert 1\bar{1}3\bar{3}5\bar{5}\rangle\right)$		$3.4439Z$
			-1	$\frac{1}{\sqrt{12}}\big(2\lvert 1\bar{1}23\bar{4}\bar{5}\rangle - \lvert 1\bar{1}2\bar{3}4\bar{5}\rangle - \lvert 1\bar{1}2\bar{3}\bar{4}5\rangle$ $+2\lvert 1\bar{1}\bar{2}\bar{3}45\rangle - \lvert 1\bar{1}\bar{2}34\bar{5}\rangle - \lvert 1\bar{1}\bar{2}3\bar{4}5\rangle\big)$	$\frac{9625711}{2799360}Z$	$3.4385Z$
	6	2	-1	$\frac{1}{\sqrt{6}}\left(2\lvert 1\bar{1}2\bar{3}45\rangle - \lvert 1\bar{1}234\bar{5}\rangle - \lvert 1\bar{1}23\bar{4}5\rangle\right)$	$\frac{4730843}{1399680}Z$	$3.3799Z$

Table 4.6: V_{ee} matrix element values for Li-C, calculated by using the integral values of Table 2.3 in the energy expressions of Tables 4.3 and 4.4.

	$\langle \underline{L}^2 \rangle$	$\langle \underline{S}^2 \rangle$	$\langle \hat{R} \rangle$	Ψ	$\langle V_{ee} \rangle$	
N	0	$\frac{3}{4}$	1	$\frac{1}{\sqrt{3}}\big(\lvert 1\bar{1}23\bar{3}4\bar{4}\rangle + \lvert 1\bar{1}23\bar{3}5\bar{5}\rangle + \lvert 1\bar{1}24\bar{4}5\bar{5}\rangle\big)$	$\frac{3843463}{839808}Z$	$4.5766Z$
	0	$\frac{15}{4}$	-1	$\lvert 1\bar{1}2\bar{2}345\rangle$	$\frac{2437421}{559872}Z$	$4.3535Z$
	2	$\frac{3}{4}$	-1	$0.9913\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}2\bar{2}34\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}35\bar{5}\rangle\right) + 0.1312\lvert 1\bar{1}34\bar{4}5\bar{5}\rangle$		$4.4008Z$
				$0.1312\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}2\bar{2}34\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}35\bar{5}\rangle\right) - 0.9913\lvert 1\bar{1}34\bar{4}5\bar{5}\rangle$		$4.7192Z$
			1	$\frac{1}{\sqrt{6}}\big(2\lvert 1\bar{1}\bar{2}3\bar{3}45\rangle - \lvert 1\bar{1}23\bar{3}\bar{4}5\rangle - \lvert 1\bar{1}23\bar{3}4\bar{5}\rangle\big)$	$\frac{15393535}{3359232}Z$	$4.5825Z$
	2	$\frac{15}{4}$	1	$\lvert 1\bar{1}23\bar{3}45\rangle$	$\frac{7549145}{1679616}Z$	$4.4946Z$
	6	$\frac{3}{4}$	-1	$\frac{1}{\sqrt{6}}\big(2\lvert 1\bar{1}2\bar{2}\bar{3}45\rangle - \lvert 1\bar{1}2\bar{2}34\bar{5}\rangle - \lvert 1\bar{1}2\bar{2}3\bar{4}5\rangle\big)$	$\frac{24551357}{5598720}Z$	$4.3852Z$
			1	$\frac{1}{\sqrt{6}}\big(2\lvert 1\bar{1}24\bar{4}5\bar{5}\rangle - \lvert 1\bar{1}23\bar{3}4\bar{4}\rangle - \lvert 1\bar{1}23\bar{3}5\bar{5}\rangle\big)$	$\frac{76337819}{16796160}Z$	$4.5445Z$
O	0	0	1	$0.1443\frac{1}{\sqrt{3}}\big(\lvert 1\bar{1}2\bar{2}3\bar{3}4\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}3\bar{3}5\bar{5}\rangle + \lvert 1\bar{1}2\bar{2}4\bar{4}5\bar{5}\rangle\big)$ $+0.9895\lvert 1\bar{1}3\bar{3}4\bar{4}5\bar{5}\rangle$		$6.0627Z$
				$0.9895\frac{1}{\sqrt{3}}\big(\lvert 1\bar{1}2\bar{2}3\bar{3}4\bar{4}\rangle + \lvert 1\bar{1}2\bar{2}3\bar{3}5\bar{5}\rangle + \lvert 1\bar{1}2\bar{2}4\bar{4}5\bar{5}\rangle\big)$ $-0.1443\lvert 1\bar{1}3\bar{3}4\bar{4}5\bar{5}\rangle$		$5.7072Z$
	2	0	-1	$\frac{1}{\sqrt{2}}\left(\lvert 1\bar{1}2\bar{3}4\bar{4}5\bar{5}\rangle - \lvert 1\bar{1}\bar{2}34\bar{4}5\bar{5}\rangle\right)$	$\frac{9897607}{1679616}Z$	$5.8928Z$
	2	2	-1	$\lvert 1\bar{1}234\bar{4}5\bar{5}\rangle$	$\frac{1224899}{209952}Z$	$5.8342Z$
			1	$\lvert 1\bar{1}2\bar{2}3\bar{3}45\rangle$	$\frac{4754911}{839808}Z$	$5.6619Z$
	6	0	1	$\frac{1}{\sqrt{6}}\big(2\lvert 2\bar{2}4\bar{4}5\bar{5}\rangle - \lvert 2\bar{2}3\bar{3}4\bar{4}\rangle - \lvert 2\bar{2}3\bar{3}5\bar{5}\rangle\big)$	$\frac{47726257}{8398080}Z$	$5.6830Z$
F	0	$\frac{3}{4}$	1	$\lvert 1\bar{1}23\bar{3}4\bar{4}5\bar{5}\rangle$	$\frac{4108267}{559872}Z$	$7.3379Z$
	2	$\frac{3}{4}$	-1	$\lvert 1\bar{1}2\bar{2}34\bar{4}5\bar{5}\rangle$	$\frac{11982943}{1679616}Z$	$7.1343Z$
Ne	0	0	1	$\lvert 1\bar{1}2\bar{2}3\bar{3}4\bar{4}5\bar{5}\rangle$	$\frac{2455271}{279936}Z$	$8.7708Z$

Table 4.7: V_{ee} matrix element values for N-Ne, calculated by using the integral values of Table 2.3 in the energy expressions of Table 4.5.

	State	E_{Pert}	E_{Exp}	ΔE_{Pert}	ΔE_{Exp}
Li	^2S	-7.0566	-7.4779		
	^2P	-6.8444	-7.4100	0.2122	0.0679
Be	^1S	-13.7629	-14.6684		
	^3P	-13.5034	-14.5683	0.2595	0.1001
	^1P	-13.2690	-14.4745	0.4939	0.1939
	^1D	-13.0112	-14.4092	0.7517	0.2592
	^3P	-13.0955	-14.3964	0.6674	0.2720
	^1S	-12.8377	-14.3212	0.9252	0.3471
B	^2P	-22.7374	-24.6581		
	^4P	-22.4273	-24.5265	0.3101	0.1316
	^2D	-22.1753	-24.4401	0.5620	0.2181
	^2S	-22.0171	-24.3685	0.7202	0.2896
	^2P	-21.9878	-24.3277	0.7496	0.3305
	^4S	-21.7612	-24.2157	0.9762	0.4424
	^2D	-21.6030	-24.2034	1.1344	0.4547
	^2P	-21.4629	-24.1319	1.2745	0.5602
C	^3P	-34.4468	-37.8558		
	^1D	-34.3202	-37.8094	0.1266	0.0464
	^1S	-34.1838	-37.7572	0.2630	0.0986
	^5S	-34.0859	-37.7021	0.3608	0.1537
	^3D	-33.7203	-37.5638	0.7265	0.2920
	^3P	-33.5938	-37.5129	0.8530	0.3429
	^1D	-33.3688	-37.4100	1.0780	0.4458
	^3S	-33.3828	-37.3737	1.0640	0.4821
	^1P	-33.2422	-37.3096	1.2046	0.5462
	^3P	-32.7641		1.6826	
	^1D	-32.6376		1.8092	
	^1S	-32.3943		2.0524	

Table 4.8: Comparison with Experimental data [Huh93], ΔE is the energy above the respective ground state (Li-C).

	State	E_{Pert}	E_{Exp}	ΔE_{Pert}	ΔE_{Exp}
N	^4S	-49.1503	-54.6117		
	^2D	-48.9288	-54.5241	0.2215	0.0876
	^2P	-48.8195	-54.4803	0.3307	0.1314
	^4P	-48.1630	-54.2101	0.3931	0.4016
	^2D	-47.8103	-54.0595	1.3400	0.5522
	^2S	-47.5888		1.5615	
	^2P	-47.5478		1.6025	
	^2P	-46.5905		2.5597	
O	^3P	-66.7048	-75.1080		
	^1D	-66.5360	-75.0357	0.1688	0.0723
	^1S	-66.3421	-74.9540	0.3627	0.1540
	^3P	-65.3265	-74.5324	1.3783	0.5756
	^1P	-64.8578		1.8470	
	^1S	-63.4984		3.2063	
F	^2P	-87.6660	-99.8060		
	^2S	-85.8342	-99.0322	1.8318	0.7738
Ne	^1S	-112.2917	-129.0500		

Table 4.9: Comparison with Experimental data [Huh93], ΔE is the energy above the respective ground state (N-Ne).

4.2 Improving the Energy

We wish to investigate why the spectral gaps in Tables 4.8 and 4.9 are a long way from the experimental values, in contrast to the reasonably accurate ground state energies. We begin by discussing the idea of screening. Physically, this is a method used to describe the 'effective potential' felt by an electron. The justification for this is that for a given electron, other electrons in the atom that have a high probability of being closer to the nucleus will screen the charge of the nucleus. Hence, shells with higher principal quantum number, which are more likely to be further away from the nucleus, should have a lower 'effective potential'.

One of the first people to introduce this concept (albeit in an empirical way) was Slater. In [Sla30] he proposed a system by which the effective potentials for each electron could be calculated empirically for a given (non-correlated) wavefunction by considering the other electrons present in the wavefunction. The ground state energies, bond lengths, and atomic and ionic radii of the resulting wavefunctions compare reasonably well with experiment for a number of atoms. These empirical results were based upon the experimental data available. We wish to use a similar method that is mathematically rigorous and needs no experimental data.

Mathematically, the effective potential is described by the value of Z in the one-electron orbitals and varying Z can be justified by the variational principle. Suppose that for $\Psi(Z)$ with corresponding energy $E(Z)$, $\Psi(Z_*)$ is the ground state of H, i.e. is the minimizer of the energy, and so $\frac{d}{dZ}E(Z*) = 0$, giving an equation which can be solved for Z_*. This method is commonly used (see, for example, [Sch01]) to find an upper bound for the Helium ground state and it improves the ground state energy from -2.5 to $-2\frac{217}{256} \sim -2.848$. This compares to an experimental value of -2.9035 and hence the variational method captures an extra 12% of the experimental energy.

We redefine the one-electron orbitals from Section 2.4 in terms of a variational parameter Z_*:

$$\psi_1 := \psi_{1s}(Z_*) := \frac{Z_*^{3/2}}{\sqrt{\pi}} e^{-Z_* r} = \psi_{1s},$$

$$\psi_2 := \psi_{2s}(Z_*) := \frac{Z_*^{3/2}}{\sqrt{8\pi}} \left(1 - \frac{Z_* r}{2}\right) e^{-Z_* r/2},$$

$$\psi_3 := \psi_{2p_z}(Z_*) := \frac{Z_*^{5/2}}{\sqrt{32\pi}} x_3 e^{-Z_* r/2},$$

$$\psi_4 := \psi_{2p_x}(Z_*) := \frac{Z_*^{5/2}}{\sqrt{32\pi}} x_1 e^{-Z_* r/2},$$
$$\psi_5 := \psi_{2p_y}(Z_*) := \frac{Z_*^{5/2}}{\sqrt{32\pi}} x_2 e^{-Z_* r/2}.$$

Note that this moves away from the perturbation theory argument of Theorem 2.2.2, although this will hopefully be justified by the resulting improvement in the energies.

In order to calculate the energy of a Slater determinant formed from these orbitals we note that the term from V_{ee} scales linearly in Z_*, as can be seen from the linear dependence on Z of the integrals in Table 2.3. However, these orbitals are no longer eigenfunctions for an atomic H_0 as this requires $Z_* = Z$, i.e. $\psi_i(Z_*)$ is an eigenfunction of $H_0(Z_*) = -\frac{1}{2}\Delta_x - \frac{Z_*}{|x|}$. Hence we rewrite H_0 as

$$H_0(Z) = H_0(Z_*) - \frac{Z - Z_*}{|x|} = H_0(Z_*) - \frac{Z_*}{|x|}\frac{(Z - Z_*)}{Z_*}.$$

We may now use the virial theorem which, since $\psi_i(Z_*)$ is an eigenfunction of $H_0(Z_*)$ gives $\langle\frac{-Z_*}{|x|}\rangle_{\psi_i(Z_*)} = 2E_0(Z_*) = -Z_*^2\frac{1}{n_i^2}$. So we now have, for an N electron wavefunction,

$$\begin{aligned}\langle H\rangle_{\Psi(Z_*)} &= -Z_*^2\Big(1 + 2\frac{Z - Z_*}{Z_*}\Big)\sum_{i=1}^{N}\frac{1}{2n_i^2} + \langle V_{ee}\rangle_{\Psi(Z_*)}\\ &= (Z_*^2 - 2ZZ_*)\sum_{i=1}^{N}\frac{1}{2n_i^2} + Z_*\langle V_{ee}\rangle_{\Psi(1)},\end{aligned}$$

and differentiating with respect to Z_* gives

$$\frac{d}{dZ_*}\langle H\rangle_{\Psi(Z_*)} = (Z_* - Z)\sum_{i=1}^{N}\frac{1}{n_i^2} + \langle V_{ee}\rangle_{\Psi(1)},$$

which is zero at $Z_* = Z - \Big(\sum_{i=1}^{N}\frac{1}{n_i^2}\Big)^{-1}\langle V_{ee}\rangle_{\Psi(1)}$. This gives the minimum value of $E_0(Z_*)$:

$$E_0 := -\frac{1}{2}\Big(\sum_{i=1}^{N}\frac{1}{n_i^2}\Big)\Big(Z - \Big(\sum_{i=1}^{N}\frac{1}{n_i^2}\Big)^{-1}\langle V_{ee}\rangle_{\Psi(1)}\Big)^2.$$

For example, in the case of Helium with the ground state $|1\overline{1}\rangle$ we have that $\sum_{i=1}^{N}\frac{1}{n_i^2} = 2$, $Z = 2$ and $\langle V_{ee}\rangle_{|1\overline{1}\rangle} = (11|11) = \frac{5}{8}$. This gives a minimizing value of

$$E_0 = -\tfrac{1}{2}(2)\left(2 - \tfrac{1}{2}\tfrac{5}{8}\right)^2 = -2\tfrac{217}{256},$$

	State	E_{Pert}	E_{Z_*}	E_{Exp}	Z_*
Li	^2S	-7.0566	-7.2891	-7.4779	2.5454
Be	^1S	-13.7629	-14.2492	-14.6684	3.3763
B	^2P	-22.7374	-23.7223	-24.6581	4.1536
C	^3P	-34.4468	-36.2168	-37.8558	4.9137
N	^4S	-49.1503	-52.0662	-54.6117	5.6605
O	^3P	-66.7048	-71.2844	-75.1080	6.3823
F	^2P	-87.6660	-94.4525	-99.8060	7.0975
Ne	^1S	-112.2917	-121.9076	-129.0500	7.8073

Table 4.10: Comparison of Perturbation theory (E_{Pert}), the one-parameter variational method (E_{Z_*}), and Experimental data (E_{Exp}). Z_* is the value of the variational parameter.

which agrees with the textbook analysis.

Applying this method to Lithium ground state of $|1\bar{1}2\rangle$ we have $\sum_{i=1}^{N}\frac{1}{n_i^2}=\frac{9}{4}$, $Z=3$, and $\langle V_{ee}\rangle_{|1\bar{1}2\rangle}=\frac{5965}{5832}$, giving a minimum energy of

$$E_0=-\frac{1}{2}\frac{9}{4}\left(3-\frac{4}{9}\frac{5965}{5832}\right)^2=-\frac{1115626801}{153055008}\sim-7.2891$$

at $Z_*=\frac{33401}{13122}\sim 2.5454$. This takes the percentage of the experimental ground state energy captured by our method from 94.3% to 97.5%.

We now consider the excited state given by $|1\bar{1}3\rangle$ which has $\sum_{i=1}^{N}\frac{1}{n_i^2}=\frac{9}{4}$, $Z=3$, and $\langle V_{ee}\rangle_{|1\bar{1}3\rangle}=\frac{57397}{52488}$ which gives

$$E_1=-\frac{88147828609}{12397455648}\sim-7.1102,$$

with a corresponding Z_* of $\frac{296897}{118098}\sim 2.5140$. This predicts a spectral gap of around 0.1789 which, whilst an improvement on the previous value, is still a factor of around 2.5 times the experimental value.

From this we expect that the ground state energies are improved considerably with the use of one variational parameter, but the effect on the spectral gaps is negligible compared to the error already present. Hence we give only the ground state energies in Table 4.10 and then investigate further ways to improve the energy, but, more importantly, to improve the spectral gaps.

4.2.1 Multiple Parameters

Our next step again has a physical motivation in terms of effective potentials. If we think of Z_* as the effective potential then the previous method forces this value to be the same for each orbital. However, intuitively, the value should be lower for the $2s$-orbital in $|1\bar{1}2\rangle$ than for the $1s$-orbitals, since the $1s$ electrons are more likely to be closer to the nucleus and so will screen the $2s$ orbital to a larger extent. Hence we wish to introduce separate variational parameters for each type orbital, i.e. we write

$$\begin{aligned}
\psi_1 &:= \psi_{1s}(Z_1) := \frac{Z_1^{3/2}}{\sqrt{\pi}} e^{-Z_1 r}, \\
\psi_2 &:= \psi_{2s}(Z_2) := \frac{Z_2^{3/2}}{\sqrt{8\pi}} \left(1 - \frac{Z_2 r}{2}\right) e^{-Z_2 r/2}, \\
\psi_3 &:= \psi_{2p_z}(Z_3) := \frac{Z_3^{5/2}}{\sqrt{32\pi}} x_3 e^{-Z_3 r/2}, \\
\psi_4 &:= \psi_{2p_x}(Z_3) := \frac{Z_3^{5/2}}{\sqrt{32\pi}} x_1 e^{-Z_3 r/2}, \text{ and} \\
\psi_5 &:= \psi_{2p_y}(Z_3) := \frac{Z_3^{5/2}}{\sqrt{32\pi}} x_2 e^{-Z_3 r/2}.
\end{aligned}$$

However, we now have a problem in that the $1s$- and $2s$-orbital are no longer orthogonal (the s- and p-orbitals are still orthogonal due to the angular terms). A large part of our analysis, including the very important Slater's rules of Corollary 2.3.2, relies on the one-electron orbitals be orthonormal. (There is in fact an analogous result for non-orthogonal orbitals [McW89]. However, it loses a lot of the simplicity and elegance of the orthogonal version and is difficult to compute by hand.) Hence we wish to find a similar basis set but with orthogonal orbitals. For this we choose to keep the $1s$-orbital unchanged and restrict the $2s$-orbital so that it is orthogonal to the $1s$-orbital.

One reason for not restricting the $1s$-orbital is that it is present in every Slater determinant. If, for example, we allowed the $2s$-orbital to vary freely and restricted the $1s$-orbital to be orthogonal, then the $1s$-orbital in, say $|1s\overline{1s}2s\rangle$ could be vastly different to that in $|1s\overline{1s}2p\rangle$. This would disagree with the conventional wisdom that the inner shells are largely unaffected by the specific choice of valence electrons.

4.2.2 Orthogonal Orbitals and their Fourier Transforms

We now discuss a number of different ways of producing a '$2s$'-orbital which is orthogonal to a given $1s$-orbital with effective charge Z_1. First we characterize what we believe to be the important properties of a '$2s$'-orbital:

- it should have no angular dependence, this would, for a start, introduce issues with orthogonality with the $2p$-orbitals;
- it must have a node, or else cannot be orthogonal to the $1s$-orbitals;
- it should be characterized by the position of the node and the decay rate, which are not independent.

The resulting question is how best to choose the new $2s$-orbital. There are a number of possible methods, the first of which would be simply to use Gram-Schmidt orthogonalization in which we simply subtract $\langle 1s|2s\rangle|1s\rangle$ from the freely varied $2s$-orbital. This method has a number of disadvantages, primarily that it destroys the relatively simple structure of the $2s$-orbital, causing it to be the sum of two terms with different long-range decay rates.

Hence, it is necessary to consider restrictions on the position of the node and the decay rate of the $2s$-orbital which ensure orthogonality. It is clear that one parameter may be freely varied, the other being fixed by orthogonality to the $1s$-orbital.

There seems to be little physical or mathematical intuition regarding the choice of which to freely vary. We therefore try both possibilities for a simple example, that of the Lithium ground state, and need to derive the restriction on the fixed parameter in each case. We begin by allowing the decay rate Z_2 to vary freely and follow, to some extent, Problem 7.61 of [tH75]. Consider

$$\psi_2 := c(1-\gamma r)e^{-Z_2 r/2},$$

which for orthogonality with ψ_1 gives

$$\begin{aligned}
&4\pi c\frac{Z_1^{3/2}}{\sqrt{\pi}}\int_0^\infty (1-\gamma r)e^{-(Z_1+Z_2/2)r}r^2dr = 0\\
\Leftrightarrow &\int_0^\infty (1-\gamma r)e^{-(Z_1+Z_2/2)r}r^2dr = 0\\
\Leftrightarrow &\int_0^\infty e^{-(Z_1+Z_2/2)r}r^2dr = \frac{3\gamma}{(Z_1+Z_2/2)}\int_0^\infty e^{-(Z_1+Z_2/2)r}r^2dr\\
\Leftrightarrow &\gamma = \tfrac{1}{6}(2Z_1+Z_2),
\end{aligned}$$

where the penultimate line is found using integration by parts. It remains to normalize ψ_2 using this value of γ, i.e. to find c such that

$$4\pi \int_0^\infty c^2(1-\gamma r)^2 e^{-Z_2 r} r^2 dr = 1,$$

giving

$$4\pi c^2 \frac{2(Z_2^2 - 6\gamma Z_2 + 12\gamma^2)}{Z_2^5} = 1,$$

and hence

$$c^2 = \frac{3Z_2^5}{8\pi(4Z_1^2 - 2Z_1Z_2 + Z_2^2)}.$$

Therefore, our new $2s$-orbital is

$$\psi_2 := c(1-\gamma r)e^{-Z_2 r/2} = \left(\tfrac{3Z_2^5}{8\pi(4Z_1^2-2Z_1Z_2+Z_2^2)}\right)^{1/2} \left(1 - \tfrac{1}{6}(2Z_1 + Z_2)r\right) e^{-Z_2 r/2}.$$

As noted before, this is no longer an eigenfunction of the non-interacting Hamiltonian H_0 and hence we need to explicitly find the kinetic and nuclear potential terms for ψ_2. These are simple one-electron integrals and are given by

$$\begin{aligned} T_2 &= \frac{Z_1Z_2}{2} + \frac{7Z_2^2}{24} - \frac{2Z_1^3Z_2}{4Z_1^2 - 2Z_1Z_2 + Z_2^2} \\ &= \frac{Z_2^2}{24}\frac{4Z_1^2 - 2Z_1Z_2 + 7Z_2^2}{4Z_1^2 - 2Z_1Z_2 + Z_2^2}, \\ V_2 &= -\frac{ZZ_2}{4}\frac{4Z_1^2 - 4Z_1Z_2 + 3Z_2^2}{4Z_1^2 - 2Z_1Z_2 + Z_2^2}. \end{aligned}$$

As expected, if we set $Z_1 = Z_2$ all the above simplify to the previous non-variational case. The kinetic and potential energies for the $1s$- and $2p$-orbitals are simply

$$T_1 = \frac{1}{2}Z_1^2, \quad V_1 = -ZZ_1, \quad T_3 = \frac{1}{8}Z_3^2, \quad V_3 = -\frac{1}{4}ZZ_3.$$

It remains to to recompute the Fourier transforms in Tables 2.2 for the new wavefunctions, the energies of which have the same integral expressions as in Tables 4.3-4.5. The method for this is exactly the same as in Section 2.5.1, starting with the result from Lemma 2.5.1 and using the normal rules of Fourier transforms. We note that for integrals containing only $1s$-orbitals or only $2p$-orbitals the Fourier transforms are simply given by replacing Z in the previous formulae by Z_1 and Z_3 respectively. The resulting Fourier transforms are shown in Table 4.11, and once again appear to be novel.

Function	Fourier Transform
$\psi_{1s}\psi_{1s}$	$\frac{16Z_1^4}{(4Z_1^2+\vert k\vert^2)^2}$
$\psi_{2s}\psi_{2s}$	$\frac{Z_2^5}{(4Z_1^2-2Z_1Z_2+Z_2^2)}\left(\frac{2(Z_1+2Z_2)}{(Z_2^2+\vert k\vert^2)^2}-\frac{Z_2(2Z_1+Z_2)(2Z_1+5Z_2)}{(Z_2^2+\vert k\vert^2)^3}+\frac{2Z_2^3(2Z_1+Z_2)^2}{(Z_2^2+\vert k\vert^2)^4}\right)$
$\psi_1\psi_2$	$\frac{\sqrt{6}Z_1^{3/2}Z_2^{5/2}}{\sqrt{(4Z_1^2-2Z_1Z_2+Z_2^2)}}\left(\frac{4(2Z_1+Z_2)}{3((Z_1+Z_2/2)^2+\vert k\vert^2)^2}-\frac{(2Z_1+Z_2)^3}{3((Z_1+Z_2/2)^2+\vert k\vert^2)^3}\right)$
$\psi_{p_j}\psi_{p_j}$	$\frac{Z_3^6}{(Z_3^2+\vert k\vert^2)^3}-\frac{6Z_3^6k_j^2}{(Z_3^2+\vert k\vert^2)^4}$
$\psi_1\psi_{p_j}$	$-\frac{2\sqrt{2}iZ_1^{3/2}Z_3^{5/2}(2Z_1+Z_3)k_j}{((Z_1+Z_3/2)^2+\vert k\vert^2)^3}$
$\psi_2\psi_{p_j}$	$\frac{\sqrt{3}iZ_2^{5/2}Z_3^{5/2}}{16\sqrt{4Z_1^2-2Z_1Z_2+Z_2^2}}\left(\frac{8(Z_2+Z_3)^2(2Z_1+Z_2)k_j}{(((Z_2+Z_3)/2)^2+\vert k\vert^2)^4}-\frac{(32Z_1+64Z_2+48Z_3)k_j}{3(((Z_2+Z_3)/2)^2+\vert k\vert^2)^3}\right)$
$\psi_{p_j}\psi_{p_\ell}$	$-\frac{6k_jk_\ell Z_3^6}{(Z_3^2+\vert k\vert^2)^4}$

Table 4.11: Fourier Transforms for Coulomb and Exchange Integrals for variational wavefunctions.

As before, the next stage is to calculate the Coulomb and exchange integrals required. The method is identical to that in Section 2.5.2, except that we use the Fourier transforms from Table 4.11 rather than from Table 2.2. The explicit expressions for the integrals are given in Table 4.12.

We are now in a position to calculate the energy of a general linear combination of Slater determinants formed from the above one-electron orbitals. However, before we continue, we return to the question of the choice of the $2s$-orbital. Recall that in the above we have freely chosen the decay rate parameter and restricted the position of the node. The calculations for the case where we freely choose the node and restrict the decay rate are analogous and are not repeated here. We then insert the resulting expressions into the energies for $|1s\overline{1s}2s_d\rangle$ and $|1s\overline{1s}2s_n\rangle$ where $2s_d$ and $2s_n$ are the $2s$ orbitals resulting from freely varying the decay rate and node position respectively. The resulting energies and parameters are given in Table 4.13 where we use the notation $2s = c(1-\alpha r)e^{-\beta r/2}$, which for the unscaled orbitals has $\alpha = 3/2$ and $\beta = 3$.

It is worth noting the large rearrangement of the orbital from the naïve unscaled form and also the relatively close agreement between the energies and parameters for the two methods. However, to be practical, we need to choose one method and based on the lower energy for the Lithium ground state when we freely vary the decay rate we choose this option.

Symbol	Expression
$(11\|11)$	$\frac{5}{8}Z_1$
$(11\|22)$	$\frac{Z_1Z_2(8Z_1^4+4Z_1^3Z_2+4Z_1Z_2^3+Z_2^4)}{(2Z_1+Z_2)^3(4Z_1^2-2Z_1Z_2+Z_2^2)}$
$(12\|21)$	$\frac{16Z_1^3Z_2^5}{(4Z_1^2-2Z_1Z_2+Z_2^2)(2Z_1+Z_2)^5}$
$(22\|22)$	$\frac{Z_2}{512}\frac{(1488Z_1^4-1952Z_1^3Z_2+1752Z_1^2Z_2^2-840Z_1Z_2^2+245Z_2^4)}{(4Z_1^2-2Z_1Z_2+Z_2^2)^2}$
$(11\|33)$	$\frac{Z_1Z_3\left(8Z_1^4+20Z_1^3Z_3+20Z_1^2Z_3^2+10Z_1Z_3^3+Z_3^4\right)}{(2Z_1+Z_3)^5}$
$(13\|31)$	$\frac{112Z_1^3Z_3^5}{3(2Z_1+Z_3)^7}$
$(22\|33)$	$\frac{Z_2Z_3}{4(4Z_1^2-2Z_1Z_2+Z_2^2)(Z_2+Z_3)^7}\times$ $[(4Z_1^2-2Z_1Z_2+Z_2^2)(Z_2^6+7Z_2^5Z_3+21Z_2^4Z_3^2+35Z_2^3Z_3^3)$ $+3Z_2^2Z_3^4(28Z_1^2-28Z_1Z_2+11Z_2^2)$ $+7Z_2Z_3^5(4Z_1^2-4Z_1Z_2+3Z_2^2)+Z_3^6(4Z_1^2-4Z_1Z_2+3Z_2^2))]$
$(23\|32)$	$\frac{Z_2^5Z_3^5(740Z_1^2+152Z_1Z_2+17Z_2^2-42Z_2Z_3-588Z_1Z_3+126Z_3^2)}{9(Z_2+Z_3)^9(4Z_1^2-2Z_1Z_2+Z_2^2)}$
$(33\|33)$	$\frac{501}{2560}Z_3$
$(33\|44)$	$\frac{447}{2560}Z_3$
$(34\|43)$	$\frac{27}{2560}Z_3$

Table 4.12: Coulomb and Exchange Integrals for variational wavefunctions.

	Energy	Z_1	α	β
$2s_d$	-7.4139	2.6937	1.1535	1.5334
$2s_n$	-7.4008	2.6969	1.1400	1.4460

Table 4.13: Comparison of the two methods of choosing the $2s$ orbital, which is of the form $2s = c(1-\alpha r)e^{-\beta r/2}$.

4.2.3 Energies in the Variational Basis

We now wish to compute the energies for the $\underline{L}^2$-$\underline{S}^2$ eigenfunctions formed from these one-electron orbitals. As discussed in Section 3.4, the simultaneous eigenspaces are identical to those given in Section 3.2.2. This demonstrates an advantage of retaining orthonormality and also restricting the orbitals in the same n and ℓ shell to have the same variational parameter.

From the expressions in Tables 4.3-4.5 and the Coulomb and exchange integrals in Table 4.12, we see that the energy expressions are going to be very complicated. For this reason we do not give explicit expressions for the energies but use Maple to substitute the explicit expressions into the energies and then minimize over the Z_i. However, in order to ensure orthogonality of wavefunctions within an eigenspace, we need to place restrictions on this minimization when an $\underline{L}^2$-$\underline{S}^2$-L_3-

S_3-$\hat{R}$ eigenspace has degeneracy greater than one. In this case we minimize each of the energies within an eigenspace and then use the values for Z_i in the lowest energy state to give the energies of the other states.

This is not an ideal situation as it further compounds the problems discussed towards the end of Section 4.1.1 with regards to orthogonality of the excited states. Now, not only may we be restricting to be orthogonal to the incorrect ground state but we are also forcing orthogonality to the lower energy terms in the same eigenspace, which again may not be accurate representations of the physical wavefunctions corresponding to the appropriate spectral line (although we believe that our wavefunction are in fact reasonable approximations to the physical wavefunctions). We are also forcing this orthogonality in a very specific and slightly arbitrary way, and hence the energies of the higher energy states in an eigenspace may be relatively poor.

The computed energies are given in Tables 4.14 and 4.15. The explicit wavefunctions are not given as they are the same as those given in Tables 4.6 and 4.7 for the non-variational case. The only difference in this case is the value of c in correlated wavefunctions where $\Psi = c\Psi_1 + \sqrt{1-c^2}\Psi_2$, in particular the ordering of eigenfunctions within an eigenspace is the same as in the non-variational case. In fact, in all cases, the ordering follows the rule that, setting $c = 1$, the lowest state has most $2s$-orbitals and the energy increases as the number of $2s$-orbitals decreases. This corresponds to the shell-filling scheme in Hund's rules.

From Table 4.16, we see that in all cases, the ground state energy differs from the experimental value by no more than 1.1%. This is a large improvement over the previous methods (see for example Table 4.1). We also see that the energy error is no longer monotonically increasing across the period which suggests that, to some extent, we have managed to remove the proposed deficiency of our previous (non-variational) basis for higher Z values. There still seems to be some decrease in the accuracy as Z increases but it is hard to say whether this is in fact due to the increase in nuclear charge or due to the decrease in the dimension of the space to be minimized over as given in Table 2.1.

Figure 4.1 shows the variational theoretical energies and the experimental energies of the ground states, also included is the Helium data that was calculated at the start of Section 4.2 (it is a one-parameter variation). As can be seen, there is very good agreement between the two sets of data.

At the start of this section we mentioned the work of Slater [Sla30] and it is interesting to compare our effective charges (i.e the exponential decay rates) with his empirical approximations. We find that the $1s$-orbitals agree very well in all cases with his prediction of $Z - 0.3$. However, one simplification in [Sla30] that

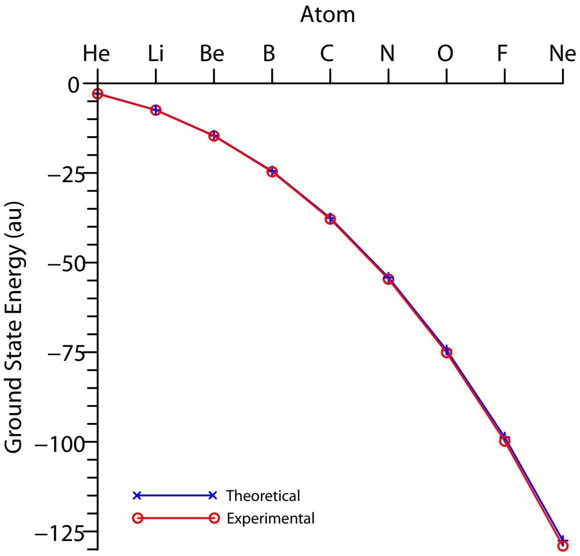


Figure 4.1: Experimental and Variational Ground State Energies.

differs from our method is the grouping of the $2s$- and $2p$-orbitals to have the same effective charge. He predicts a value of $Z - (2 \times 0.85) - ((Z - 3) \times 0.35)$, which generally lies between our variational values for Z_2 and Z_3. In general Slater's value becomes closer to that of Z_3 as Z increases. This is unsurprising if we take Slater's effective potential to be the weighted average of our Z_2 and Z_3, where there are progressively more $2p$-orbitals than $2s$-orbitals.

It is of course trivial to restrict our minimization to set $Z_2 = Z_3$ and we find that the energies are poorer by about 0.2% whilst the spectral gaps are generally very slightly larger than with three variational parameters. This shows that the gain in adding a second parameter is much greater than that of adding the third.

Comparing the calculated energy levels to the experimental values (Tables 4.17 and 4.18), we see that, although the general agreement with experiment for the ordering of the spectral levels is excellent, there are two pairs of levels which are swapped. These are the ${}^1D$ and ${}^3P$ levels of Beryllium, and ${}^1S$ and ${}^5S$ levels of Carbon.

There does not seem to be any obvious reason why these levels should be in poorer agreement than the others, for example, no presence of other experimental lines of the same term nearby that could cause a large perturbation of the energy when the correlation effects are included. However, it should be noted that, for the Beryllium case, the MDHF calculations of [TTST94] predict the same energy ordering as our method. The data for the MDHF method for the required states of Carbon do not seem to be available but it would be interesting to see if the prediction agrees with our method or with experiment.

As in the non-variational case we can compute the virial ratio, recall from Table 4.2 that the non-variational method gave values of around 1.6 when the ideal value is 2. We find that, for all cases, the virial ratio for the three-parameter variational calculation is 2 to at least 6 decimal places. This is unsurprising since the virial theorem follows from dilation of wavefunctions.

Looking at the spectral gaps we see that there has once again been a large improvement in agreement with experimental results. Most gaps differ by no more than a few percent, which is clearly a vast improvement on the $150 - 200\%$ or so error from the unscaled and one-parameter cases. There are a few spectral gaps, such as the ^5S state of Carbon that are 40% or so off from the experimental values, but, in general, the agreement is very good.

From Tables 4.14 and 4.15 we see that minimization over the Z_i does not, in general, affect the ordering of the energies. In fact, the only levels that change order are in Carbon. The most significant interchange is that of the ^1S and ^5S (2nd and 3rd excited states), which cause a disagreement with experiment. However, our results agree with the MDHF data and it would be somewhat unreasonable to expect our data to improve on this method. In addition there are interchanges of the higher ^1D and ^3S wavefunctions, in this case there is no experimental data available but the ordering agrees with MDHF.

The general overall independence of the ordering with respect to the values of the Z_i will be partially explained in Section 4.3. As mentioned previously (see Section 4.1.2), this will take the form of a mathematical justification of Hund's rules.

Another interesting point, which can be seen in Table 4.14, is that variational parameter Z_3 in the ^1P case is less than one. If these values are interpreted as an 'effective charge' felt by the $2p$ electron then this indicates some form of self-screening. This effect can also been seen in the MDHF method [Fro67]. In addition this shows that there is a significant rearrangement in the spatial orbitals between the singlet and triplet cases. In terms of charge density, the low value of Z_3 corresponds to a high mean distance from the nucleus.

The explanation for this in terms of the energy expressions given in Table 4.3 is analogous to that given by Hartree [HH36]. The only term by which the energies of the ^1P and ^3P states differ is the $2s-2p$ exchange term $(23|32)$. In the ^1P state this has a positive coefficient and in the ^3P case it has the more usual negative coefficient. Hence, in the ^1P case, this term acts in a repulsive way, forcing the $2s$- and $2p$-orbitals to be further apart.

This effect is not seen in the ^3P case where the $2p$-orbital should in fact be

	State	E	Z_1	Z_2	Z_3	c	E_Z
Li	^2S	-7.4139	2.6937	1.5334			-7.0566
	^2P	-7.3504	2.6858		1.0458		-6.8444
Be	^1S	-14.5795	3.7052	2.3669	1.9944	0.9410	-13.7629
	^3P	-14.4823	3.6944	2.4045	1.7807		-13.5034
	^1P	-14.3688	3.6962	2.6684	0.9324		-13.2690
	^3P	-14.3128	3.6806		1.7502		-13.0955
	^1D	-14.2764	3.6813		1.7025		-13.0112
	^1S	-14.1439	3.7052	2.3669	1.9944	0.9410	-12.8377
B	^2P	-24.4885	4.7086	3.1628	2.4660	0.9663	-22.7374
	^4P	-24.3969	4.6925	3.2440	2.4757		-22.4273
	^2D	-24.2448	4.6930	3.2432	2.3470		-22.1753
	^2S	-24.1719	4.6938	3.2710	2.2573		-22.0171
	^2P	-24.1010	4.6932	3.3746	2.1187		-21.9878
	^4S	-24.0776	4.6732		2.4432		-21.7612
	^2D	-24.0010	4.6742		2.3960		-21.6030
	^2P	-23.9076	4.7086	3.1628	2.4660	0.9663	-21.4629
C	^3P	-37.5689	5.7107	3.9670	3.1116	0.9858	-34.4468
	^1D	-37.5039	5.7114	3.9790	3.0520	0.9860	-34.3202
	^5S	-37.4974	5.6893	4.0713	3.1623		-34.0859
	^1S	-37.4656	5.7096	3.9998	3.0265	0.9545	-34.1838
	^3D	-37.2698	5.6894	4.0501	3.0739		-33.7203
	^3P	-37.2053	5.6899	4.0599	3.0389		-33.5938
	^1D	-37.0173	5.6885	4.0265	2.9773		-33.3688
	^3S	-36.9869	5.6873	3.9731	2.9938		-33.3828
	^1P	-36.9550	5.6892	4.0577	2.9316		-33.2422
	^3P	-36.7965	5.7107	3.9670	3.1116	0.9858	-32.7641
	^1D	-36.7331	5.7114	3.9790	3.0520	0.9860	-32.6376
	^1S	-36.5799	5.7096	3.9998	3.0265	0.9545	-32.3943

Table 4.14: Eigenspace energies for the variational method, Li-C. Z_i are the minimizing variational parameters and c is the correlation coefficient in the eigenfunctions, i.e. $\Psi = c\Psi_1 + \sqrt{1-c^2}\Psi_2$, Ψ_i are given in Tables 4.3 and 4.4. E_Z is the energy obtained by setting $Z_i = Z$.

	State	E	Z_1	Z_2	Z_3	c	E_Z
N	^4S	-54.1597	6.7117	4.7535	3.7924		-49.1503
	^2D	-54.0407	6.7124	4.7711	3.7317		-48.9288
	^2P	-54.0075	6.7110	4.7893	3.7162	0.9788	-48.8195
	^4P	-53.7666	6.6854	4.8658	3.7592		-48.1630
	^2D	-53.5340	6.6850	4.8414	3.7065		-47.8103
	^2S	-53.4173	6.6857	4.8575	3.6669		-47.5888
	^2P	-53.3071	6.6830	4.7591	3.6794		-47.5478
	^2P	-52.9277	6.7110	4.7893	3.7162	0.9788	-46.5905
O	^3P	-74.3931	7.7118	5.5613	4.4117		-66.7048
	^1D	-74.3004	7.7122	5.5709	4.3828		-66.5360
	^1S	-74.2328	7.7103	5.5967	4.3628	0.9749	-66.3421
	^3P	-73.7784	7.6805	5.6490	4.3916		-65.3265
	^1P	-73.4204	7.6785	5.5620	4.3549		-64.8578
	^1S	-72.8054	7.7103	5.5967	4.3628	0.9749	-63.4984
F	^2P	-98.7503	8.7112	6.3576	5.0587		-87.6660
	^2S	-97.8704	8.6748	6.4189	5.0416		-85.8342
Ne	^1S	-127.5695	9.7101	7.1469	5.7177		-112.2917

Table 4.15: Eigenspace energies for the variational method, N-Ne. Z_i are the minimizing variational parameters and c is the correlation coefficient in the eigenfunctions, i.e. $\Psi = c\Psi_1 + \sqrt{1-c^2}\Psi_2$, Ψ_i are given in Table 4.5. E_Z is the energy obtained by setting $Z_i = Z$.

Atom	Li	Be	B	C	N	O	F	Ne
Error	0.9%	0.6%	0.7%	0.8%	0.8%	1.0%	1.0%	1.1%

Table 4.16: Percentage error in the ground state energy.

closer to the $2s$ orbital than in the case where exchange energy is ignored. This is true for the results of [HH36] and can be checked in our case by performing a minimization without the exchange term. This results in the values $Z_1 = 3.6953$, $Z_2 = 2.4700$ and $Z_3 = 1.7807$, which are higher than the triplet values by 0.009, 0.0655 and -0.1768 respectively. In other words, the p-orbital contracts and the $2s$-orbital become more diffuse when including the exchange term. This indeed indicates that the p-orbital is generally closer to the to $2s$-orbital in the triplet case than when no exchange term is present.

This effect is also present in MDHF results, from [TTST94] we see that the expected radius of the p-orbital is 5.025au for the ^1P state compared to 2.808au for the ^3P state. We note here that we cannot directly compare the exponential decay rates as the MDHF wavefunctions are formed from a sum of orbitals with different decay rates and do not have such clear shell distinctions.

4.2.4 Comparison with Other Methods

We now wish to compare our results (which appear to be novel) with those of other methods of computing ground (and excited) state energies of atoms. To begin with we compare with the most basic theory, true single determinant HF. As noted previously, all (uncorrelated) ground states are single determinants, which to some extent can be seen as a consequence of aligning of the spins in the ground state, which leads to Slater determinants with few symmetries under interchanging spins.

This single determinant representation is not possible in a large number of the excited states, where a single Slater determinant cannot be placed in one space. The ground state energies can be seen in Tables 4.17 and 4.18, noting that for all ground states this is a true HF calculation despite the labelling of 'MDHF'.

The next step of approximation needed to enable excited state energies to be computed is to allow multiple determinants in a single configuration, i.e. MDHF. The fact that these energies (e.g. [TTST94]) are better than ours can be attributed predominantly to the differing ways of minimizing the energy. Whilst [TTST94] uses a numerical approach (and is generally unrestricted by a choice of basis) we have essentially the minimal basis set possible for these calculations.

It is interesting to note the energies of the Beryllium ground state where our method out-performs even numerical MDHF. This suggests that a careful choice of basis to allow wavefunctions with appropriate angular momentum and spin eigenvalues to be formed and, most importantly, correlated, is more important than brute force computation in gaining accurate energies.

Of course, if we compare to a MCHF calculation in which the near-degeneracy of states differing by the replacement of $2s\overline{2s}$ with $2p\overline{2p}$ (i.e. correlation) is taken into account, we find that the MCHF values are much better. This is simply a consequence of minimization over a much larger basis set. Comparing to [CV66] we see that, in general, the single configuration energies are similar to those of [TTST94] and so we may take the MCHF results to be representative of the accuracy despite the age of the paper. We show the results for the modified ground states of Beryllium, Boron and Carbon in Table 4.19 and it is easy to see the improvement in energy in these cases. Unsurprisingly the improvement in the energy between MDHF and MCHF is inversely proportional to the size of c in Table 4.14. This is analogous to c being a measure of the amount of correlation between the $2s\overline{2s}$ and $2p\overline{2p}$ states in our calculated wavefunctions.

4.2.5 Spectral Plots

Figures 4.2 and 4.3 show the spectra of Lithium to Fluorine (the cases of Helium and Neon are uninteresting as they consist of a single line each) in a number of different ways. Within each plot, levels of the same colour correspond to levels of the same term (e.g. ^1S, ^3P). The dashed black line is the lowest experimental energy containing an orbital with principal quantum number 3, for clarity, higher energy levels of such configurations are not shown. The dotted black line is the ground state energy of the first positive ion.

The reason for plotting these two black lines is that we expect our method to give poorer results above each of these levels. One would expect that lines close to those involving $n = 3$ or higher orbitals, and of the same term, should experience a relatively large perturbation when the extra configurations (and hence extra correlations) are included. Above the energy of the ion, the configuration with $N = Z$ electrons will no longer be the minimizer, which calls into question the validity of the method.

However, as has been noted in previous discussions, the ordering of the calculated spectra is in excellent agreement with experiment, when restricted to $n = 1$ and $n = 2$ orbitals. In general it appears that the agreement with experiment does indeed become poorer above these levels, although it is not clear if this is a direct effect of this or simply that the approximation is poorer for higher levels in general.

Levels for which there is no experimental data are not plotted, this is essentially a cut-off at a given energy level. The experimental values and variational theoretical values are those in Tables 4.17 and 4.18 (E_{exp} and E_{Z_i} respectively).

	State	E_{Z_i}	E_{Exp}	E_{MDHF}	ΔE_{Z_i}	ΔE_{Exp}	ΔE_{MDHF}
Li	^2S	-7.4139	-7.4779	-7.4327			
	^2P	-7.3504	-7.4100	-7.3651	0.0635	0.0679	0.0677
Be	^1S	-14.5795	-14.6684	-14.5730			
	^3P	-14.4823	-14.5683	-14.5115	0.0972	0.1001	0.0615
	^1P	-14.3688	-14.4745	-14.3947	0.2107	0.1939	0.1783
	^1D	-14.2764	-14.4092		0.3030	0.2592	
	^3P	-14.3128	-14.3964		0.2667	0.2720	
	^1S	-14.1439	-14.3212		0.4356	0.3471	
B	^2P	-24.4885	-24.6581	-24.5291			
	^4P	-24.3969	-24.5265	-24.4507	0.0915	0.1316	0.0784
	^2D	-24.2448	-24.4401	-24.3119	0.2437	0.2181	0.2172
	^2S	-24.1719	-24.3685	-24.2481	0.3165	0.2896	0.2810
	^2P	-24.1010	-24.3276	-24.1790	0.3875	0.3305	0.3500
	^4S	-24.0776	-24.2157		0.4807	0.4424	
	^2D	-24.0010	-24.2034		0.4876	0.4547	
	^2P	-23.9076	-24.1319		0.5808	0.5062	
C	^3P	-37.5689	-37.8558	-37.6886			
	^1D	-37.5039	-37.8094	-37.6313	0.0650	0.0464	0.0573
	^1S	-37.4656	-37.7572	-37.5496	0.1033	0.0986	0.1390
	^5S	-37.4974	-37.7021	-37.5992	0.0715	0.1537	0.0894
	^3D	-37.2698	-37.5638	-37.3944	0.2991	0.2920	0.2945
	^3P	-37.2053	-37.5129	-37.3377	0.3636	0.3429	0.3509
	^1D	-37.0173	-37.4100	-37.1696	0.5516	0.4458	0.5190
	^3S	-36.9869	-37.3737	-37.1421	0.5820	0.4821	0.5465
	^1P	-36.9550	-37.3096	-37.1158	0.6139	0.5462	0.5728
	^3P	-36.7965			0.7724		
	^1D	-36.7331			0.8358		
	^1S	-36.5799			0.9889		

Table 4.17: Comparison of three-parameter variational energies with Multi-Determinant Hartree-Fock [TTST94] and experimental energies. ΔE is the energy above the respective ground state (Li-C).

	State	E_{Z_i}	E_{Exp}	E_{MDHF}	ΔE_{Z_i}	ΔE_{Exp}	ΔE_{MDHF}
N	^4S	-54.1597	-54.6117	-54.4009			
	^2D	-54.0407	-54.5241	-54.2962	0.1190	0.0876	0.1048
	^2P	-54.0075	-54.4803	-54.2281	0.1523	0.1314	0.1728
	^4P	-53.7666	-54.2101	-53.9883	0.3932	0.4016	0.4127
	^2D	-53.5340	-54.0595	-53.7836	0.6257	0.5522	0.6173
	^2S	-53.4173		-53.6834	0.7424		0.7175
	^2P	-53.3071		-53.5839	0.8526		0.8170
	^2P	-52.9277			1.2320		
O	^3P	-74.3931	-75.1080	-74.8094			
	^1D	-74.3004	-75.0357	-74.7293	0.0928	0.0723	0.0801
	^1S	-74.2328	-74.9540	-74.6110	0.1603	0.1540	0.1984
	^3P	-73.7784	-74.5324	-74.1839	0.6147	0.5756	0.6255
	^1P	-73.4204		-73.8720	0.9727		0.9374
	^1S	-72.8054			1.5877		
F	^2P	-98.7503	-99.8060	-99.4093			
	^2S	-97.8704	-99.0322	-98.5312	0.8800	0.7738	0.8781
Ne	^1S	-127.5695	-129.0500	-128.5471			

Table 4.18: Comparison of three-parameter variational energies with Multi-Determinant Hartree-Fock [TTST94] and experimental energies. ΔE is the energy above the respective ground state (N-Ne).

	E_{Z_i}	E_{MDHF}	E_{MCHF}	E_{Exp}
Be	-14.5795	-14.5730	-14.6514	-14.6684
B	-24.4885	-24.5291	-24.5600	-24.6581
C	-37.5689	-37.6886	-37.7059	-37.8558

Table 4.19: Ground state energies of Beryllium, Boron and Carbon using three-parameter variational, Multi-Determinant Hartree-Fock [TTST94] and Multi-Configuration Hartree-Fock methods [CV66], compared to Experimental results.

The non-variational energies, taken from from Tables 4.14 and 4.15, are included to demonstrate the vast benefit of the variational scaling, and also the relative invariance of the ordering of the spectra under this variational procedure.

The plots in the left hand column show the absolute energy levels, those in the middle column show the spectral gaps by simply subtracting the relevant ground state energy from all energy levels. Finally, the plots in the right hand column demonstrate the multi-scale nature of the results, i.e. the spectral gaps are of a different order to the overlying total energies.

4.3 Validity and Failure of Hund's Rules

As noted previously, minimization over the variational parameters had little effect on the spectral ordering. This suggests that there are relationships that hold independently of the value of the coupling constant λ and the specific choice of the one-electron basis.

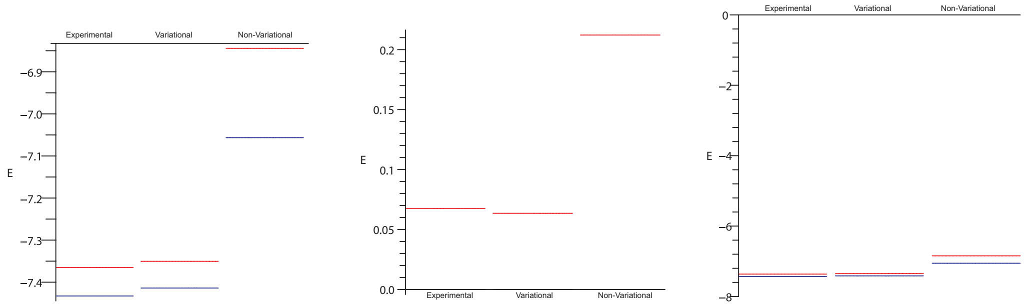
The standard empirical rules for this are Hund's rules, which were described in Section 4.1.2, the spin alignment part of which is rigorous for single Slater determinants. An interesting question is how Hund's rules translate to linear combinations of Slater determinants, for which we need to consider the extra terms generated by the interaction of different Slater determinants under the action of V_{ee}.

An additional complication is the correlation of wavefunctions, the effects of which are expected to be more difficult as there seems to be no *a priori* way to compare the magnitudes of two terms such as $(11|22)$ and $(11|33)$ for all possible $1s$-, $2s$- and $2p$-orbitals, a problem that is directly related to the lack of rigorous proof of shell ordering. This prevents us from deriving any inequalities on the energies for configurations with different numbers of $2s$- and $2p$-orbitals.

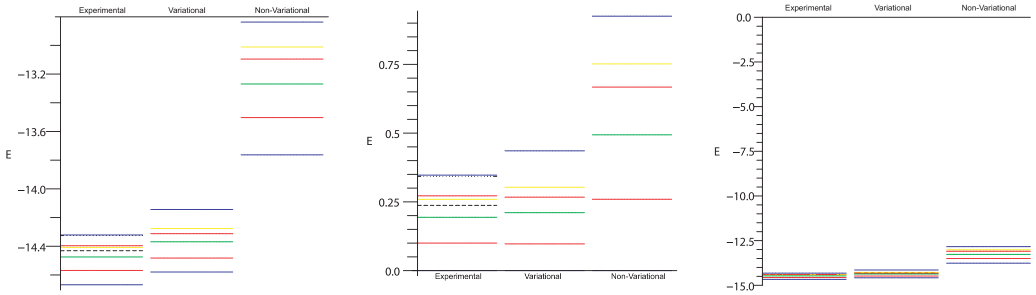
For example, from the eigenspaces in Table 3.1 we see that it is not possible to derive the ordering of $|2s\rangle$ and $|2p\rangle$ without explicit computation of the integrals. Thus we obtain no *a priori* information about Lithium.

Moving on to Beryllium we have a more interesting case. We begin by ignoring correlation (setting $c = 1$) and, as mentioned above, partition the space of $\underline{L}^2$-$\underline{S}^2$ eigenfunctions into wavefunctions containing Slater determinants with orbitals $1s^2 2s^2 2p^{N-4}$, $1s^2 2s 2p^{N-3}$ and $1s^2 2p^{N-2}$, where in this case $N = 4$. These spaces are well defined as the number of $2p$-orbitals in the sets $\{\underline{L}^2|\Psi\rangle\}$ and $\{\underline{S}^2|\Psi\rangle\}$ is invariant for any given Slater determinant $|\Psi\rangle$ (see Lemmas 3.3.7 and 3.3.9).

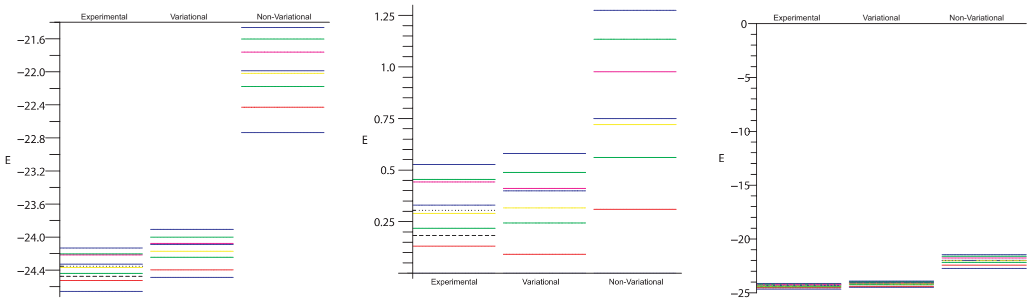
There is only one case with no $2p$-orbitals, lying in the ${}^1S$ eigenspace, and hence we consider the eigenfunctions with one $2s$-orbital. Using the term notation, these



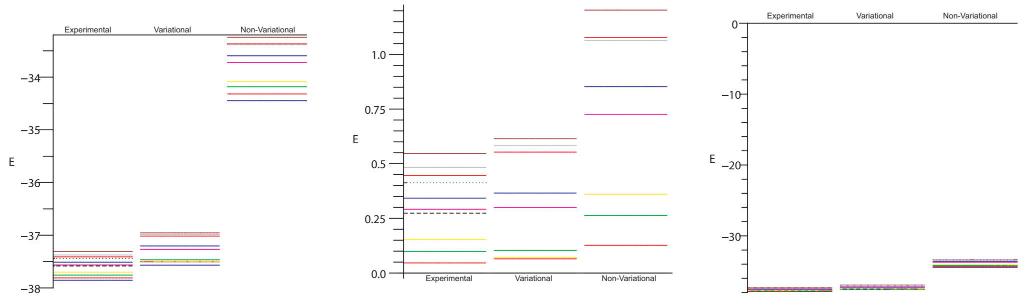
Lithium



Beryllium

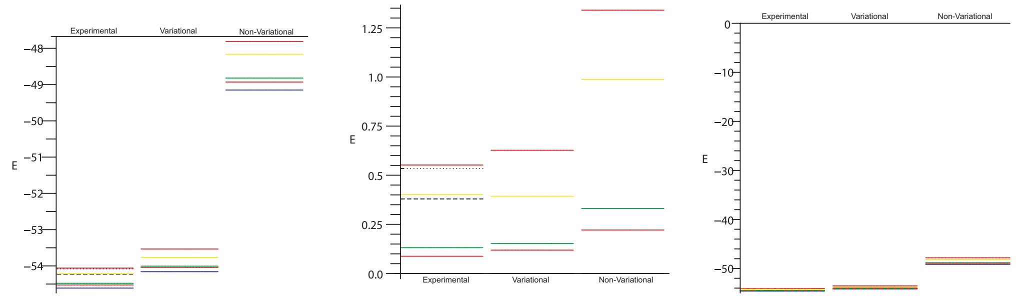


Boron

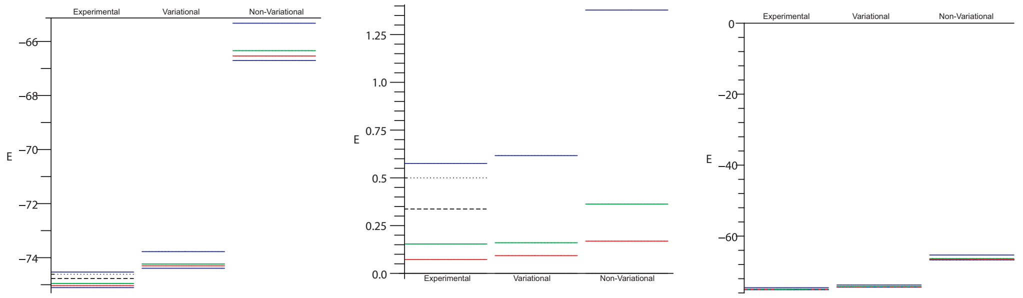


Carbon

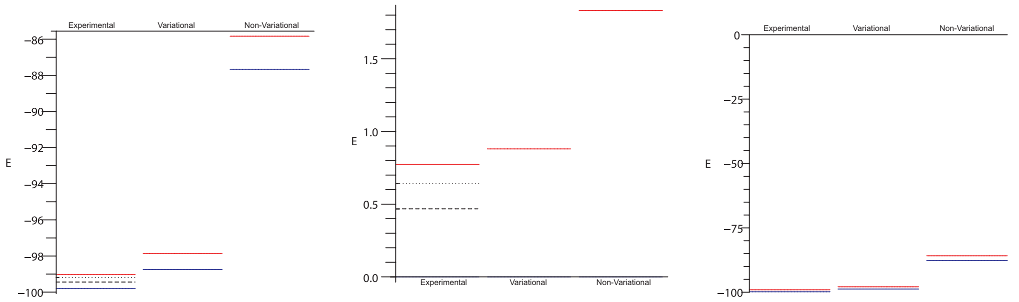
Figure 4.2: Atomic spectra of Li-C, see text for details.



Nitrogen



Oxygen



Fluorine

Figure 4.3: Atomic spectra of N-F, see text for details.

lie in the spaces ^1P and ^3P and from Hund's rules we would expect the triplet state to have lower energy. Comparing the energies in Table 4.3 we see that $E_{^3P} - E_{^1P} = -2(23|32) < 0$ (since exchange terms are positive) and hence the triplet term does have lower energy. Note that this quantitative result differs from that of the more naïve argument which states that $|2s2p\rangle$ has one exchange term and $|2s\overline{2p}\rangle$ has none.

Next we consider Slater determinants containing no $2s$-orbitals. In this case we have a wavefunction in each of the spaces ^1S, ^3P and ^1D. Hund's rules suggest that the ^3P state should have lowest energy so we use this as a reference energy, giving

$$E_{^3P} = E_{^1D} - (33|33) + (33|44) = E_{^1S} - (33|33) + (33|44) - 3(34|43),$$

from which we easily conclude that $E_{^1D} < E_{^1S}$. However, the relationships involving $E_{^3P}$ are less clear and we need an inequality involving $(33|33)$ and $(33|44)$.

Lemma 4.3.1. *For all orbitals p_1, p_2, p_3 for which we have*

$$\underline{L}p_1 = \begin{pmatrix} 0 \\ -ip_3 \\ ip_2 \end{pmatrix}, \quad \underline{L}p_2 = \begin{pmatrix} ip_3 \\ 0 \\ -ip_1 \end{pmatrix}, \quad \textit{and} \quad \underline{L}p_3 = \begin{pmatrix} -ip_2 \\ ip_1 \\ 0 \end{pmatrix}, \tag{4.1}$$

the following identity holds:

$$\begin{aligned} \int_{\mathbb{R}^6} \frac{1}{|x-y|} |p_i(x)|^2 |p_i(y)|^2 dxdy = & \int_{\mathbb{R}^6} \frac{1}{|x-y|} |p_i(x)|^2 |p_j(y)|^2 dxdy \\ & + 2\int_{\mathbb{R}^6} \frac{1}{|x-y|} p_i^*(x) p_j(x) p_i(y) p_j^*(y) dxdy \end{aligned}$$

for all $i \neq j$.

Proof We begin by noting that the $\langle \underline{L}^2 \rangle = 6$, $\langle \underline{S}^2 \rangle = 0$ eigenspace in Table 3.2 is of dimension 5 and hence is non-degenerate. This is true for any choice of p-orbitals satisfying the conditions in (4.1). It follows from Lie Algebra theory that all eigenfunctions within this space have the same expected value for V_{ee}. We choose two representatives, namely

$$\Psi_1 := \frac{1}{\sqrt{2}}\big(|1s\overline{1s}p_1\overline{p_2}\rangle - |1s\overline{1s}\overline{p_1}p_2\rangle\big), \quad \Psi_2 := \frac{1}{\sqrt{2}}\big(|1s\overline{1s}p_1\overline{p_1}\rangle - |1s\overline{1s}p_2\overline{p_2}\rangle\big),$$

and using Slater's rules from Theorem 2.3.1 we find the two expectations to be

$$\langle V_{ee}\rangle_{\Psi_1} = (1s1s|1s1s) + 4(1s1s|p_1p_1) - 2(1sp_1|p_11s) + (p_1p_1|p_2p_2) + (p_1p_2|p_2p_1)$$
$$\langle V_{ee}\rangle_{\Psi_2} = (1s1s|1s1s) + 4(1s1s|p_1p_1) - 2(1sp_1|p_11s) + (p_1p_1|p_1p_1) - (p_1p_2|p_2p_1).$$

Hence, since the expected values are equal, we see that

$$(p_1p_1|p_2p_2) + (p_1p_2|p_2p_1) = (p_1p_1|p_1p_1) - (p_1p_2|p_2p_1)$$

and the result follows by the rotational symmetry of the integrals. □

From this Lemma we see that, in particular, $-(33|33) + (33|44) < 0$ and $-(33|33) + (33|44) + (34|43) < 0$ and so obtain the ordering

$$E_{^3P} < E_{^1D} < E_{^1S},$$

which holds for all $1s$-, $2s$- and $2p$-orbitals regardless of the radial parts. We now have three groups of ordered energies for the uncorrelated wavefunctions:

$$E_{^1S,2s^2}, \quad E_{^3P,2s2p} < E_{^1P,2s2p}, \quad E_{^3P,2p^2} < E_{^1D,2p^2} < E_{^1S,2p^2},$$

but as discussed earlier we cannot determine a complete ordering.

We now wish to introduce correlation, which in Beryllium is only present in the ^1S eigenspace. Diagonalizing the $H = H_0 + V_{ee}$ matrix we see that the two correlated eigenvalues are

$$\lambda_\pm = \tfrac{1}{2}\big(\langle H\rangle_{\Psi_1} + \langle H\rangle_{\Psi_2}\big) \pm \sqrt{\tfrac{1}{4}\big(\langle H\rangle_{\Psi_1} - \langle H\rangle_{\Psi_2}\big)^2 + |\langle\Psi_1|V_{ee}|\Psi_2\rangle|^2},$$

where we have used that evenly excited Slater determinants are not correlated by H_0. We see that this causes the lower eigenvalue to be shifted down and the higher to be shifted up; it is not possible for correlation to cause the eigenvalues to cross each other. It is however possible for correlation to cause crossing of eigenvalues that are not correlated with each other.

It follows that, if the original spectrum has the ordering $E_{^1S,2s^2} < E_{^1S,2p^2}$ then the inequalities above are maintained for the correlated wavefunctions (as there is no inequality involving $E_{^1S,2s^2}$ and $E_{^1S,2p^2}$ is shifted higher). However, this does not necessarily maintain the overall ordering of the spectrum unless the ground state is $E_{^1S,2s^2}$ and the highest excited state is $E_{^1S,2p^2}$. For our particular choice of orbitals this is the case (see Tables 4.8 and 4.17) but there seems to be no *a* *priori* reason why this should be the case for all choices of orbitals.

Furthermore, if we have $E_{{}^1S,2s^2} > E_{{}^1S,2p^2}$ then we can no longer be sure that, when correlation is introduced, we retain the ordering $E_{{}^3P,2p^2} < E_{{}^1D,2p^2} < E_{{}^1S,2p^2}$ as $E_{{}^1S,2p^2}$ is now shifted down whilst $E_{{}^3P,2p^2}$ and $E_{{}^1D,2p^2}$ are unaffected by correlation.

Since correlation seems to cause such difficulties even in the simplest case, we derive only the inequalities for the non-correlated wavefunctions of the remaining atoms.

It is worth noting that these rules are also independent of the value of λ, the coupling coefficient in the perturbation description of the Hamiltonian. Therefore, for uncorrelated wavefunctions, if the states are correctly ordered with respect to the true eigenstates as $\lambda \to 0$ then they are correctly ordered for the physical value $\lambda = 1$. They are also independent of the choice of radial terms, provided that all orbitals are orthogonal.

Theorem 4.3.2 (Rigorous Spectral Orderings)**.** *Let*

$$X_N^Q = \mathit{Span}\{|\phi_{1s}\overline{\phi_{1s}}\psi_1 \dots \psi_N\rangle =: \Psi \mid \psi_1, \dots, \psi_N \in V,\ \hat{Q}\Psi = Q\Psi\}$$

where $\hat{Q} = \sum_{i=1}^3 a^\dagger(\phi_{2p_i})a(\phi_{2p_i}) + a^\dagger(\overline{\phi_{2p_i}})a(\overline{\phi_{2p_i}})$,

$$V = \{\phi_{2s}, \overline{\phi_{2s}}, \phi_{2p_1}, \overline{\phi_{2p_1}}, \phi_{2p_2}, \overline{\phi_{2p_2}}, \phi_{2p_3}, \overline{\phi_{2p_3}}\},$$

and the one-electron orbitals are orthonormal and the spatial parts satisfy

$$\underline{L}^2\phi_{1s} = \underline{L}^2\phi_{2s} = 0, \quad \underline{L}^2\phi_{2p_i} = 2\phi_{2p_i}, \quad L_i\phi_{2p_j} \in \mathit{Span}\{\phi_{2p_1}, \phi_{2p_2}, \phi_{2p_3}\}, \forall i, j.$$

Then the ordering of the energies of the $\underline{L}^2$*-*$\underline{S}^2$ *eigenspaces formed from* X_N^Q *for Lithium-Neon (*$N = 1, \dots, 8$*) satisfy the inequalities in Table 4.20.*

Proof The result follows trivially from the expressions for the energies given in Tables 4.3-4.5, the positivity of the Coulomb and exchange integrals and the identity in Lemma 4.3.1. It should be noted that, although the restriction on the *p*-orbitals in the Theorem is less strict than that in Lemma 4.3.1, the result still holds as the energies are independent of the *p*-shell basis. ☐

Corollary 4.3.3. *In particular, the results of Theorem 4.3.2 hold for the MCSCF GS with V the MCSCF orbitals as well as for MDHF (also known as "angular momentum - spin adapted" HF) with V the HF orbitals. This is true both in the*

	Orbitals	Inequalities
Li	$2s$ $2p$	$E_{^2S}$ $E_{^2P}$
Be	$2s^2$ $2s2p$ $2p^2$	$E_{^1S}$ $E_{^3P} < E_{^1P}$ $E_{^3P} < E_{^1D} < E_{^1S}$
B	$2s^22p$ $2s2p^2$ $2p^3$	$E_{^2P}$ $E_{^4P} < E_{^2D} < E_{^2S}$ $E_{^4P} < E_{^2P}$ $E_{^4S} < E_{^2D} < E_{^2P}$
C	$2s^22p^2$ $2s2p^3$ $2p^4$	$E_{^3P} < E_{^1D} < E_{^1S}$ $E_{^5S} < E_{^3D} < \{E_{^3P}, E_{^1D}\} < E_{^1P}$ $E_{^5S} < E_{^3S}$ $E_{^3P} < E_{^1D} < E_{^1S}$
N	$2s^22p^3$ $2s2p^4$ $2p^5$	$E_{^4S} < E_{^2D} < E_{^2P}$ $E_{^4P} < E_{^2D} < E_{^2S}$ $E_{^4P} < E_{^2P}$ $E_{^2P}$
O	$2s^22p^4$ $2s2p^5$ $2p^6$	$E_{^3P} < E_{^1D} < E_{^1S}$ $E_{^3P} < E_{^1P}$ $E_{^1S}$
F	$2s^22p^5$ $2s2p^6$	$E_{^2S}$ $E_{^2P}$

Table 4.20: Inequalities concerning the ordering of the atomic spectra, derived using only positivity of the exchange integrals and the identity in Lemma 4.3.1. The notation $\{A, B\}$ means that there is insufficient information to order the two states.

limit of a complete basis set as well as for any suitable finite basis set such as STO-3G.

Although the idea of deriving spectral orderings from the V_{ee} matrix elements is not new ([Sla29, Joh32, Uff38]), previous works have considered the spectral orderings in only the cases in which there is a complete ordering. The more interesting cases are not investigated and [Sla29] claimed that the experimental spectral gaps should correspond to these integrals, which, due to minimization over orbitals, is most definitely not true. Further to this, the expressions are derived using a number of results which are not rigorously proven, such as the fact that $\hat{Q}$ commutes with $\underline{L}^2$, $\underline{S}^2$, L_3 and S_3.

Before we discuss these inequalities and their relationship to Hund's rules we first note the obvious symmetry of the energy orderings in Table 4.20 under taking the dual (see Definition 3.1.5). We wish to show that this is a rigorous result, independent of the choice of basis and eigenfunctions. Also, we hope that the proof of such a result will give insight into the important properties that cause this effect.

4.3.1 Dual Invariance of the Spectral Orderings

As well as the ideas of Section 3.1.6 we recall the operators $\hat{P}$, $\hat{Q}_\uparrow$ and $\hat{Q}_\downarrow$ given in Definition 3.3.5, which count the number of pairs of $2p$-orbitals, and the number of spin up / spin down $2p$-orbitals respectively. We also define two new operators that count the number of $2s$ and $\overline{2s}$ orbitals

$$\hat{S}_\uparrow := a^\dagger(2s)a(2s), \quad \hat{S}_\downarrow := a^\dagger(\overline{2s})a(\overline{2s}),$$

the expected values of which can be written in terms of N, $\langle S_3\rangle_{|\Psi\rangle}$, $\langle \hat{Q}_\uparrow\rangle_{|\Psi\rangle}$ and $\langle \hat{Q}_\downarrow\rangle_{|\Psi\rangle}$; their inclusion merely simplifies notation.

The sum of operators $\hat{S} := \hat{S}_\uparrow + \hat{S}_\downarrow$, i.e. the total number of $2s$-orbitals, commutes with $\underline{L}^2$, $\underline{S}^2$, L_3, S_3 and $\hat{R}$. The proofs are straightforward and follow by writing

$$\hat{S} = N - 2 - \hat{Q}$$

and noting that $\hat{Q}$ commutes with each operator.

We define the symbolic expectation of V_{ee} for a wavefunction to be the expectation expressed in Coulomb and exchange integral notation ($(ii|jj)$ and $(ij|ji)$) as in Tables 4.3-4.5.

Theorem 4.3.4 (Dual Invariance of Spectral Orderings). *Let* $\chi_1 = \sum_{i=1}^{M_1} \alpha_i |\Psi_i\rangle$ *and* $\chi_2 = \sum_{i=1}^{M_2} \beta_i |\Phi_i\rangle$*, where* $\alpha_i, \beta_i \in \mathbb{C}$*, and* $|\Psi_i\rangle, |\Phi_i\rangle \in X_N$*, be eigenfunctions* *of* $\underline{L}^2$ *and* $\underline{S}^2$*. Further suppose*

$$\langle\hat{Q}\rangle_{|\Psi_1\rangle} = \langle\hat{Q}\rangle_{|\Psi_i\rangle} = \langle\hat{Q}\rangle_{|\Phi_j\rangle}$$

for all $i = 1, \dots, M_1$*,* $j = 1, \dots, M_2$*. Then for the dual operator given by Definition 3.1.5,* $*\chi_1$ *and* $*\chi_2$ *are also eigenfunctions of* $\underline{L}^2$ *and* $\underline{S}^2$*, with* $\langle\underline{L}^2\rangle_{*\chi_i} = \langle\underline{L}^2\rangle_{\chi_i}$ *and* $\langle\underline{S}^2\rangle_{*\chi_i} = \langle\underline{S}^2\rangle_{\chi_i}$*. Furthermore, the difference in the symbolic expected values of* V_{ee} *of* χ_1 *and* χ_2 *is invariant under* $*$*, i.e.*

$$\langle V_{ee}\rangle_{\chi_1} - \langle V_{ee}\rangle_{\chi_2} = \langle V_{ee}\rangle_{*\chi_1} - \langle V_{ee}\rangle_{*\chi_2}.$$

This does not show that the explicit energy difference is invariant under taking the dual; different orbitals are used for the p and $8-p$ valence electron cases. However, if the sign of the energy difference is well-defined for two eigenfunctions, then it is also well-defined under the dual and has the same sign. As with the duality results for the eigenspaces in Section 3.1.6 this result appears to be novel.

Proof of Theorem 4.3.4 The fact that $*\chi_i$ are eigenfunctions of $\underline{L}^2$ and $\underline{S}^2$ with appropriate eigenvalues is simply the result of Theorem 3.1.10.

By the invariance of the energy under linear combinations of L_3 and S_3 eigenfunctions, we may choose χ_i to be eigenfunctions of S_3 and L_3 with L_3 eigenvalue zero. The restriction on $\hat{Q}$ trivially shows that χ_i are eigenfunctions of $\hat{R}$, the parity. Hence we may use Theorem 3.3.1 and all Slater determinants within each of χ_i and $*\chi_i$ are evenly excited.

In second quantization notation we may write two Slater determinants of length K that differ by m valence orbitals as (up to sign)

$$\begin{aligned}|\Psi_1\rangle &= a^\dagger(\psi_1)\dots a^\dagger(\psi_K)|0\rangle\\ |\Psi_2\rangle &= a^\dagger(\psi_{K+1})\dots a^\dagger(\psi_{K+m})a^\dagger(\psi_{m+1})\dots a^\dagger(\psi_K)|0\rangle\end{aligned}$$

where wlog we have chosen to replace the first m orbitals in the Slater determinant and $|0\rangle = |1s\overline{1s}\rangle$. It is clear that

$$|\Psi_2\rangle = a^\dagger(\psi_{K+1})\dots a^\dagger(\psi_{K+m})a(\psi_m)\dots a(\psi_1)|\Psi_1\rangle. \tag{4.2}$$

From the definition of $*$, we have

$$\begin{aligned}*|\Psi_1\rangle &= a(\psi_K)\dots a(\psi_1)|1\rangle\\ *|\Psi_2\rangle &= a(\psi_K)\dots a(\psi_{m+1})a(\psi_{K+m})\dots a(\psi_{K+1})|1\rangle.\end{aligned}$$

We now note that $a^\dagger(\psi_1)\dots a^\dagger(\psi_m)a(\psi_m)\dots a(\psi_1)|1\rangle = |1\rangle$, from which we have

$$\begin{aligned}
*|\Psi_2\rangle &= a(\psi_K)\dots a(\psi_{m+1})a(\psi_{K+m})\dots a(\psi_{K+1}) \\
&\qquad\qquad a^\dagger(\psi_1)\dots a^\dagger(\psi_m)a(\psi_m)\dots a(\psi_1)|1\rangle \\
&= (-1)^{2(K-m)m} a(\psi_{K+m})\dots a(\psi_{K+1})a^\dagger(\psi_1)\dots a^\dagger(\psi_m) \\
&\qquad\qquad a(\psi_K)\dots a(\psi_{m+1})a(\psi_m)\dots a(\psi_1)|1\rangle \\
&= a(\psi_{K+m})\dots a(\psi_{K+1})a^\dagger(\psi_1)\dots a^\dagger(\psi_m)\,(*|\Psi_1\rangle)\,. \qquad (4.3)
\end{aligned}$$

From (4.2) and (4.3) we see that $*|\Psi_1\rangle$ and $*|\Psi_2\rangle$ differ by the same orbitals as $|\Psi_1\rangle$ and $|\Psi_2\rangle$.

We are now in a position to compute the energy differences. Firstly we see that

$$*\chi_1 = \sum_{i=1}^{M_1} \alpha_i^*\big(*|\Psi_i\rangle\big), \quad *\chi_2 = \sum_{i=1}^{M_2} \beta_i^*\big(*|\Phi_i\rangle\big)$$

and hence the four expected values we are interested in are

$$\begin{aligned}
\langle V_{ee}\rangle_{\chi_1} &= \sum_{i,j=1}^{M_1} \alpha_i^*\alpha_j\langle\Psi_j|V_{ee}|\Psi_i\rangle, \quad \langle V_{ee}\rangle_{\chi_2} = \sum_{i,j=1}^{M_2} \beta_i^*\beta_j\langle\Phi_j|V_{ee}|\Phi_i\rangle, \\
\langle V_{ee}\rangle_{*\chi_1} &= \sum_{i,j=1}^{M_1} \alpha_i\alpha_j^*\langle *\Psi_j|V_{ee}|*\Psi_i\rangle = \sum_{i,j=1}^{M_1} \alpha_i^*\alpha_j\langle *\Psi_j|V_{ee}|*\Psi_i\rangle, \\
\langle V_{ee}\rangle_{*\chi_2} &= \sum_{i,j=1}^{M_2} \beta_i\beta_j^*\langle *\Phi_j|V_{ee}|*\Phi_i\rangle = \sum_{i,j=1}^{M_2} \beta_i^*\beta_j\langle *\Phi_j|V_{ee}|*\Phi_i\rangle,
\end{aligned}$$

where we have used that V_{ee} is self adjoint and real, and all the Slater determinants are real. We wish to show that

$$\langle V_{ee}\rangle_{\chi_1} - \langle V_{ee}\rangle_{\chi_2} = \langle V_{ee}\rangle_{*\chi_1} - \langle V_{ee}\rangle_{*\chi_2},$$

which is equivalent to

$$\langle V_{ee}\rangle_{\chi_1} - \langle V_{ee}\rangle_{*\chi_1} = \langle V_{ee}\rangle_{\chi_2} - \langle V_{ee}\rangle_{*\chi_2},$$

i.e. that the difference in expected values of V_{ee} between a wavefunction and its dual depend only upon the number of $2s$- and $2p$-orbitals, which are the only conserved quantities between χ_1 and χ_2.

Consider first

$$\langle \chi_1 | V_{ee} | \chi_1 \rangle - \langle *\chi_1 | V_{ee} | *\chi_1 \rangle = \sum_{i,j=1}^{M_1} \alpha_i \alpha_j^* \big(\langle \Psi_j | V_{ee} | \Psi_i \rangle - \langle *\Psi_j | V_{ee} | *\Psi_i \rangle \big) \qquad (4.4)$$

and note that, by Slater's rules from Theorem 2.3.1, we need only consider terms that differ by zero or two orbitals. We begin with the case of two Slater determinants $|\Psi_1\rangle$ and $|\Psi_2\rangle$ that differ by two orbitals. As noted above, the duals differ by the same pairs of orbitals, with the same sign and hence $\langle \Psi_1 | V_{ee} | \Psi_2 \rangle = \langle *\Psi_1 | V_{ee} | *\Psi_2 \rangle$. Therefore the cross terms cancel and (4.4) becomes

$$\langle \chi_1 | V_{ee} | \chi_1 \rangle - \langle *\chi_1 | V_{ee} | *\chi_1 \rangle = \sum_{i=1}^{M_1} |\alpha_i|^2 \big(\langle \Psi_i | V_{ee} | \Psi_i \rangle - \langle *\Psi_i | V_{ee} | *\Psi_i \rangle \big). \qquad (4.5)$$

For a single Slater determinant we may define the following quantities:

$$P := \langle \hat{P} \rangle_{|\Psi\rangle}, \ \ Q_\uparrow := \langle \hat{Q}_\uparrow \rangle_{|\Psi\rangle}, \ \ Q_\downarrow := \langle \hat{Q}_\downarrow \rangle_{|\Psi\rangle}, \ \ S_\uparrow := \langle \hat{S}_\uparrow \rangle_{|\Psi\rangle}, \ \ S_\downarrow := \langle \hat{S}_\downarrow \rangle_{|\Psi\rangle},$$

(along with the sums $\hat{Q} := \hat{Q}_\uparrow + \hat{Q}_\downarrow$ and $\hat{S} := \hat{S}_\uparrow + \hat{S}_\downarrow$). Further, by the restriction on $\hat{Q}$, Corollary 3.3.15 shows that each Slater determinant in χ_1 must have the same value of P, as must each in χ_2.

Now note that the terms involving $1s$ and $\overline{1s}$ depend only on $S := S_\uparrow + S_\downarrow$ and $Q := Q_\uparrow + Q_\downarrow$. This is because each valence electron will form one Coulomb term with each of $1s$ and $\overline{1s}$ and exactly one exchange term with either $1s$ or $\overline{1s}$. Furthermore, by the rotational invariance of the $1s$ spatial orbital, $(1s1s|2p_i2p_i)$ and $(1s2p_i|2p_i1s)$ are independent of i. Denote the sum of these terms by $f(S,Q)$.

Using the expansion from Corollary 2.3.2 we obtain

$$\begin{aligned} \langle V_{ee} \rangle_{|\Psi\rangle} &= f(S,Q) + SQ(22|33) - (S_\uparrow Q_\uparrow + S_\downarrow Q_\downarrow)(23|32) + P(33|33) \\ &\quad + \left(\binom{Q}{2} - P \right)(33|44) - \left(\binom{Q_\uparrow}{2} + \binom{Q_\downarrow}{2} \right)(34|43). \end{aligned}$$

Denoting eigenvalues for the dual by the addition of a tilde, we have

$$\tilde{S}_\uparrow = 1 - S_\uparrow, \quad \tilde{S}_\downarrow = 1 - S_\downarrow, \quad \tilde{Q}_\uparrow = 3 - Q_\uparrow, \quad \tilde{Q}_\downarrow = 3 - Q_\downarrow,$$

$$\text{and } \tilde{P} = 3 + P - (Q_\uparrow + Q_\downarrow),$$

where the first four identities follow trivially from the definition of $*$. The identity for $\tilde{P}$ follows by considering that there are a possible 3 pairs of $2p$-orbitals and the total number of different spatial $2p$-orbitals occupied in $|\Psi\rangle$ is given by $Q_\uparrow + Q_\downarrow - P$, none of which can form pairs in the dual.

It now follows that

$$\tilde{S} = 2 - S, \quad \tilde{Q} = 6 - Q, \text{ and}$$

$$\tilde{S}_\uparrow \tilde{Q}_\uparrow + \tilde{S}_\downarrow \tilde{Q}_\downarrow = 6 - 3S - Q + (S_\uparrow Q_\uparrow + S_\downarrow Q_\downarrow).$$

Thus in order to show that $\langle V_{ee} \rangle_{|\Psi\rangle} - \langle V_{ee} \rangle_{*|\Psi\rangle}$ depends only on S and Q it remains to consider the binomial coefficients in the above. We have

$$\binom{Q}{2} - \binom{\tilde{Q}}{2} = \binom{Q}{2} - \binom{6-Q}{2},$$

and

$$\begin{aligned} &\binom{Q_\uparrow}{2} + \binom{Q_\downarrow}{2} - \binom{\tilde{Q}_\uparrow}{2} - \binom{\tilde{Q}_\downarrow}{2} = \binom{Q_\uparrow}{2} + \binom{Q_\downarrow}{2} - \binom{3-Q_\uparrow}{2} - \binom{3-Q_\downarrow}{2} \\ &\quad = \tfrac{1}{2}\left[Q_\uparrow(Q_\uparrow - 1) + Q_\downarrow(Q_\downarrow - 1) - (3-Q_\uparrow)(2-Q_\uparrow) - (3-Q_\downarrow)(2-Q_\downarrow)\right] \\ &\quad = -6 + 2Q. \end{aligned}$$

It is therefore clear that all coefficients in $\langle V_{ee} \rangle_{|\Psi\rangle} - \langle V_{ee} \rangle_{*|\Psi\rangle}$ depend only upon the two quantities S and Q and hence only on N and Q, i.e. each of the differences in the sum of (4.5) are independent of i. The result follows from the fact that $\sum_{i=1}^{M_1} |\alpha_i|^2 = 1$. □

Note that the result is independent of the choice of basis for the $2p$-orbitals, we have simply chosen a basis whose symmetry allows easy computation.

4.3.2 Comparison with the original Hund's Rules

We now return to a comparison of Hund's rules and our results. The first thing of note is that, with the addition of shell ordering, Hund's rules provides a complete ordering of the spectrum (order by number of p-orbitals, then by total spin and then by total angular momentum), whereas our results are not a total ordering.

However, in the cases where we do have an explicit ordering our results agree with Hund's rules. Although this is interesting in itself, the main interest comes from the cases where the rigorous orderings are less restrictive than those of Hund. As an example, Hund's rules would order the $2s2p^2$ states of Boron as $E_{^4P} < E_{^2D} < E_{^2P} < E_{^2S}$ whereas our inequalities could allow any of $E_{^4P} < E_{^2D} < E_{^2P} < E_{^2S}$, $E_{^4P} < E_{^2P} < E_{^2D} < E_{^2S}$ and $E_{^4P} < E_{^2D} < E_{^2S} < E_{^2P}$.

We now consider the states of Carbon corresponding to $2s2p^3$ and in particular the first ^3S and second ^1D states in Table 4.17. Hund's rules would predict that the ^3S state should be lower lying that the ^1D state whereas our inequalities do not allow us to order these two states. As can be seen from the experimental energies in the table, the physical ordering actually disagrees with Hund's rules but is allowed by our ordering. This is an important example as it justifies the use of the inequalities in Table 4.20 for true angular momentum and spin eigenstates, rather than the original Hund's rules.

Moreover, the fact that the calculated ^5S level and first ^1S levels of Carbon swap between Tables 4.8 and 4.17 (i.e. for the non-variational and variational basis sets respectively) shows that there is no basis-independent total ordering of the spectrum. In particular it shows that there probably does not exist an *a priori* ordering of terms involving different numbers of $2s$- and $2p$-orbitals.

However, the experimental ^1D and ^3P $2p^2$ states of Beryllium are not in agreement with Hund's rules nor our inequalities (with the ^1D state lower than the ^3P state). This suggests that, in a minority of cases, correlation between states containing different configurations is important in predicting the spectral ordering.

4.3.3 Counterexample to the Intuition of Hund's Rules

The common intuitive explanation of Hund's rules (see e.g. [Sch01]), at least for simple cases like the Beryllium $2s2p$ singlet and triplet, is that the aligned spins in the triplet case cause the electrons to be, on average, further apart and thus reduce the inter-electron repulsion. However, using the Z_i values from Table 4.14 we may calculate the kinetic, nuclear potential, and electron-repulsion terms for each of the singlet and triplet cases. These values are given in Table 4.21 and it is clear that, whilst the ordering is correct, it is not actually due to a smaller V_{ee} term, a result that has also been noted for more complex CI calculations, e.g. [TT74]. This suggests that the true physical explanation of Hund's rules should not simply involve the electron repulsion term but should also describe how this term is interlinked with the kinetic and nuclear potential terms.

	^1P	^3P
KE+PE	-18.4535	-18.9084
V_{ee}	4.0847	4.4261
Total	-14.3688	-14.4823

Table 4.21: Kinetic, nuclear-electron potential and electron-electron potential energy splitting of the ^1P and ^3P $2s2p$ terms of Beryllium.

Chapter 5

Other Atomic Properties

5.1 Atomic Ground State Dimensions

Along with the predicted ground state energies, our method also gives the dimension of the ground state eigenspace, simply by noting that it is the minimal dimension of the relevant space, as given by d in Table 3.8. The resulting dimensions are given in Table 5.1.

These dimensions are interesting as they are a measure of the 'flexibility' within the ground state, i.e. the number of degrees of freedom (or parameters) which can be varied without affecting the energy of the state.

On a qualitative level, we expect that an atom with a high-dimensional ground state will form a wider range of molecules than an atom with a similar number of valence (bonding) electrons but with a lower dimensional ground state. This should be true both in terms of molecular geometry (e.g. linear, bent, tetrahedral) as well as in terms of which atoms it will bond with.

These two ideas are interlinked, in that an atom that can form molecules with a wider range of geometries should also be able to bond with a wider range of types of atom; the other atoms in the molecule will have a greater freedom to form other bonds in the molecule, and also be able to form geometries that have lower nucleus-nucleus repulsion terms.

A more quantitative example will be given in Chapter 6 where we first see that

Atom	Li	Be	B	C	N	O	F	Ne
GS dimension	2	1	6	9	4	9	2	1

Table 5.1: Perturbation theory ground state dimensions of Li-Ne.

the pair density is independent of the choice of S_3 state (see Lemma 6.2.1) and then show that if there are spatial degrees of freedom, such as in the biologically important cases of Carbon and Oxygen, the pair density has a free parameter that can be varied without affecting the energy of the atom. This is not true of other atoms such as Nitrogen. Considering the pair density to be a measure of how likely it is for two electrons to bond simultaneously at two different points, we see that this parameter results in a range of bond angles for a given bond length.

5.2 Electron Densities and Atomic Radii

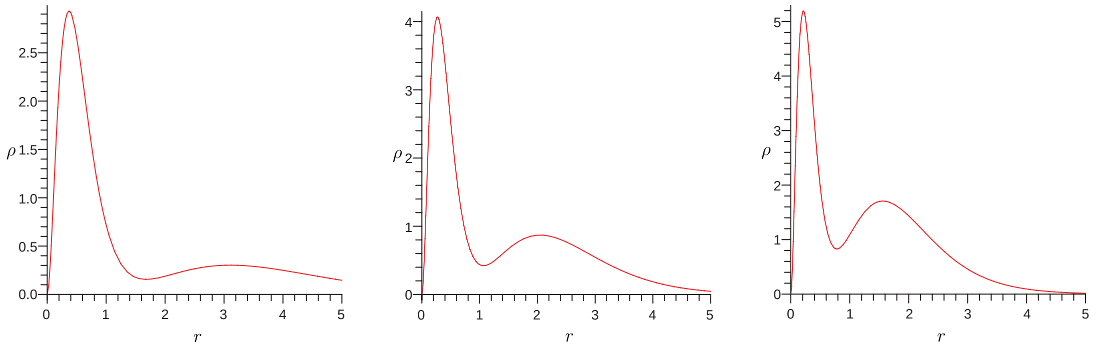
One major advantage of our method compared to numerical methods such as numerical HF and DFT is that we obtain explicit analytic wavefunctions, as compared to ones determined only at a finite number of points. One use of this concerns a common plot in the literature, the radial electron densities of atoms, which are given in [FF77] for the HF case.

We define the radial distribution of a wavefunction to be the electron density ρ normalized such that $\int_0^\infty \rho r^2 dr = 1$ (i.e. integrated over the angular part). Figure 5.1 shows the radial distributions for Lithium to Neon, in which the two 'shells' of electrons are immediately visible. It is also obvious that, despite the differing exponential decay rates of the $2s$- and $2p$-orbitals (see Tables 4.14 and 4.15), we do not see two distinct shells for these two sets of orbitals.

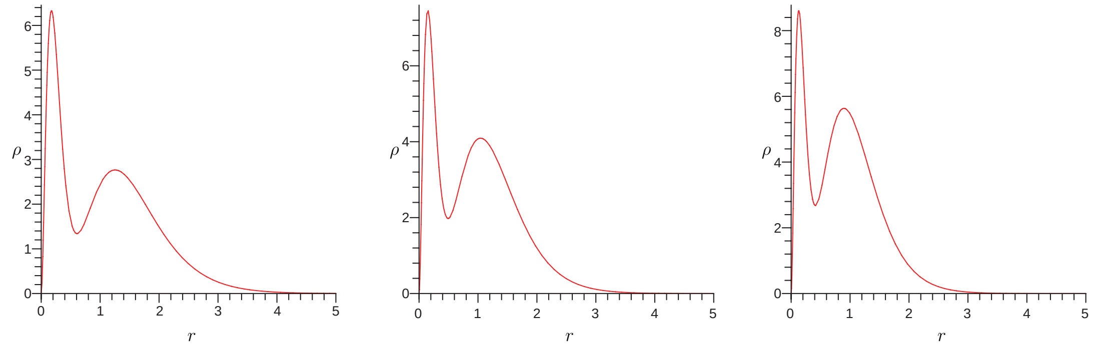
Figure 5.1 also shows the contraction of the size of an atom across the second row of the periodic table. This is a well known effect that is attributed to the greater nuclear charge causing the electrons to be more attracted to the nucleus. Another piece of data linked to this idea is the position of the outer maximum of the radial distribution. Intuitively, this is a good measure of the 'size' of the atom and should also be a good first approximation to the length of bonds formed between atoms. Of course, there must be some flexibility in these estimates to allow for hybridization of orbitals and other bonding effects.

A number of authors including Slater (see [Sla64]) attempted to derive empirical atomic radii which could be summed to predict covalent and crystal bond lengths. These values are in reasonable agreement with the position of the outer maximum of the radial distribution, as shown in Table 5.2. However, the error when assuming purely additive bond lengths can be large. Also shown are the maximum of the radial distributions for the SCF calculations in [CRR67].

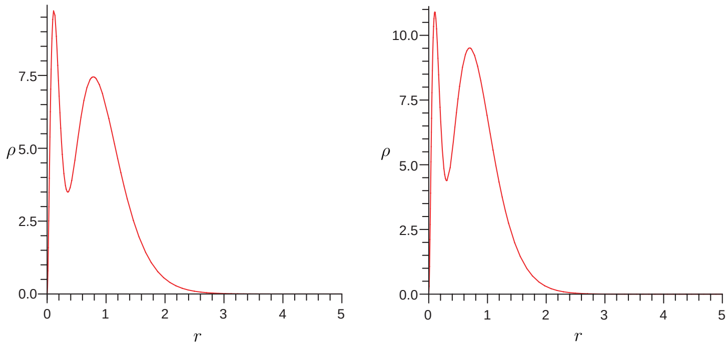
We give the data both in Angstroms and in atomic units as the SCF data is in Angstoms to only 2d.p. and hence conversion is inaccurate. The agreement is in general very good, although we find our predicted maxima to be slightly lower



Lithium Beryllium Boron



Carbon Nitrogen Oxygen



Fluorine Neon

Figure 5.1: Radial distributions for the ground states of Li-Ne.

than the SCF data. The contraction of atoms across the second row can be seen clearly in both sets of data.

	Li	Be	B	C	N	O	F	Ne
r_{max} (Å)	1.65	1.08	0.83	0.66	0.55	0.47	0.41	0.37
r^{SCF}_{max} (Å)	1.67	1.12	0.87	0.67	0.56	0.48	0.42	0.38
r^{Slater} (Å)	1.45	1.05	0.85	0.70	0.65	0.60	0.50	
r_{max} (au)	3.110	2.047	1.562	1.254	1.044	0.897	0.784	0.697
r^{SCF}_{max} (au)	3.157	2.117	1.645	1.267	1.059	0.907	0.794	0.718
r^{Slater} (au)	2.741	1.985	1.607	1.323	1.229	1.134	0.945	

Table 5.2: The position of the outer maximum of the electron density for our data (r_{max}) and the Self Consistent Field method of [CRR67]. Also shown are Slater's empirical values for atomic radii [Sla64].

How these maxima correspond non-empirically to the experimental bond lengths is an interesting question. From calculations on simple molecules such as H_2 (e.g. [Sch01]) it is obvious that molecular bonding is related to an increase in the electron density between the atoms. This suggests that we should, in some sense, maximize the overlap of the electron densities. However, the best approach would be to perform a full multi-centre calculation to obtain the molecular bond lengths. These should then be compared to the atomic radii obtained by the same method of calculation in order to determine any non-empirical relation.

It is easy to demonstrate that, in general, it is not possible to find a set of atomic radii from which bond lengths can be constructed through addition of these values, i.e. $r_{\text{X-Y}} = r_{\text{X}} + r_{\text{Y}}$. Due to symmetry, one would expect that the atomic radii must be defined to be almost exactly half the X-X bond length. Whilst this gives good results in some cases, it gives relatively poor agreement with experiment in other cases.

In the following examples, all data are taken from [NIS05] and are given in atomic units. The experimental bond length of H-H in H_2 is 1.401, F-F in F_2 is 2.668 and the average length of a C-C bond is 2.836. This gives atomic radii of approximately 0.7 for H, 1.34 for F and 1.42 for C. This would predict a C-H bond length of 2.12 compared to the experimental average of 2.06. For H-F the predicted value is 2.04 compared to the experimental value of 1.733 in HF.

It is clear from these examples that any best-fit data set for atomic radii will cause some predicted bond lengths to be too short and some too long. This demonstrates the importance of hybridization in the formation of molecules.

5.3 Ionization Energies

Along with the spectral gaps, another well known physical property our data can be used to predict are the ionization energies of atoms. We define $I(Z) = E(Z, N-1) - E(Z, N)$ where $E(Z, N)$ is the ground state energy of an atom/ion with nuclear charge Z and N electrons. From analogous calculations with methods such as HF we suspect that the results will depend strongly on the model and thus give us a good insight into possible improvements. For example, Koopmans' theorem [SO96] uses a 'frozen orbital' approach where the ionization energy is taken to be the energy of the orbital in the highest energy state. Since this does not allow a rearrangement of the electrons in the excited state it generally causes the calculated $E(Z, N-1)$ to be too large, and thus tends to produce ionization energies that are too positive.

This problem should not be as noticeable with our method as we also variationally minimize the energy of the ionized state. We must also minimize over the $\underline{L}^2$-$\underline{S}^2$ eigenspaces but, as expected, this minimizer always occurs for the state that is the minimizer for the atom with the same number of electrons as the ion.

We expect the relative error for the ionized states to be smaller than that for the atomic states, as the basis will allow a better approximation to the true minimizer. This is due to the increased nuclear charge relative to number of electrons (Z/N), which will favour one-electron states with lower principle quantum number n. This difference in quality of the minimizers for different N with fixed Z could cause a relatively large error in the ionization energies, and we would predict that the ionization energies should be too small when compared to experiment.

Table 5.3 shows the ground state energies for the atom (obtained for experiment by summing all ionization energies) and the first positive ion, along with the corresponding ionization energy, both for our variational calculations and for experiment [Huh93].

It is clear that the ratio of ionization energy and ground state energy decreases rapidly as Z increases. In addition, the theoretical energies of the ions are better than those of the atoms, which is what we predicted. However, the theoretical ionization energies, in general, become much worse as Z increases.

Neither the atomic nor ionic energies themselves are poor, and hence this must follow from accumulation of errors when subtracting the two energies. We propose that this is due to the variational part of our procedure, which causes the ionic energies to be more accurate than the atomic ones.

Given two energies $0 > E^+ > E^-$ and relative errors ΔE^+ and ΔE^- giving two new energies $\tilde{E}^+ := E^+ + \Delta E^+ |E^+|$, $\tilde{E}^- := E^- + \Delta E^- |E^-|$, we have that

	Experimental			Theoretical		
	$E(Z,N)$	$E(Z,N-1)$	$I(Z)$ % of $E(Z,N)$	$E(Z,N)$ % error	$E(Z,N-1)$ % error	$I(Z)$ % error
He	-2.9036	-2.0000	0.9036 31.12	-2.8477 1.93	-2.0000 0.00	0.8477 6.19
Li	-7.4779	-7.2799	0.1980 2.65	-7.4139 0.86	-7.2227 0.79	0.1912 3.43
Be	-14.6684	-14.3258	0.3426 2.33	-14.5795 0.61	-14.2558 0.49	0.3237 5.51
B	-24.6581	-24.3532	0.3049 1.24	-24.4885 0.69	-24.2539 0.41	0.2346 23.06
C	-37.8558	-37.4420	0.4138 1.09	-37.5689 0.76	-37.2542 0.50	0.3124 23.95
N	-54.6177	-54.0836	0.5341 0.98	-54.1597 0.84	-53.7637 0.59	0.3960 25.86
O	-75.1080	-74.6080	0.5000 0.67	-74.3931 0.95	-74.1223 0.65	0.2708 45.84
F	-99.8060	-99.1658	0.6402 0.64	-98.8038 1.00	-98.4080 0.76	0.3958 38.18
Ne	-129.0500	-128.2575	0.7925 0.61	-127.5690 1.15	-127.2195 0.81	0.3495 55.90

Table 5.3: Experimental [Huh93] and Theoretical energies for atoms and their first positive ions, along with the first ionization energies. (% error is the percentage of the experimental energy not captured by the theoretical result).

the relative error in the energy differences is given by

$$\begin{aligned}\frac{\tilde{E}^+ - \tilde{E}^-}{E^+ - E^-} &= \frac{(E^+ + \Delta E^+|E^+|) - (E^- + \Delta E^-|E^-|)}{E^+ - E^-} \\ &= 1 + \frac{\Delta E^+|E^+|}{E^+ - E^-} - \frac{\Delta E^-|E^-|}{E^+ - E^-} \\ &= 1 + \frac{\Delta E^+(|E^+| - |E^-|) + |E^-|(\Delta E^+ - \Delta E^-)}{E^+ - E^-} \\ &= 1 - \Delta E^+ + \frac{|E^-|}{E^+ - E^-}(\Delta E^+ - \Delta E^-).\end{aligned}$$

Since ΔE^+ is small, the error in the ionization energy is large if the ratio of the difference in relative errors $(\Delta E^+ - \Delta E^-)$, and the experimental ionization energy as a proportion of the experimental ground state energy $((E^+ - E^-)/E^-)$ is large. This agrees well with the values in Table 5.3.

Figure 5.2 clearly demonstrates that, as already noted, all the theoretical ionization energies are too low. The Neon result can further be explained by the fact that the ground state has only one choice of Slater determinant and the only improvement on the energy comes from the variational procedure whereas the Ne$^+$ calculation minimizes over an 8-dimensional ground state before variation so we would expect a better energy and thus a large error in the ionization energy. It is also clear that, with the exception of Neon, the qualitative prediction for the ionization energies is very good when compared to experimental data.

We can demonstrate the dependence of $I(Z)$ on the relative quality of the basis set for atoms and ions in a much more dramatic way by considering the non-variational results. As noted in Section 4.1, the qualitative non-variational spectral gaps are quite good (they are all roughly the same factor larger than the experimental energy gaps), but if our explanation holds then we would expect the non-variational ionization energies to be very poor. Scaling the orbitals has more of an effect on the valence orbitals so in the non-variational model these should be relatively poor. Hence, when we remove one in ionization, there will be one fewer poorly approximated orbital and the relative energy error should be smaller than in the atomic case.

As can clearly be seen from Table 5.3 the ionization energies for the non-variational method are very poor and in fact, for Li-O predict that the ion has a lower energy than the atom. This clearly shows how important the variational procedure is in producing accurate ground and excited state wavefunctions.

As a final demonstration of the possibilities of the model we calculate the 10th ionization of Chlorine, which experimentally is 43962 kJ mol^{-1} or 16.747au

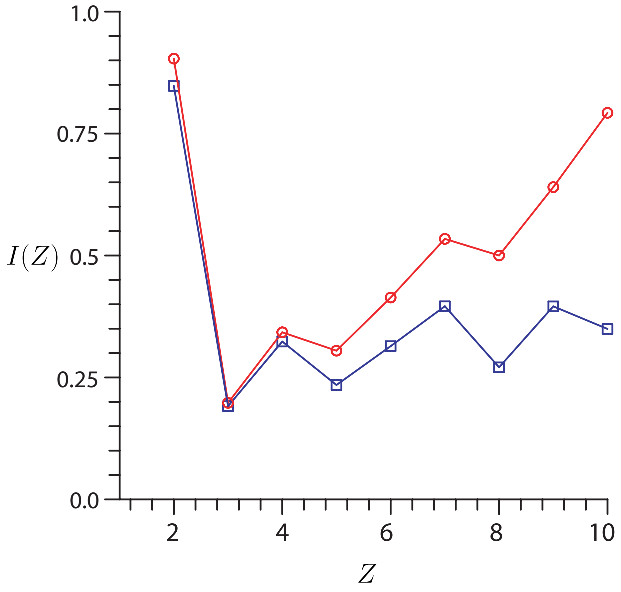


Figure 5.2: Experimental (red) and three-parameter variational (blue) Ionization Energies.

Atom	$E(Z,N)$	$E(Z,N-1)$	$I(Z)$
Li	-7.0566	-7.125	-0.0684
Be	-13.7629	-13.9088	-0.1459
B	-22.7374	-23.4536	-0.7162
C	-34.4468	-35.5348	-1.0880
N	-49.1503	-50.6879	-1.5376
O	-66.7048	-69.1717	-2.4669
F	-99.8060	-90.7929	9.0131
Ne	-127.5690	-116.2621	11.3069

Table 5.4: Non-variational Atomic and Ionic energies, along with Ionization Energies.

[Huh93]. We calculate the energies of Cl^{9+} and Cl^{10+} to be -416.8777au and -400.3836au, which gives an ionization energy of 16.4941au, within 1.5% of the experimental value. These values are calculated in the same way as the Oxygen ionization energy except the nuclear charge is now 17. This helps to justify the idea that method becomes more accurate when ionization causes a smaller change in the ratio Z/N.

From this investigation we see that determination of accurate ionization energies requires different qualities of a basis set than those necessary to accurately predict ground state energies. For ground state energies, the basis must model the ground state wavefunction as accurately as possible. When used for ionization energies, this accuracy of the ground state is not sufficient unless the basis set models the ionic wavefunction with very similar accuracy. If this is not the case then the errors will accumulate to give poor ionization energies. Also, a method could give very accurate ionization energies whilst at the same time giving very poor total energies for the atom and ion.

These results suggest that our basis choice is more accurate when there is more nuclear charge per electron, and it is unclear at present how best to extend or modify our basis so that it is equally accurate irrespective of the ratio of nuclear charge and number of electrons.

5.4 Conclusions and Open Problems

We have seen that a minimal basis of one-electron orbitals, inspired by perturbation theory and physical arguments regarding the screening of electrons, can be chosen such that the Hamiltonian matrix consists of purely rational functions of the screening parameters (Chapters 2 and 3). Further, this matrix has a number of interesting structural properties, as described in Chapter 3. These results include a particle-hole duality of eigenfunctions (Theorem 3.1.10), as well as a restriction on the excitation between Slater determinants in the same spin-angular-momentum-parity eigenspaces (Theorem 3.3.1).

In Section 4.2 we saw that using three variationally determined screening parameters leads to accurate ground state and excited energy levels. This contrasts to the use of zero or one variational parameters (Sections 4.1 and 4.2) which, whilst giving relatively accurate ground state energies, give very poor spectral gaps.

In all cases the agreement of the ordering of the spectra with the experimental data is excellent. We have shown (Theorem 4.3.2) that, to some extent, the ordering of the spectrum is determined by the choice of one-electron orbitals and

Slater determinants, irrespective of the exact form of the orbitals. We have also shown that the duality result for eigenfunctions extends to the symbolic energy differences (Theorem 4.3.4).

Finally (in Chapter 5) we have investigated a number of other atomic properties, including the dimension of the ground state, the form of the radial electron density, and the first ionization energies. In particular, the calculations of the ionization energies give insight as to a possible mathematical justification for the screening parameters.

There are a number of obvious open problems or extensions to these results, possibly the most important of which would be to understand why the screening constants (variational parameters of Section 4.2) are so effective. In particular, we would hope that they arise naturally in an order expansion of a suitably scaled problem. From the existing results, and in particular the discussion of the ionization energies in Section 5.3, we expect this limit to be one of large Z for a fixed number of electrons.

We would further expect that the biggest gain in order should appear when introducing two parameters (one for each principal quantum number shell). This is expected due to the relative lack of improvement in the spectral gaps after introducing one parameter and the similarities of the calculations with two or three parameters.

It would also be interesting to extend these results to larger atoms, especially as this would allow us to investigate the 'filling order' of the $3d$- and $4s$-orbitals. As noted previously, the drawback to these calculations is the rapidly increasing dimension of the problem. For such problems it seems sensible to automate some aspects of the calculations, such as the formation of eigenfunctions and calculation of symbolic energies.

The method could also be extended to small molecules, such as dimers. Whilst $\underline{L}^2$ is no longer a well-defined quantum number we may use an analogous method to form the spin-L_3-parity eigenspaces and continue from there. The difficulty in these cases (assuming that the core shells are kept filled, keeping the dimensions relatively small) is the calculation of multi-center integrals, which in general will not be anywhere near as simple as in the atomic case. For example, the Hydrogen calculation of Sugiura [Sug27] results in expressions containing the exponential integral.

Chapter 6

Pair Densities and Bond Angles

6.1 Introduction to Bonding Theory

We begin by reviewing a number of methods for determining bond angles in relatively simple molecules such as AH_2, where A is an atom from the second period. These methods can be split into two types: those that aim to predict the physical values to a high level of accuracy, and those that aim to give a physical understanding of molecular bonding but tend to give less quantitatively accurate results.

6.1.1 Numerical Methods

Methods of the first type include HF, SCF, and DFT, which are described briefly in Chapter 1. The results for small molecules are in very good agreement with experiment. As an example we consider the water molecule, which is frequently studied due to its abundance, chemical and biological significance, and its unusual bond angle of 104.5°. The best DFT results for the bond angle are accurate to within a tenth of a degree [KSM99], those using HF/SCF methods within a degree or so [Bel78, SO96], and those using methods that include almost all correlation, such as MPPT or CI, to within one tenth of a degree [NIS05]. The relative poorness of SCF calculations can be attributed to the lack of correlation terms which produces a physically inaccurate wavefunction.

These methods (other than DFT) commonly use very large basis sets and, although the numerical results are generally impressive, they give little intuition as to which orbitals are important for bonding and how they are deformed between the atomic and molecular cases. All of the above methods may use molecular orbitals, which are formed from linear combinations of atomic orbitals centred on each of the atoms. This method is easily demonstrated for the case of H_2, see e.g. [Sch01].

One major drawback of these methods is the difficulty involved in computing the integrals for orbitals centred on different atoms, especially when physically realistic Slater-type orbitals are used (cf. the relatively simple case of two Hydrogen $1s$-orbitals in [Sug27]). This problem can be partially overcome by using Gaussian orbitals which simplify the integrals due to the fact that the product of two Gaussians centred at different points is again a Gaussian centred at a third point. The disadvantage of using Gaussians is that, in order to obtain an accurate approximation to the physical wavefunction, many more basis functions are needed than when using Slater type orbitals.

We now move onto the methods that aim to give a more intuitive understanding of molecular bonding, often at the expense of accuracy of numerical predictions. Three of the main methods used are Hybrid orbitals, Valence Shell Electron Pair Repulsion theory (VSEPR) and Walsh diagrams.

6.1.2 Hybrid Orbitals

The use of hybrid orbitals was first suggested by Pauling in the 1930's [Pau31] and is meant to explain bonding in simple molecules. Hybrid orbitals are one-electron orbitals that are linear combinations of the standard atomic one-electron orbitals of one atom. This method can lead to inaccurate predictions as it only takes into account the central atom. For example, H_2O has a bond angle of 104.5° whereas Li_2O is a straight molecule [BSKW63, BB01] but the hybrid orbital method cannot differentiate between these two cases. However, hybrid orbitals do give a good description of a large number of molecules where covalent bonding is important, failing in cases where the bonds are more ionic in character as in Li_2O.

Pauling [Pau31] made three simple physically-motivated assumptions:

1. The electron-pair bond is formed through the interaction of an unpaired electron on each of two atoms.

2. The spins of the electrons are opposed when the bond is formed, so that they cannot contribute to the paramagnetic susceptibility of the substance.

3. Two electrons which form a shared pair cannot take part in forming additional pairs.

Note that point 1 implies that bonding only involves valence electrons.

Combining these assumptions with a physical example demonstrates the necessity of a hybrid-orbital-like explanation for bonding. Consider the ground state of the Carbon atom which, in HF theory, is given by $|1s\overline{1s}2s\overline{2s}p_1p_2\rangle$ and hence by the above can form only two bonds in a molecule. However physically we find that CH_2 is highly unstable whereas CH_4 is stable. Furthermore, the bonds in CH_4 are all of the same length, have the same breaking energy, and the molecule is symmetric, which suggests that the orbitals forming the bonds are also symmetric.

Thus we need to do two things in order to intuitively explain the symmetry and stability of CH_4. Firstly we must allow at least one orbital in the ground state of carbon to be promoted to obtain $|1s\overline{1s}2sp_1p_2p_3\rangle$, which may now form four bonds (as there are now four unpaired electrons). Secondly, in order to obtain symmetry of the bonds, we must take linear combinations of the $2s$ and $2p$ orbitals (described as sp^3 hybrids), which allow all four bonds to be of the same type .

We now consider a few examples which demonstrate the effectiveness of this method. Working across the second period we begin with BeH_2, the HF ground state of the Beryllium atom being $|1s\overline{1s}2s\overline{2s}\rangle$. This is then excited to the state $|1s\overline{1s}2s2p\rangle$ (to provide two bonding orbitals) and forms hybrid orbitals to give a wavefunction of the form $|1s\overline{1s}(sp_+)(sp_-)\rangle$ where $sp_\pm := \frac{1}{\sqrt{2}}(2s \pm 2p)$. The two sp orbitals point in opposite directions and hence the resulting molecule is linear.

Next we consider Boron, and in particular BH_3. As with Beryllium, one of the $2s$-orbitals is excited to a $2p$-orbital and this time we form three sp^2 hybrid orbitals, which all lie in a plane and point to the corners of an equilateral triangle. Hybrid orbital theory thus predicts that BH_3 is trigonal planar.

Finally, we consider the remaining cases of Carbon, Nitrogen and Oxygen, each of which forms four sp^3 hybrid orbitals, which point to the corners of an equilateral tetrahedron. Therefore hybrid orbital theory predicts that CH_4 will be a tetrahedron. However, we note that in NH_3 and H_2O will also form 'tetrahedrons' but with one and two of the vertices occupied only by electrons, leading to a pyramidal shape for NH_3 and a bent planar molecule for H_2O.

These results all agree qualitatively with the experimental molecular shapes and the method can be extended to include d shell electrons. The clear disadvantage of this implementation of hybrid orbital method is that it does not give a quantitative bond angle and, in particular, does not explain why H_2O and NH_3 have bond angles 104.5° and 107.3° respectively [DFRSP82], rather than the predicted tetrahedral angle of 109.5°. It does however show that H_2O should have a bond angle closer to 109.5° than to 120°, which would be the predicted bond angle for a trigonal planar molecule. It is also not applicable for molecules such as NCH where the outer atoms are of different species, but hybrid orbital theory is useful in organic chemistry where, in general, only Carbon and Hydrogen atoms need be considered.

We now move on to discuss VSEPR which can be combined with hybrid orbital theory (although it was not originally formulated using such orbitals, merely under the assumption that all bonds we equivalent) to produce more accurate predictions of bond angles.

6.1.3 Valence Shell Electron Pair Repulsion

VSEPR was first introduced by Sidgwick and Powell [SP40] and proposes that the angular arrangement of electrons in the central atom of a simple molecule is a function of only the total number of valence shell electrons of the atom plus the valence (number of bonding orbitals) of the attached atoms. They speculated that there is a bijection between the resulting number of electron pairs and the shape of the molecule. Two pairs gives a linear molecule, three a trigonal planar molecule, four pairs a tetrahedral molecule and so on.

These results are consistent with all the AH_n molecules studied and also with the hybrid orbital theory. In this original formulations there is no differentiation between bonding and non-bonding (or lone) pairs of electrons. It should be noted that the results are consistent with the classical model of electrons minimizing their interaction through $1/r_{12}$ [Gil60].

An important improvement in the VSEPR model was introduced by Gillespie and Nyholm [GN57] who introduced a distinction between lone pairs and bonding pairs and specifically in the ordering of their repulsive strengths as:

$$\text{lone pair - lone pair} > \text{lone pair - bond pair} > \text{bond pair - bond pair}.$$

The intuitive explanation for this is that the bonding orbitals are pulled away from the nucleus and hence are further apart than the lone pairs and will repel each other more weakly than the lone pairs.

This improves the accuracy of the predicted angles, for example in H_2O we see that the two lone pairs should repel each other and the bond pairs more strongly than the bond pairs repel each other, leading to a contraction of the bond angle from 109.5° towards the experimental value of 104.5°. There is a similar effect (but less so due to the presence of only one lone pair) in NH_3.

6.1.4 Walsh Diagrams

The two methods so far require information only about the valence of atoms forming the molecule. An alternative description of bonding would be to consider the formation of simple molecular orbitals and the resulting change in their energies as the geometry of the molecule is varied. This is precisely what is done when using Walsh diagrams, which were first introduced by Walsh in [Wal53] and later modified by Mulliken [Mul55]. Walsh's determination of the effects of geometry on molecular energy was based on experimental data for molecular geometries, excitation energies and intuitive arguments regarding overlap of the orbitals. The resulting Walsh diagrams are essentially a plot of valence shell orbital energies against bond angle.

The assumption is then made that the total energy of the molecule is minimized when the sum of orbital energies is minimized. This is plausible when the other terms in the Hamiltonian, such as the nuclear-nuclear repulsion are relatively independent of angle but breaks down when these terms differ largely with angle. For example, in molecules with highly ionic bonds, the internuclear forces are clearly not independent of the bond angle. In this case the Walsh diagrams tend to fail to predict the correct geometry, for example, predicting that Li_2O is bent whereas experimentally it is straight.

A further assumption is that the curves are independent of the charge of the nuclei involved, in other words the diagram should look qualitatively the same for all atoms [BP72]. Thus we are led to the conclusion that, for example, the triplet state of CH_2 has the same geometry as the ground state of BH_2.

For molecules which are formed from covalent bonds, Walsh diagrams give very good qualitative agreement with experiment. This is fairly unsurprising since the original curves in the diagrams were fitted somewhat to experimental data. The reason for this empirical fitting is that, whilst the general shapes of the curves can be well explained by the change in overlap of bonding and antibonding orbitals, as is done in [Wal53] and extended/corrected (Walsh neglected a $2s$ contribution to the $3a_1$ state for bent molecules, leading to an incorrect slope) in [Mul55], the precise forms of these curves cannot be determined *a priori*. The slopes of the curves were therefore chosen to fit experimental data, Walsh going so far as determining where certain curves should cross.

We now give some examples of the results of using a Walsh diagram (see Figure 6.1) to predict the geometry of AH_2 molecules. For this it is necessary to explain the commonly used notation for molecular orbitals. There are two distinct cases when labelling the orbitals, namely whether the molecule is linear or bent.

In linear molecules, the orbitals are denoted by σ, π, etc, which denote the symmetry with respect to rotation about the bond. σ orbitals are completely symmetric under this rotation and π bonds are antisymmetric under rotation by 180°. These orbitals are further categorized with subscripts g and u to denote

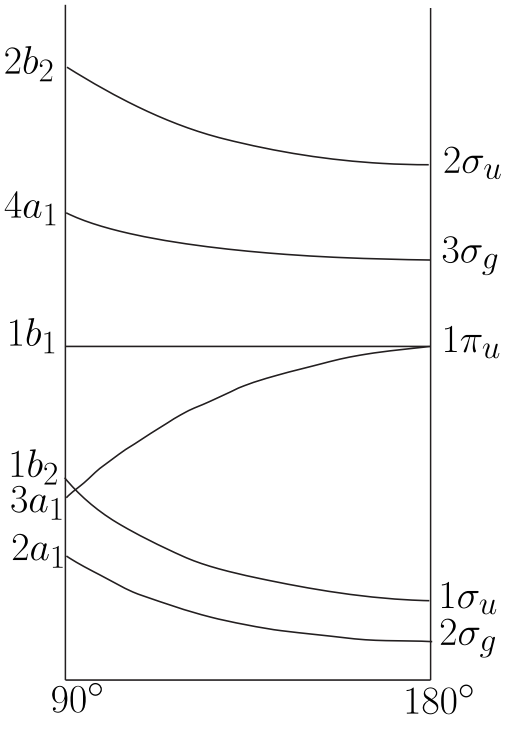


Figure 6.1: An example Walsh Diagram, essentially that of [Her66]. Labels are as described in the text, with the additional first number a simple numerical ordering of the orbitals of the same symmetry. The lowest lying $\sigma_g - a_1$ filled by the $1s$ orbitals is not shown.

whether they are symmetric or antisymmetric respectively under inversion about the central atom. Here inversion about the central atom means taking the central atom to be at 0 and mapping $(x, y, z) \mapsto -(x, y, z)$.

In the case of a bent molecule the orbitals are labelled as a or b depending if they are symmetric or antisymmetric respectively under rotation by 180° about z, the bisector of the two bonds in the plane. Further there are subscripts 1 or 2 denoting symmetry and antisymmetry respectively under reflection in the plane perpendicular to the plane of the molecule and passing through z.

Upon bending, a σ_g orbital becomes an a_1 orbital, a σ_u orbital becomes a b_2 orbital and the degenerate π_u orbitals split to become one a_1 and one b_1 orbital.

Beginning with BeH_2 we see that the four valence electrons will be placed in the $\sigma_g - a_1$ and $\sigma_u - b_2$ orbitals which favour linearity, leading to a linear molecule.

For CH_2 the first four electrons are again placed in the $\sigma_g - a_1$ and $\sigma_u - b_2$ orbitals and it could be argued that, in the degenerate linear molecule, the remaining two electrons would be placed with the same spin in the π_u orbitals and hence in the bent case there would be one each in the b_1 and a_1 orbitals. However, it is not clear whether the extra exchange terms are more important than the lowering of energy when bent caused by placing both electrons in the a_1 orbital with opposite spin. This could only be determined by explicit calculation and would depend on the equilibrium angle, the more bent it is the more likely it is to favour the singlet state.

For H_2O the eight orbitals will fill the $\sigma_g - a_1$ (linear), $\sigma_u - b_2$ (linear), $\pi_u - a_1$ (bent) and $\pi_u - b_1$ (neither) orbitals where the term in parenthesis denotes the favoured configuration for the orbital. The next question is whether the lowering of the $\pi_u - a_1$ orbital energy in bending overcomes the raising of the $\sigma_g - a_1$ and $\sigma_u - b_2$ energies. Walsh argued that this is indeed the case and hence H_2O is (strongly) bent.

The most difficult case is probably BH_2 as this is the transition between straight and bent molecules and will have the smallest energy difference between the two cases (having only one upper a_1 orbital as compared to two in H_2O). It is once again important to be able to accurately determine the slopes of the curves. Walsh claimed that the curves lead to a bent molecule but without explicit computation of the curves this cannot be completely decided. It should be noted that the experimental geometry of BH_2 was known to be bent at this time.

In order to investigate the accuracy of the Walsh diagrams, a number of studies have been performed using SCF methods to generate the orbital-energy vs. angle curves, for example [SD73] and [CC98]. These studies generally use a HF-like method so that the individual orbital energies are well defined and thus the appropriate curves can be determined. The equilibrium bond angle is then determined by taking the configuration that minimizes the sum of orbital energies.

These methods tend to give good agreement with the qualitative shapes of the Walsh diagram curves but do not always give accurate predictions for the ground state geometry. For example, as pointed out in [SD73] the Gaussian orbital results of [MH65] for H_2O do not give a minimum for this energy sum even for angles as small as 45°. This clearly predicts a bent molecule but the vast difference between this and the experimental value throws some doubt on the accuracy of the method.

Similarly, when using a RHF basis, [SD73] predicts that BeH_2 should be linear (valence orbital energy sum) but bent with an angle of less that 90° for the full energy including the $1s$-orbitals. Analogously, CH_2 is predicted to be bent with symmetry $^3B_1$ and angle 123° (valence) compared to symmetry $^1A_1$ and angle $< 90°$ (full). Both NH_2 and H_2O are predicted to be bent but with angles $< 90°$ for both valence and full energy.

This clearly shows that RHF is not an effective method to use and better qualitative results are obtained in the same paper using an internally consistent SCF method. Even so, the predicted angles are still quite a way from the experimental values, being off by $10 - 20°$ in a number of cases.

The Walsh diagrams produced by these models are in good agreement with each other and with the qualitative ones produced by Walsh. There are however a number of notable differences such as the lower a_1 curve slightly favouring 90° as compared to the linear prediction of Walsh. Also, the b_2 curve actually increases faster than the upper a_1 curve decreases as the angle decreases. It is the combined effect with the change in the lower a_1 curve that produces the predicted effects.

The fact that the RHF and ICSCF curves are qualitatively similar but give vastly different results suggests that the correct predictions of the Walsh diagrams is heavily dependent on the exact forms of the curves and hence an *a priori* qualitative prediction without fitting to experimental results would be unlikely to give good correlation with experiment.

It should be noted that it is not only Walsh diagram type arguments that give results differing significantly from experiment. An example is the hybrid-orbital-like study of BH_2 [JLH62] which predicted (before the physical geometry was known) that the BH_2 ground state should be linear. It has since been shown [HJ67] that the BH_2 ground state is in fact bent with a bond angle of 131°.

6.1.5 Mathematical Justification

The numerical methods used to determine molecular geometries are usually mathematically rigorous in the limit of infinite basis sets but there seem to be no error estimates or rate of convergence arguments for quantites such as the bond angle for different basis sets.

Allen [All72] notes that VSEPR can be made mathematically and physically rigorous and also that the results are equivalent to those of Walsh diagrams. This equivalence is shown using trace invariance and symmetry arguments, and a mathematical justification for Walsh diagrams (i.e. the fact that the minimum of the valence electron energies determines the global minimum) is given under suitable assumptions (little charge transfer between atoms, i.e. predominantly covalent bonding).

6.1.6 Pair density and bond angle of AH_2

As has been noted above, methods such as Walsh diagrams are somewhat empirical and fitted to experimental values. When rigorous *a priori* calculations are performed, the agreement with experiment seems to provide either information about the effect of bonding on atomic orbitals, with erratic numerical results, or very good numerical results with little physical understanding.

We wish to investigate another method which will hopefully give both a good physical understanding of why some molecules are bent and others are linear, as well as giving reasonably accurate bond angles. This could also be used as the first step in an explanation of more complicated molecular bonding, just as the calculation of simple atomic energies and wavefunctions can be seen as a step towards more complicated larger atoms.

It is clear from methods such as VSEPR that good qualitative bonding descriptions can be derived from models containing very little information. However, it is also clear that more information than simply the number of electron pairs is needed to produce accurate numerical results.

Our claim is that the pair density (classically the probability of finding two electrons simultaneously at spatial positions x and y) is a good approximation to the probability of two Hydrogen atoms forming bonds with the central atom that pass through x and y. Hence the most likely bond angle should be given by the maximum of the pair density.

We start by making a few simplifications to the model in order to make it analytically soluble in a relatively simple way. Firstly, we assume that the two bonding electrons will be at the same, fixed, distance r from the nucleus of the central atom. The justification for this is that the Hydrogen atoms are indistinguishable and hence the electrons in each bond should, on average, be the same distance from the nucleus of the centre atom. This is confirmed by experimental data, as commented on above, with the two bonds in AH_2 having the same bond length and bond energy.

Secondly, we ignore the bonding electrons contributed by the Hydrogen atoms. Methods such Walsh diagrams and molecular orbitals make great use of the fact that the Hydrogen atoms have a $1s$-orbital and model the way these $1s$-orbitals interact with the central atomic valence orbitals. However, VSEPR makes no use at all of the type of orbitals on the outer atoms, it simply requires knowledge of the number of binding electrons. This is analogous to using the pair density, if we were to find the equilibrium bond angle of AH_3 we would need to use the three point density and so on.

Thirdly, we ignore the interaction of the Hydrogen atoms (nuclei and electrons) both with each other and with the central atom. This can be thought of as assuming that all inter-nuclear terms and nuclear-electron terms between different atoms are independent of angle. In practice this is clearly not the case.

Effects similar to this can be seen in the other models where it is normally argued that, upon bonding, the Hydrogen atoms gain a small positive charge and hence repel each other. This is taken to mean that the predicted bond angles should be too small. Whether this is indeed the over-riding effect or whether the tendency for two Hydrogen atoms to form H_2 and thus be attracted to each other, leading to the bond angle prediction being too large, is not clear.

In conclusion, we propose the following, which we have attempted to justify above but for which we have no mathematical proof:

Proposition 6.1.1 (Pair Density and Bond Angle)**.** *The bond angle in the tri-atomic molecule AH_2 is well-approximated by the angle between the two vectors x and y which maximize the pair density of the atomic ground state of A with $|x| = |y| = r$, where r represents the 'typical distance' of the bonding electrons from the nucleus.*

Although the pair density is a well-known quantity in the chemistry literature, it appears that this is the first time it has been used to predict bond angles. As a consequence of this, all results in this chapter are new.

6.1.7 Choice of r

In order to predict the experimental bond angle and not just the general behaviour of the bond angle with r we need to choose the value of r which best represents the experimental position of the binding electrons of the central atom A. One possibility would be to use the experimental bond length; this is not the correct quantity as it gives the separation of the A and H nuclei and not the position of the electrons.

A second option would be to use the expected radius of the valence electrons in the ground state of the central atom. This seems a more sensible definition and has the advantage that we may use our theoretical values rather than relying on experimental ones. It seems probable that this is not the ideal value to use as it does not take into account hybridization when forming molecules. We take these values to be representative of typical bond lengths rather than direct predictions of experimental values. For consistency we take the position of the electrons in the bond to be the position of the maximum valence shell density, as given in Table 5.2.

The experimental bond lengths and angles are given in Table 6.1.

	Bond Angle (°)	Bond Length (au)
BeH_2	180	2.507
BH_2	131	2.2283
CH_2	136	2.034
NH_2	103.3	1.932
OH_2	104.5	1.808

Table 6.1: Experimental bond lengths and angles for XH_2 [DFRSP82].

6.1.8 Mathematical Formulation of the Pair Density

We define the spatial pair density of a wavefunction $\Psi(x_1, \dots, x_N)$ to be

$$\rho_2^\Psi(x_1, x_2) := \sum_{s_1,\dots,s_N} \int |\Psi(x_1, x_2, x_3, \dots, x_N)|^2 dx_3 \dots dx_N,$$

where $x_i = (r_i, s_i)$. The quantity $\rho_2^\Psi(x_1, x_2) dv_1 dv_2$ is the probability of finding one electron in the volume dv_1 around position x_1 and another electron in the volume dv_2 around position x_2, independently of the positions of all other electrons and of all spins. We have not performed the usual normalization (see e.g. [Löw55]) to the the number of pairs, which would have a prefactor $\binom{N}{2}$; we are only interested in the critical points of the pair density and hence the extra prefactor serves no useful purpose.

We now rewrite the pair density at a pair of points $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3$ as

$$\rho_2^\Psi(x, y) =: \langle \Psi | V_{x,y} | \Psi \rangle,$$

where the operator $V_{x,y}$, for $x \neq y$, is a two particle operator given by

$$V_{x,y} = \sum_{1 \leq i < j \leq N} \delta(x_i - x)\delta(x_j - y).$$

It is clear that the wavefunctions will need to be sufficiently well-behaved to allow integration against a delta distribution. This is true for the true wavefunctions which are continuous [Kat57] and also clearly true for our wavefunctions. We also note that $V_{x,y}$ is of the form required to apply Slater's rules from Theorem 2.3.1. The integrals in Slater's rules are given by

$$\begin{aligned}\langle ij|k\ell\rangle :&= \int_{(\mathbb{R}^3 \times \mathbb{Z}_2)^2} dx_1 dx_2 \psi_i^*(x_1, s_1)(x_1, s_1)\psi_j^*(x_2, s_2)V_{x,y}\psi_k(x_1, s_1)\psi_\ell(x_2, s_2)\\ &= \sum_{s_1, s_2} \psi_i^*(x, s_1)\psi_j^*(y, s_2)\psi_k(x, s_1)\psi_\ell(y, s_2)\end{aligned}$$

and, analogously to the exchange terms in the energy calculation, we see that the integral vanishes unless the pairs ψ_i-ψ_k and ψ_j-ψ_ℓ have the same spin. The pair densities will therefore have very similar forms to the expected values of V_{ee}, except that we no longer have the symmetries that, e.g. $[2sp_1|p_1 2s] = [2sp_2|p_2 2s] = [2sp_3|p_3 2s]$.

This is equivalent to the well known fact that the expected value of V_{ee} can be written purely in terms of the pair density:

$$\langle V_{ee} \rangle_\Psi := \int_{\mathbb{R}^6} \frac{\rho_2^\Psi(x, y)}{|x - y|} dx dy.$$

There is a relatively simple form for the pair density of a single Slater determinant:

$$\begin{aligned}\langle \Psi | V_{x,y} | \Psi \rangle &= \tfrac{1}{2} \sum_m^N \sum_n^N \langle mn|mn\rangle - \langle mn|nm\rangle\\ &= \tfrac{1}{2} \sum_{s_1, s_2} \sum_m^N \sum_n^N m^*(x)m(x)n^*(y)n(y) - m^*(x)n(x)m(y)n^*(y)\\ &= \tfrac{1}{2} \sum_{m<n} \big(|m(x)|^2|n(y)|^2 + |n(x)|^2|m(y)|^2\big) - \mathrm{Re} \sum_{\substack{m<n\\ m,n \text{ same spin}}} m^*(x)n(x)m(y)n^*(y), \quad (6.1)\end{aligned}$$

where all the sums run over the orbitals in $|\Psi\rangle$. We do not compute the other cases from Theorem 2.3.1 as we do not yet know whether we will be dealing with singly and/or doubly excited orbitals. (Recall from Section 3.3 that the energy calculations did not require the use of singly excited Slater determinants.)

6.2 Spin and Angular Momentum dependence of the Pair Density

In order to consider the pair density of a general atomic ground state (for which $\langle \underline{L}^2\rangle_\Psi$ and $\langle \underline{S}^2\rangle_\Psi$ are fixed) we must consider the effects of the choice of linear combinations of the L_3 and S_3 eigenfunctions within a given $\underline{L}^2$-$\underline{S}^2$ eigenspace. Since the spatial pair density does not depend on spin, and the energy is the same for any given S_3 eigenfunction, we expect the pair density to be independent of the choice of S_3 eigenvalue and indeed of the fact that the function may or may not be an eigenfunction of S_3. However, the effects of taking linear combinations of L_3 eigenfunctions seem less clear and this will be discussed later.

6.2.1 Spin Invariance of the Pair Density

In order to show that the pair density is independent of the choice of linear combination of S_3 eigenfunctions we wish to show that the ladder operator $S_+ := S_1 + iS_2$, which for Ψ an eigenfunction of S_3 with eigenvalue s satisfies $S_3S_+\Psi = (s+1)S_+\Psi$, commutes with pair density.

Lemma 6.2.1. *The expected value of the pair density* $\rho_2^\Psi(x,y) := \langle\Psi|V_{x,y}|\Psi\rangle$ *where*

$$V_{x,y} = \sum_{1\le i<j\le N} \delta(x_i - x)\delta(x_j - y)$$

is invariant under the application of the spin ladder operator $S_+ := S_1 + iS_2$ *for all* $\Psi \notin \ker S_+$ *with* $S_3\Psi = s\Psi$, $\underline{S}^2 = \mathcal{S}\Psi$. *I.e. we have*

$$\frac{\langle V_{x,y}\rangle_{S_+\Psi}}{\langle 1\rangle_{S_+\Psi}} = \frac{\langle V_{x,y}\rangle_\Psi}{\langle 1\rangle_\Psi}.$$

Before we prove this lemma, which we will do using second quantization, we need to find the second quantized form of the pair density.

Lemma 6.2.2. *Let* $V = \sum_{i=1}^N \sum_{j>i}^N v(x_i, x_j)$ *be the sum of spin-independent operators* $v(\cdot,\cdot)$ *which is invariant under permutation of the coordinates. Then the* *action of* V *on a wavefunction formed from orbitals* ψ_i *and* $\overline{\psi_i}$ *is given by*

$$V = \tfrac{1}{2} \sum_{p,q,r,s} (pq|rs) \left(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps} \right)$$

where the sum runs over all orbitals,

$$\hat{E}_{ij} := a^\dagger(\psi_i)a(\psi_j) + a^\dagger(\overline{\psi_i})a(\overline{\psi_j})$$

and $(pq|rs)$ *is the purely spatial integral*

$$(pq|rs) = \iint_{\mathbb{R}^6} dx_1 dx_2 \psi_p^*(x_1)\psi_q(x_1)v(x_1,x_2)\psi_r^*(x_2)\psi_s(x_2).$$

Proof We begin with the second quantized form of a general two-particle operator as given in [SO96]:

$$V = \tfrac{1}{2} \sum_{i,j,k,\ell} \langle ij|k\ell\rangle a^\dagger(\chi_i)a^\dagger(\chi_j)a(\chi_\ell)a(\chi_k).$$

Recall that $\langle ij|k\ell\rangle = [ik|j\ell]$ and so

$$V = \tfrac{1}{2} \sum_{p,q,r,s} [pq|rs] a^\dagger(\chi_p)a^\dagger(\chi_r)a(\chi_s)a(\chi_q),$$

where the sum runs over spatial-spin orbitals. Noting that, for a spin independent operator V, $[pq|rs]$ is zero unless the pairs p,q and r,s are the same spin, we have, after integrating $[pq|rs]$ over spin,

$$\begin{aligned} V = \tfrac{1}{2} \sum_{p,q,r,s} (pq|rs) \big[& a^\dagger(\psi_p)a^\dagger(\psi_r)a(\psi_s)a(\psi_q) + a^\dagger(\overline{\psi_p})a^\dagger(\overline{\psi_r})a(\overline{\psi_s})a(\overline{\psi_q}) \\ & + a^\dagger(\psi_p)a^\dagger(\overline{\psi_r})a(\overline{\psi_s})a(\psi_q) + a^\dagger(\overline{\psi_p})a^\dagger(\psi_r)a(\psi_s)a(\overline{\psi_q}) \big], \end{aligned} \tag{6.2}$$

where the sum runs over only spatial orbitals. We now note that the two annihilation terms either correspond to different orbitals or give zero and so we may anticommute them:

$$\begin{aligned} V = -\tfrac{1}{2} \sum_{p,q,r,s} (pq|rs) \big[& a^\dagger(\psi_p)a^\dagger(\psi_r)a(\psi_q)a(\psi_s) + a^\dagger(\overline{\psi_p})a^\dagger(\overline{\psi_r})a(\overline{\psi_q})a(\overline{\psi_s}) \\ & + a^\dagger(\psi_p)a^\dagger(\overline{\psi_r})a(\psi_q)a(\overline{\psi_s}) + a^\dagger(\overline{\psi_p})a^\dagger(\psi_r)a(\overline{\psi_q})a(\psi_s) \big]. \end{aligned}$$

We now swap the terms corresponding to the indices q and r:

$$\begin{aligned} V = -\tfrac{1}{2} \sum_{p,q,r,s} (pq|rs) \big[& a^\dagger(\psi_p)[\delta_{q,r} - a(\psi_q)a^\dagger(\psi_r)]a(\psi_s) \\ & + a^\dagger(\overline{\psi_p})[\delta_{q,r} - a(\overline{\psi_q})a^\dagger(\overline{\psi_r})]a(\overline{\psi_s}) \\ & - a^\dagger(\psi_p)a(\psi_q)a^\dagger(\overline{\psi_r})a(\overline{\psi_s}) - a^\dagger(\overline{\psi_p})a(\overline{\psi_q})a^\dagger(\psi_r)a(\psi_s)\big] \end{aligned}$$

and expanding out we obtain

$$\begin{aligned} V = \tfrac{1}{2} \sum_{p,q,r,s} (pq|rs) \big[& a^\dagger(\psi_p)a(\psi_q)a^\dagger(\psi_r)a(\psi_s) + a^\dagger(\overline{\psi_p})a(\overline{\psi_q})a^\dagger(\overline{\psi_r})a(\overline{\psi_s}) \\ & + a^\dagger(\psi_p)a(\psi_q)a^\dagger(\overline{\psi_r})a(\overline{\psi_s}) + a^\dagger(\overline{\psi_p})a(\overline{\psi_q})a^\dagger(\psi_r)a(\psi_s) \\ & - \delta_{q,r}a^\dagger(\psi_p)a(\psi_s) - \delta_{q,r}a^\dagger(\overline{\psi_p})a(\overline{\psi_s})\big]. \end{aligned}$$

Introducing the operator

$$\hat{E}_{ij} := a^\dagger(\psi_i)a(\psi_j) + a^\dagger(\overline{\psi_i})a(\overline{\psi_j})$$

gives the result. □

Proof of Lemma 6.2.1 We first show that $V_{x,y}$ and S_+ commute. We begin by writing both operators in second quantization, where it is clear that $S_+ = \sum_i a^\dagger(\psi_i)a(\overline{\psi_i})$ and, by Lemma 6.2.2, we have

$$V_{x,y} = \tfrac{1}{2} \sum_{p,q,r,s} \psi_p^*(x)\psi_q(x)\psi_r^*(y)\psi_s(y) \left(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}\right),$$

with

$$\hat{E}_{ij} := a^\dagger(\psi_i)a(\psi_j) + a^\dagger(\overline{\psi_i})a(\overline{\psi_j}).$$

For the first term of $\hat{E}_{ij}$ and a general term from S_+ denoted $a^\dagger(\psi_k)a(\overline{\psi_k})$, we have, noting that $a^\dagger(\varphi)a^\dagger(\varphi) = 0$,

$$a^\dagger(\psi_i)a(\psi_j)a^\dagger(\psi_k)a(\overline{\psi_k}) = \delta_{kj}a^\dagger(\psi_i)a(\overline{\psi_k}) + a^\dagger(\psi_k)a(\overline{\psi_k})a^\dagger(\psi_i)a(\psi_j),$$

which gives

$$[a^\dagger(\psi_i)a(\psi_j), a^\dagger(\psi_k)a(\overline{\psi_k})] = \delta_{kj}a^\dagger(\psi_i)a(\overline{\psi_k}). \tag{6.3}$$

Similarly we find that

$$[a^\dagger(\overline{\psi_i})a(\overline{\psi_j}), a^\dagger(\psi_k)a(\overline{\psi_k})] = -\delta_{ik}a^\dagger(\psi_k)a(\overline{\psi_j}). \tag{6.4}$$

combining (6.3) and (6.4) and summing over k to get the whole of S_+ gives

$$[\hat{E}_{ij}, S_+] = \sum_k \left(\delta_{kj} a^\dagger(\psi_i) a(\overline{\psi_k}) - \delta_{ik} a^\dagger(\psi_k) a(\overline{\psi_j}) \right) = 0,$$

from which it follows trivially that $[S_+, V_{x,y}] = 0$.

We now wish to show that this commutation implies that the pair density is invariant under the application of S_+ to a wavefunction Ψ with the stated properties. First note that $S_+^* = S_- = S_1 - iS_2$ and hence we see that

$$\begin{aligned} S_+^* S_+ = S_- S_+ &= (S_1 - iS_2)(S_1 + iS_2) = S_1^2 + S_2^2 + i[S_1, S_2] \\ &= S_1^2 + S_2^2 - S_3 = \underline{S}^2 - S_3^2 - S_3, \end{aligned}$$

where we have used the identities $[S_1, S_2] = iS_3$ and $\underline{S}^2 = S_1^2 + S_2^2 + S_3^2$. Therefore we have

$$\begin{aligned} \frac{\langle S_+\Psi | V_{x,y} | S_+\Psi \rangle}{\langle S_+\Psi | S_+\Psi \rangle} &= \frac{\langle S_+\Psi | S_+ V_{x,y}\Psi \rangle}{\langle S_+\Psi | S_+\Psi \rangle} = \frac{\langle S_+^* S_+ \Psi | V_{x,y} | \Psi \rangle}{\langle S_+^* S_+ \Psi | \Psi \rangle} \\ &= \frac{\langle (\underline{S}^2 - S_3^2 - S_3)\Psi | V_{x,y} | \Psi \rangle}{\langle (\underline{S}^2 - S_3^2 - S_3)\Psi | \Psi \rangle} = \frac{(\mathcal{S}^2 - s^2 - s)\langle \Psi | V_{x,y} | \Psi \rangle}{(\mathcal{S}^2 - s^2 - s)\langle \Psi | \Psi \rangle} \\ &= \frac{\langle \Psi | V_{x,y} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \end{aligned}$$

and the result holds, since $\mathcal{S}^2 - s^2 - s = 0$ if and only if $\Psi \in \ker S_+$. □

It is clear that, since S_i, $i = 1, 2, 3$ commute with L_3, $\underline{L}^2$ and $\underline{S}^2$, S_+ also commutes with these three operators. For a given $\underline{L}^2$-$\underline{S}^2$-L_3 simultaneous eigenspace with $\underline{S}^2$ eigenvalue $k(k+1)$ there exists a wavefunction Ψ_{-k} with S_3 eigenvalue $-k$ and energy E. The remaining S_3 eigenfunctions within the simultaneous eigenspace with the same energy can be generated by

$$\Psi_{-k+i} = \frac{S_+^i \Psi_{-k}}{\langle S_+^i \Psi_{-k} | S_+^i \Psi_{-k} \rangle^{1/2}}.$$

Further we may write a general normalized wavefunction with energy E within the $\underline{L}^2$-$\underline{S}^2$-L_3 simultaneous eigenspace as

$$\Psi = \sum_{i=-k}^{k} \alpha_i \Psi_i$$

where $\sum_{i=-k}^{k} |\alpha_i|^2 = 1$. By Lemma 6.2.1 we see that $\rho_2^{\Psi_i} = \rho_2^{\Psi_{-k}}$ for all $i =$

$-k,\dots,k$. Furthermore, since S_3 is self-adjoint and the Ψ_j satisfy $S_3\Psi_j = j\Psi_j$, the Ψ_j are all orthogonal with respect to integration over spin. Thus, since $V_{x,y}$ is independent of spin, it follows that $\langle\Psi_i|V_{x,y}|\Psi_j\rangle = 0$ for all $i \neq j$. Therefore we have that

$$\begin{aligned}\frac{\langle\Psi|V_{x,y}|\Psi\rangle}{\langle\Psi|\Psi\rangle} &= \sum_{i=-k}^{k} |\alpha_i|^2\langle\Psi_i|V_{x,y}|\Psi_i\rangle + \sum_{i\neq j}\alpha_i\alpha_j^*\langle\Psi_i|V_{x,y}|\Psi_j\rangle \\ &= \sum_{i=-k}^{k} |\alpha_i|^2\langle\Psi_{-k}|V_{x,y}|\Psi_{-k}\rangle = \langle\Psi_{-k}|V_{x,y}|\Psi_{-k}\rangle\end{aligned}$$

and the following corollary holds:

Corollary 6.2.3. *Let Ψ_{-k} satisfy $\underline{L}^2\Psi_{-k} = \mathcal{L}\Psi_{-k}$, $L_3\Psi_{-k} = \ell\Psi_{-k}$, $\underline{S}^2\Psi_{-k} = k(k+1)\Psi_{-k}$, and $S_3\Psi_{-k} = -k\Psi_{-k}$. Then for any wavefunction $\Psi = \sum_{i=-k}^{k}\alpha_i\Psi_i$ where Ψ_i are given by*

$$\Psi_{-k+j} = \frac{S_+^j\Psi_{-k}}{\langle S_+^j\Psi_{-k}|S_+^j\Psi_{-k}\rangle^{1/2}}, \quad j = 1,\dots,2k,$$

with $\sum_{i=-k}^{k}|\alpha_i|^2 = 1$, the pair density $\langle\Psi|V_{x,y}|\Psi\rangle$ is independent of the choice of α_i.

This means that we may freely pick an S_3 eigenfunction for a given $\underline{L}^2$-$\underline{S}^2$ eigenspace from which to calculate the pair density. Thus the only choice that remains for a given $\underline{L}^2$-$\underline{S}^2$ eigenspace is which linear combination of L_3 eigenfunctions to choose.

6.2.2 Angular Momentum Dependence of the Hydrogen Density

We claim that, unlike for spin, the pair density depends on the choice of linear combination of the L_3 eigenfunctions. This is motivated by the simple example of the density of the Hydrogen p-orbital for which there is a one-parameter family of densities. Consider

$$\Psi(x) = \sum_{i=1}^{3}\alpha_i p_i(x) = k\big((a+ib)\cdot x\big)e^{-|x|/2}$$

where $\alpha_i \in \mathbb{C}$, $\sum_{i=1}^{3}|\alpha_i|^2 = 1$, $k \in \mathbb{R}$, $a, b \in \mathbb{R}^3$ and $|a|^2 + |b|^2 = 1$, then

$$\rho^\Psi = \big(|a\cdot x|^2 + |b\cdot x|^2\big)k^2e^{-|x|} = \big(x\cdot(a\otimes a + b\otimes b)x\big)k^2e^{-|x|}.$$

We now apply a general rotation $R \in SO(3)$ and note that $\mathrm{Trace}(R(a \otimes a + b \otimes b)R^T) = |a|^2 + |b|^2 = 1$ and $\mathrm{Rank}(a \otimes a + b \otimes b) \leq 2$, and so we may choose R such that

$$R(a \otimes a + b \otimes b)R^T = \begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \lambda \in [0, 1/2].$$

This gives a one-parameter family of densities with canonical form

$$\rho^\Psi = \big((1-\lambda)x_1^2 + \lambda x_2^2\big)k^2 e^{-|x|}. \qquad (6.5)$$

Figure 6.2 shows this density for a number of choices of the parameter λ, the case $\lambda = 0$ is the traditional view of a p-orbital. Note that, despite the relatively simple proof, this flexibility of the hydrogen orbitals does is not mentioned in the literature.

This analysis suggests that a general pair density containing p-orbitals is also likely to depend on a number of parameters and we must derive a canonical pair density in each case. We expect a parametrized pair density only when p-orbitals are present as s-orbitals are invariant under rotation.

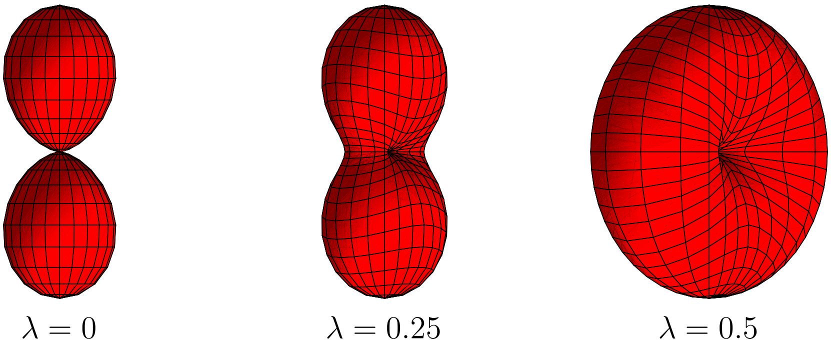

$\lambda = 0$ $\lambda = 0.25$ $\lambda = 0.5$

Figure 6.2: Plots of the one-parameter family of Hydrogen p-orbital densities given in (6.5).

We now investigate the physical meaning of λ and in particular justify why we have described this as an angular momentum dependence of the density.

6.2.3 Physical Interpretation of λ

Intuitively, λ is some measure of the axis under which the wavefunction is rotationally invariant and hence we expect it to be connected with angular momentum.

Consider the total angular momentum vector

$$\underline{L} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix},$$

where L_i are defined in (3.1). We note that $\underline{L}\psi$ is the zero vector on s-orbitals and hence we consider the effect only on the three dimensional space spanned by p_i. For $\Psi := \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$, $\alpha_i \in \mathbb{C}$, by (3.6) we see that

$$\begin{aligned} \langle \Psi | \underline{L} | \Psi \rangle = {} & |\alpha_1|^2 \left\langle \begin{array}{c|c} p_1 & 0 \\ p_1 & -ip_3 \\ p_1 & ip_2 \end{array} \right\rangle + |\alpha_2|^2 \left\langle \begin{array}{c|c} p_2 & ip_3 \\ p_2 & 0 \\ p_2 & -ip_1 \end{array} \right\rangle + |\alpha_3|^2 \left\langle \begin{array}{c|c} p_3 & -ip_2 \\ p_3 & ip_1 \\ p_3 & 0 \end{array} \right\rangle \\ & + 2\mathrm{Re}(\alpha_1^* \alpha_2) \left\langle \begin{array}{c|c} p_1 & ip_3 \\ p_1 & 0 \\ p_1 & -ip_1 \end{array} \right\rangle + 2\mathrm{Re}(\alpha_2^* \alpha_3) \left\langle \begin{array}{c|c} p_2 & -ip_2 \\ p_2 & ip_1 \\ p_2 & 0 \end{array} \right\rangle \\ & + 2\mathrm{Re}(\alpha_3^* \alpha_1) \left\langle \begin{array}{c|c} p_3 & 0 \\ p_3 & -ip_3 \\ p_3 & ip_2 \end{array} \right\rangle . \end{aligned}$$

Using the orthonormality of the p_i, we find

$$\langle \Psi | \underline{L} | \Psi \rangle = 2\mathrm{Re} \begin{pmatrix} -i\alpha_2^* \alpha_3 \\ -i\alpha_3^* \alpha_1 \\ -i\alpha_1^* \alpha_2 \end{pmatrix} = 2\mathrm{Im} \begin{pmatrix} \alpha_2^* \alpha_3 \\ \alpha_3^* \alpha_1 \\ \alpha_1^* \alpha_2 \end{pmatrix}.$$

Inserting our example of $\Psi := \sqrt{1-\lambda}\, p_1 + i\sqrt{\lambda}\, p_2$ gives

$$\langle \Psi | \underline{L} | \Psi \rangle = 2\mathrm{Im} \begin{pmatrix} 0 \\ 0 \\ i\sqrt{1-\lambda}\sqrt{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2\sqrt{1-\lambda}\sqrt{\lambda} \end{pmatrix},$$

therefore $|\langle \underline{L} \rangle_\Psi|^2 = 4\lambda(1-\lambda) \in [0,1]$, and since $\lambda \in [0, \frac{1}{2}]$, this uniquely determines λ.

An analogous calculation shows that setting $\Psi := \alpha_1 |p_2 p_3\rangle + \alpha_2 |p_1 p_3\rangle + \alpha_3 |p_1 p_2\rangle$ (which is the case where the spin-aligned p-shell is 'missing' an orbital) gives

$$\langle \Psi | \underline{L} | \Psi \rangle = 2\mathrm{Im} \begin{pmatrix} -\alpha_1^* \alpha_2 \\ \alpha_1^* \alpha_3 \\ -\alpha_2^* \alpha_3 \end{pmatrix}.$$

which, setting $\alpha_1 = \sqrt{1-\lambda}$, $\alpha_2 = i\sqrt{\lambda}$ and $\alpha_3 = 0$ once again gives $|\langle \underline{L} \rangle_\Psi|^2 = 4\lambda(1-\lambda)$.

6.3 Oxygen

As mentioned previously, one of the most commonly studied bond angles is that of water. We therefore begin our analysis of pair density prediction of bond angles with the central atom being Oxygen. We will cover this case in some detail, deriving a general method which will be applicable to a number of other atoms.

6.3.1 General Ground State

We begin by finding a general Oxygen ground state wavefunction from which we will derive a canonical Oxygen ground state pair density. We note that, from the restriction to the ground state energy and fixing $S_3\Psi = \Psi$ (which we may do by Corollary 6.2.3) the wavefunction is

$$\Psi = \sum_{i=1}^{3} \alpha_i |1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_i}\rangle = |1s\overline{1s}2s\overline{2s}p_1p_2p_3(\sum_{i=1}^{3} \alpha_i \overline{p_i})\rangle,$$

where $\alpha_i \in \mathbb{C}$ and $\sum_i |\alpha_i|^2 = 1$. Using Slater's rules from Theorem 2.3.1, the pair density of this general wavefunction is given by

$$\begin{aligned}\rho_2^\Psi(x,y) = \rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle} &+ \tfrac{1}{2}\Big(\big|\sum_i \alpha_i p_i(x)\big|^2 \rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(y) + x \leftrightarrow y\Big) \\ &- \tfrac{1}{2}\Big(\big(1s(x)1s^*(y) + 2s(x)2s^*(y)\big)\big(\sum_i \alpha_i p_i(x)\big)^*\big(\sum_j \alpha_j p_j(y)\big) + x \leftrightarrow y\Big).\end{aligned}$$

In order to find a canonical pair density we rewrite the sum of p-orbitals as

$$\sum_i \alpha_i p_i(x) = k\big((a+ib)\cdot x\big)e^{-Z|x|/2}, \quad k \in \mathbb{R},\ a,b \in \mathbb{R}^3,\ |a|^2 + |b|^2 = 1,$$

which gives

$$\big|\sum_i \alpha_i p_i(x)\big|^2 = \big(|a\cdot x|^2 + |b\cdot x|^2\big)k^2 e^{-Z|x|} = \big(x\cdot(a\otimes a + b\otimes b)x\big)k^2 e^{-Z|x|},$$

and also

$$\begin{aligned}
&\Big(\sum_i \alpha_i p_i(x)\Big)^* \Big(\sum_j \alpha_j p_j(y)\Big) + x \leftrightarrow y \\
&\quad = \big[\big((a-ib)\cdot x\big)\big((a+ib)\cdot y\big) + \big((a-ib)\cdot y\big)\big((a+ib)\cdot x\big)\big] k^2 e^{-Z(|x|+|y|)/2} \\
&\quad = 2\big[(a\cdot x)(a\cdot y) + (b\cdot x)(b\cdot y)\big] k^2 e^{-Z(|x|+|y|)/2} \\
&\quad = \big(x\cdot(a\otimes a + b\otimes b)y\big) k^2 e^{-Z(|x|+|y|)/2}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\rho_2^{\Psi}(x,y) = {}& \rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x,y) \\
& + \tfrac{1}{2}\Big[\big(x\cdot(a\otimes a + b\otimes b)x\big) k^2 e^{-Z|x|} \rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(y) + x \leftrightarrow y\Big] \\
& - \big(1s(x)1s(y) + 2s(x)2s(y)\big)\big(x\cdot(a\otimes a + b\otimes b)y\big) k^2 e^{-Z(|x|+|y|)/2}.
\end{aligned} \tag{6.6}$$

We now perform a simultaneous rotation of all coordinates and define $\Psi^R(x_1,\ldots,x_N) := \Psi(R^T x_1,\ldots,R^T x_N)$ where $R \in SO(3)$. In particular we wish to show that $\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x)$ and $\rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x,y)$ are invariant under this simultaneous rotation.

From Table 4.18 we see that the ground state of Nitrogen is given by the ${}^4S$ state, which, from Table 3.5 and Lemma 6.2.1 has a unique pair density generated by $\Psi = |1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle$. It is therefore sensible to consider the case of Nitrogen before proceeding with the discussion of Oxygen.

6.3.2 Nitrogen Density and Pair Density

From the above calculations for Oxygen, it is clear that we need to find the unique density and unique pair density for the ground state of Nitrogen, given by $\Psi^N = |1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle$. The uniqueness follows from Corollary 6.2.3. The single particle density is trivial to compute and is given by

$$\begin{aligned}
\rho^N(x) &= 2|1s(x)|^2 + 2|2s(x)|^2 + |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2 \\
&= 2|1s(|x|)|^2 + 2|2s(|x|)|^2 + k^2(x_1^2 + x_2^2 + x_3^3)e^{-Z|x|} \\
&= 2|1s(r)|^2 + 2|2s(r)|^2 + k^2 r^2 e^{-Zr},
\end{aligned} \tag{6.7}$$

where we use the notation $r := |x|$ and $p_i(x) = k x_i e^{-Zr/2}$.

Using the expansion (6.1) we obtain

$$\begin{aligned}
&\rho_2^N(x,y)\\
&= \tfrac{1}{2}\Bigg[|1s(x)|^2|1s(y)|^2 + 4|1s(x)|^2|2s(y)|^2 + |2s(x)|^2|2s(y)|^2\\
&\qquad + 2\big(|1s(x)|^2 + |2s(x)|^2\big)\sum_{i=1}^{3}|p_i(y)|^2 + \sum_{1\le i<j\le 3}|p_i(x)|^2|p_j(y)|^2 + x \leftrightarrow y\Bigg]\\
&\quad - \Bigg[2\cdot 1s(x)2s(x)1s(y)2s(y) + \sum_{1\le i<j\le 3} p_i(x)p_i(y)p_j(x)p_j(y)\\
&\qquad\qquad + \big(1s(x)1s(y) + 2s(x)2s(y)\big)\sum_{1\le i<j\le 3} p_i(x)p_i(y)\Bigg].
\end{aligned}$$

We now define notation for the radial terms:

$$R_1^2(r) := |1s(x)|^2, \quad R_2^2(r) := |2s(x)|^2, \quad R_3^2(r) := k^2 e^{-Zr}. \tag{6.8}$$

Fixing $|x| = |y| = r$ gives

$$\begin{aligned}
\rho_2^N(x,y) &= F(r) - R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)(x_1y_1 + x_2y_2 + x_3y_3)\\
&\quad + \tfrac{1}{2}R_3^4(r)\big[x_1^2y_2^2 + x_2^2y_1^2 + x_1^2y_3^2 + x_3^2y_1^2 + x_2^2y_3^2 + x_3^2y_2^2\\
&\qquad\qquad - 2x_1y_1x_2y_2 - 2x_1y_1x_3y_3 - 2x_2y_2x_3y_3\big]\\
&= F(r) - R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)\sum_{i=1}^{3}x_iy_i\\
&\quad + \tfrac{1}{2}R_3^4(r)\Bigg[\sum_{i=1}^{3}x_i^2\sum_{j=1}^{3}y_j^2 - \Big|\sum_{i=1}^{3}x_iy_i\Big|^2\Bigg]\\
&= \tilde{F}(r) - R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)r^2\cos\Theta - \tfrac{1}{2}R_3^4(r)r^4\cos^2\Theta
\end{aligned} \tag{6.9}$$

where $F(r)$ and $\tilde{F}(r)$ are purely radial functions and Θ is the angle between x and y, which in spherical polar coordinates is given by

$$\cos\Theta = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2(\cos\phi_1\cos\phi_2 + \sin\phi_1\sin\phi_2). \tag{6.10}$$

In particular

$$\tilde{F}(r) = R_1^4(r) + 2R_1^2(r)R_2^2(r) + R_2^4(r) + 2\big(R_1^2(r) + R_2^2(r)\big)r^2R_3^2(r) + \tfrac{1}{2}r^4R_3^4(r). \tag{6.11}$$

For the density, using (6.7) and the above notation, we have

$$\rho^N(x) = 2R_1^2(r) + 2R_2^2(r) + r^2 R_3^2(r). \tag{6.12}$$

It is clear that both the density and pair density are invariant under simultaneous rotation of x and y.

6.3.3 Maximizing the Nitrogen Pair Density

Theorem 6.3.1 (Maximum of Nitrogen Pair Density)**.** *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define*

$$A_0(r) := R_1^4(r) + 2R_1^2(r)R_2^2(r) + R_2^4(r) + 2\big(R_1^2(r) + R_2^2(r)\big)R_3^2(r) + \tfrac{1}{2}R_3^4(r),$$

$A_1(r) := \big(R_1^2(r) + R_2^2(r)\big)R_3^2(r)$, $A_2(r) := R_3^4(r)$. *Then the maximum of the Nitrogen ground state pair density*

$$\rho_2^N = A_0(r) - A_1(r)\cos\Theta - \tfrac{1}{2}A_2(r)\cos^2\Theta$$

is at $\cos\Theta = \max\left\{-1, -\frac{A_1(r)}{A_2(r)}\right\}$.

Proof Denoting the pair density in (6.9) by

$$\rho_2^N =: A_0(r) - A_1(r)\cos\Theta - \tfrac{1}{2}A_2(r)\cos^2\Theta$$

where $A_1(r) := R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)$ and $A_2(r) := R_3^4(r)$ (note the difference in definition of $R_3^2(r)$), with $A_i(r) \geq 0$, we have

$$\begin{aligned}
\frac{\partial}{\partial\Theta}\rho_2^N &= A_1(r)\sin\Theta + A_2(r)\cos\Theta\sin\Theta,\\
\frac{\partial^2}{\partial\Theta^2}\rho_2^N &= A_1(r)\cos\Theta + A_2(r)(\cos^2\Theta - \sin^2\Theta),
\end{aligned}$$

and the critical points of ρ_2^N are clearly at $\sin\Theta = 0$ and $\cos\Theta = -\frac{A_1(r)}{A_2(r)}$.

If $\sin\Theta = 0$ then $\frac{\partial^2}{\partial\Theta^2}\rho_2^N = \pm A_1(r) + A_2(r)$. The sub-case when $\cos\Theta = 1$ is clearly never a maximum as the second derivative is $A_1(r) + A_2(r) > 0$, and the sub-case $\cos\Theta = -1$ is a maximum if and only if $-A_1(r) + A_2(r) < 0$, which is equivalent to $-\frac{A_1(r)}{A_2(r)} < -1$.

If $\cos\Theta = -\frac{A_1(r)}{A_2(r)}$ then $\frac{\partial^2}{\partial\Theta^2}\rho_2^N = \frac{A_1^2(r) - A_2^2(r)}{A_2(r)}$, which is negative if and only if $A_1(r) < A_2(r)$ which is equivalent to $-\frac{A_1(r)}{A_2(r)} > -1$. Hence, the maximum pair density is at $\cos\Theta = \max\left\{-1, -\frac{A_1(r)}{A_2(r)}\right\}$. $\square$

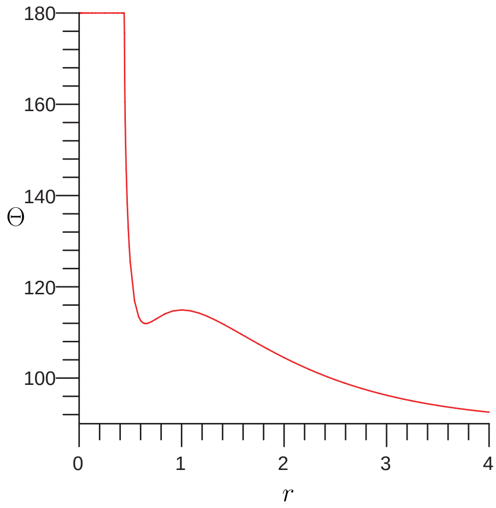


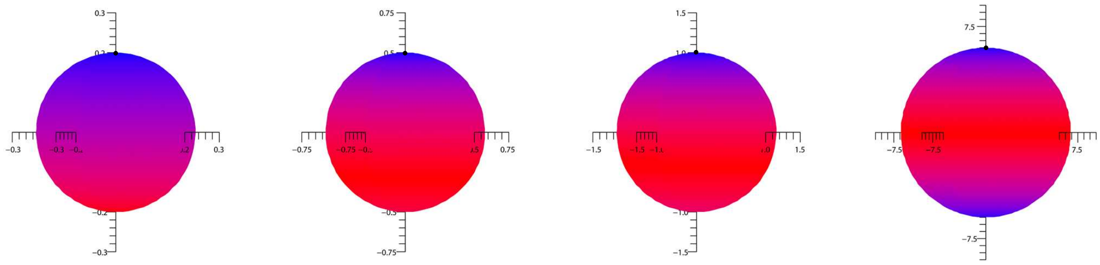
Figure 6.3: Bond angle range for Nitrogen

It follows that, since $A_2(r)$ decays (exponentially) more slowly than $A_1(r)$, the bond angle ranges from $\arccos(-1) = 180°$ close to the nucleus and tends to to $\arccos(0) = 90°$ as $r \to \infty$. This is consistent with the idea that, close to the origin, the pair density is dominated by s-p interactions which favour 180° and passes to the p-p case, which favours 90°, as the radius increases.

Figure 6.3 shows more explicitly how the predicted bond angle varies with r. In particular, we note that at a distance of 1.044au, which is the position of the outer maximum of the radial density as given in Table 5.2, our method predicts a bond angle of 114.9° as compared to the experimental value of 103.3°. Possible reasons for the discrepancy will be discussed later. Another interesting feature of the bond angle is the non-monotonicity, which follows from the non-monotonicity of $|\psi_{2s}(r)|^2$, which enters the bond angle through $A_1(r)$.

Along with the predicted bond angle at various distances, we may also plot the pair density on the sphere for a given distance. We do this by noting that the rotational invariance allows us to choose the position of the first electron, which we fix to be the north pole (marked by the black dot) and then plotting the pair density on the rest of the sphere via a colour map. The pair densities for a range of distances are plotted in Figure 6.4, blue areas denote low pair density, red high, and shades of purple the areas in between.

The pair densities are clearly invariant under rotation about the north-south axis, which follows trivially from the fact that (6.9) depends only on the angle between the two electrons. The distances chosen are such that they show the change from a bond angle of 180° near the origin ($r = 0.2$) to a bond angle of almost 90° as r becomes large ($r = 6.0$). The distances in between show the progression of the maxima, $r = 1.044$ being the 'bonding distance'.



$r = 0.2$ $r = 0.5$ $r = 1.044$ $r = 6.0$

Figure 6.4: Nitrogen pair densities for various distances. The first electron is fixed at the north pole, red denotes high probability of finding the second electron in a given region, blue low probability and shades of purple those areas in between. In particular, the third figure shows a predicted non-linear NH_2 molecule at a radius given by the outer maximum of the single particle density.

6.3.4 Simultaneous Rotation of the Oxygen Pair Density

From Section 6.3.2 and in particular the invariance of $\rho^N(r)$ under rotation and $\rho_2^N(x,y)$ under simultaneous rotation of x and y, we see that the only terms in the Oxygen pair density (6.6) that vary under simultaneous rotation are given by the tensor products.

We therefore have

$$\begin{aligned}\rho_2^{\Psi^R}(x,y) = \rho_2^N(x,y) + \tfrac{1}{2}\Big[\big(x \cdot R(a\otimes a + b\otimes b)R^T x\big)k^2 e^{-Z|x|}\rho^N(y) + x \leftrightarrow y\Big]\\ - \big(x \cdot R(a\otimes a + b\otimes b)R^T y\big)\\ \times \big(1s(x)1s(y) + 2s(x)2s(y)\big)k^2 e^{-Z(|x|+|y|)/2}.\end{aligned}$$

We now note that $\mathrm{Trace}(R(a\otimes a + b\otimes b)R^T) = |a|^2 + |b|^2 = 1$ and $\mathrm{Rank}(a\otimes a + b\otimes b) \leq 2$ and hence we may choose R such that

$$R(a\otimes a + b\otimes b)R^T = \begin{pmatrix} 1-\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & 0\end{pmatrix}, \ \lambda \in [0, 1/2].$$

It follows that there is a one-parameter family of pair densities with canonical form

$$\begin{aligned}\rho_2^{\Psi^R}(x,y) = \rho_2^N(x,y) &+ \tfrac{1}{2}\big[k^2\big((1-\lambda)x_1^2+\lambda x_2^2\big)e^{-Z|x|}\rho^N(y) + x \leftrightarrow y\big] \\ &- \big(1s(x)1s(y)+2s(x)2s(y)\big)\big((1-\lambda)x_1y_1+\lambda x_2y_2\big)k^2e^{-Z(|x|+|y|)/2}.\end{aligned} \tag{6.13}$$

The canonical form for the wavefunction representing these pair densities is given by $a = \sqrt{1-\lambda}e_1$ and $b = i\sqrt{\lambda}e_2$, with a wavefunction

$$\Psi^O = |1s\overline{1s}2s\overline{2s}p_1p_2p_3(\sqrt{1-\lambda}\overline{p_1} + i\sqrt{\lambda}\overline{p_2})\rangle. \tag{6.14}$$

6.3.5 Maximizing the Oxygen Pair Density

We wish to find the global maximum of the canonical Oxygen pair density (6.13). We first note that, to simplify the spherical polar coordinate representation, instead of $\sqrt{1-\lambda}p_1 + i\sqrt{\lambda}p_2$ we may take $\sqrt{1-\lambda}p_3 + i\sqrt{\lambda}p_1$ in (6.14). Then, from (6.13) and using the notation in (6.8),

$$\begin{aligned}\rho_2^O(x,y) &= \rho_2^N(x,y) - \big(R_1^2(r)+R_2^2(r)\big)R_3^2(r)\big((1-\lambda)x_3y_3+\lambda x_1y_1\big) \\ &\quad + \tfrac{1}{2}R_3^2(r)\big[\big((1-\lambda)x_3^2+\lambda x_1^2\big)\rho^N(y) + x \leftrightarrow y\big] \\ &= \rho_2^N(x,y) - \big(R_1^2(r)+R_2^2(r)\big)R_3^2(r)\big((1-\lambda)x_3y_3+\lambda x_1y_1\big) \\ &\quad + \tfrac{1}{2}R_3^2(r)\rho^N(r)\big((1-\lambda)(x_3^2+y_3^2)+\lambda(x_1^2+y_1^2)\big).\end{aligned}$$

Inserting the Nitrogen density (6.12) and pair density (6.9) we have

$$\begin{aligned}\rho_2^O(x,y) = \tilde{F}(r) &- R_3^2(r)\big(R_1^2(r)+R_2^2(r)\big)r^2\cos\Theta - \tfrac{1}{2}R_3^4r^4\cos^2\Theta \\ &- \big(R_1^2(r)+R_2^2(r)\big)R_3^2(r)\big((1-\lambda)x_3y_3+\lambda x_1y_1\big) \\ &+ \tfrac{1}{2}R_3^2(r)\big(2R_1^2(r)+2R_2^2(r)+r^2R_3^2(r)\big)\big((1-\lambda)(x_3^2+y_3^2)+\lambda(x_1^2+y_1^2)\big),\end{aligned} \tag{6.15}$$

where $\tilde{F}$ is given in (6.11).

We now see that finding the critical points of this pair density is highly nontrivial, it being a quartic polynomial in six variables. However the fact that pair density is symmetric in x and y and suggests that there may be some symmetry in the position of critical points, which would reduce the degree of the problem.

Lemma 6.3.2. *Let $f : S_r^2 \times S_r^2 \to \mathbb{R}$, $r > 0$, be a symmetric function given by*

$$\begin{aligned}f(x,y) = \alpha_1(x_1y_1+x_2y_2+x_3y_3) &+ \alpha_2(x_1y_1+x_2y_2+x_3y_3)^2 \\ &+ \gamma_1\big(\beta_1(x_1^2+y_1^2)+\beta_2(x_2^2+y_2^2)+\beta_3(x_3^2+y_3^2)\big) \\ &+ \big(\beta_1x_1y_1+\beta_2x_2y_2+\beta_3x_3y_3\big)\big(\gamma_2+\gamma_3(x_1y_1+x_2y_2+x_3y_3)\big)\end{aligned}$$

where $x = (x_1, x_2, x_3)$*,* $y = (y_1, y_2, y_3) \in S_r^2$*,* $\alpha_i = \alpha_i(r)$*,* $\gamma_i = \gamma_i(r)$ *and* $\beta_i \in \mathbb{R}$*. Then every isolated critical point* $(a, b) = ((a_1, a_2, a_3), (b_1, b_2, b_3))$ *satisfies either* $a_i^2 = b_i^2$*,* $i = 1, 2, 3$ *or* $a \cdot b = r^2\gamma_2/(2\gamma_1 - r^2\gamma_3)$*. Furthermore, for any continuum of critical points with* $a \cdot b \neq r^2\gamma_2/(2\gamma_1 - r^2\gamma_3)$*, a representative can be chosen which satisfies* $a_i^2 = b_i^2$*,* $i = 1, 2, 3$*.*

Proof We use the method of Lagrange multipliers where our function to be maximized is f and we have two constraints, namely $g_1(x, y) = x_1^2 + x_2^2 + x_3^2 - r^2 = 0$ and $g_2(x, y) = y_1^2 + y_2^2 + y_3^2 - r^2 = 0$. Hence we wish to solve $\nabla f = \lambda_1\nabla g_1 + \lambda_2\nabla g_2$ where our gradient is over the 6-dimensional space with coordinates $\{x_1, x_2, x_3, y_1, y_2, y_3\}$. Denote an extremum of f under the constraints g_1 and g_2 by $(a, b) = ((a_1, a_2, a_3), (b_1, b_2, b_3))$. We therefore have, along with the two constraints, the six equations given by

$$\begin{aligned}
2\lambda_1 a_i &= \alpha_1 b_i + 2\alpha_2 b_i(a_1b_1 + a_2b_2 + a_3b_3) + 2\gamma_1\beta_i a_i \\
&\quad + \beta_i b_i(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)) \\
&\quad + \gamma_3 b_i(\beta_1 a_1b_1 + \beta_2 a_2b_2 + \beta_3 a_3b_3) \qquad (6.16)\\
2\lambda_2 b_i &= \alpha_1 a_i + 2\alpha_2 a_i(a_1b_1 + a_2b_2 + a_3b_3) + 2\gamma_1\beta_i b_i \\
&\quad + \beta_i a_i(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)) \\
&\quad + \gamma_3 a_i(\beta_1 a_1b_1 + \beta_2 a_2b_2 + \beta_3 a_3b_3). \qquad (6.17)
\end{aligned}$$

We now consider the two sets of equations $a_i \times (6.16)_i$ and $b_i \times (6.17)_i$ where the subscript i denotes the equation for a_i and b_i. We then have

$$\begin{aligned}
a_i^2(2\lambda_1 - 2\gamma_1\beta_i) &= a_ib_i\big[\alpha_1 + 2\alpha_2(a_1b_1 + a_2b_2 + a_3b_3) \\
&\quad + \beta_i\big(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)\big) \\
&\quad + \gamma_3(\beta_1a_1b_1 + \beta_2a_2b_2 + \beta_3a_3b_3)\big] \qquad (6.18)\\
b_i^2(2\lambda_2 - 2\gamma_1\beta_i) &= a_ib_i\big[\alpha_1 + 2\alpha_2(a_1b_1 + a_2b_2 + a_3b_3) \\
&\quad + \beta_i\big(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)\big) \\
&\quad + \gamma_3(\beta_1a_1b_1 + \beta_2a_2b_2 + \beta_3a_3b_3)\big]. \qquad (6.19)
\end{aligned}$$

Now noting that, for each i, the right hand sides of (6.18) and (6.19) are equal gives the set of three equations

$$a_i^2(\lambda_1 - \gamma_1\beta_i) = b_i^2(\lambda_2 - \gamma_1\beta_i). \qquad (6.20)$$

Summing these three equations over i and using g_1 and g_2 gives

$$r^2(\lambda_1 - \lambda_2) = \gamma_1\big(\beta_1(a_1^2 - b_1^2) + \beta_2(a_2^2 - b_2^2) + \beta_3(a_3^2 - b_3^2)\big). \tag{6.21}$$

Similarly, we construct $a_i \times (6.17)_i$ and $b_i \times (6.16)_i$ which give

$$\begin{aligned} a_ib_i(2\lambda_1 - 2\gamma_1\beta_i) = a_i^2\big[&\alpha_1 + 2\alpha_2(a_1b_1 + a_2b_2 + a_3b_3) \\ &+ \gamma_3(\beta_1a_1b_1 + \beta_2a_2b_2 + \beta_3a_3b_3) \\ &+ \beta_i\big(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)\big)\big] \end{aligned} \tag{6.22}$$

$$\begin{aligned} a_ib_i(2\lambda_2 - 2\gamma_1\beta_i) = b_i^2\big[&\alpha_1 + 2\alpha_2(a_1b_1 + a_2b_2 + a_3b_3) \\ &+ \gamma_3(\beta_1a_1b_1 + \beta_2a_2b_2 + \beta_3a_3b_3) \\ &+ \beta_i\big(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)\big)\big]. \end{aligned} \tag{6.23}$$

Taking (6.22) – (6.23) gives the set of equations

$$\begin{aligned} 2a_ib_i(\lambda_1 - \lambda_2) = (a_i^2 - b_i^2)\big[&\alpha_1 + 2\alpha_2(a_1b_1 + a_2b_2 + a_3b_3) \\ &+ \gamma_3(\beta_1a_1b_1 + \beta_2a_2b_2 + \beta_3a_3b_3) \\ &+ \beta_i(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3))\big]. \end{aligned}$$

Again, summing the set of equations over i and using g_1 and g_2 gives

$$\begin{aligned} &2(\lambda_1 - \lambda_2)(a_1b_1 + a_2b_2 + a_3b_3) \\ &= \big(\beta_1(a_1^2 - b_1^2) + \beta_2(a_2^2 - b_2^2) + \beta_3(a_3^2 - b_3^2)\big)\big(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)\big). \end{aligned} \tag{6.24}$$

From (6.21) we see that if either $\gamma_1 = 0$, or $\beta_i = 0$, $i = 1, 2, 3$ then $\lambda_1 = \lambda_2$. Else we have $\beta_1(a_1^2 - b_1^2) + \beta_2(a_2^2 - b_2^2) + \beta_3(a_3^2 - b_3^2) = r^2\frac{\lambda_1 - \lambda_2}{\gamma_1}$ and inserting this into (6.24) gives

$$2\gamma_1(\lambda_1 - \lambda_2)(a_1b_1 + a_2b_2 + a_3b_3) = r^2(\lambda_1 - \lambda_2)(\gamma_2 + \gamma_3(a_1b_1 + a_2b_2 + a_3b_3)).$$

It follows that that either $\lambda_1 = \lambda_2$ or $a_1b_1 + a_2b_2 + a_3b_3$ has a value independent of λ_i, α_i and β_i, specifically, $a_1b_1 + a_2b_2 + a_3b_3 = r^2\frac{\gamma_2}{2\gamma_1 - r^2\gamma_3}$.

Suppose $\lambda_1 = \lambda_2 = \lambda$ then (6.20) gives

$$(a_i^2 - b_i^2)(\lambda - \gamma_1\beta_i) = 0$$

for each i. If $\beta_1 = \beta_2 = \beta_3$ then f depends only on $x \cdot y$ and hence has a continuum of critical points. It is clear that we may therefore choose $a_i^2 = b_i^2$ for all i.

If $\beta_1 \neq \beta_2 \neq \beta_3$ then at most one of $\lambda - \gamma_1\beta_i$ is zero and hence at least two of $a_i^2 = b_i^2$ and then g_1 and g_2 give that $a_i^2 = b_i^2$ for all i.

The remaining case, without loss of generality, is $\beta_1 = \beta_2 \neq \beta_3$. If $\lambda - \gamma_1\beta_3 = 0$ then $a_1^2 = b_1^2$ and $a_2^2 = b_2^2$ and again g_1 and g_2 give of the $a_3^2 = b_3^2$. If $\lambda - \gamma_1\beta_1 = \lambda - \gamma_1\beta_1 = 0$ then we must have $a_3^2 = b_3^2$ and since $\beta_1 = \beta_2$ we see that f is invariant under simultaneous rotation about the e_3 axis, giving a continuum of critical points. We may then simultaneously rotate a and b about the e_3 axis to choose $a_i^2 = b_i^2$ for all i. ☐

We move on to consider the case from Lemma 6.3.2 for which there is a fixed inner product, namely $a \cdot b = r^2\gamma_2/(2\gamma_1 - r^2\gamma_3)$. Inserting this into f gives

$$\begin{aligned} f(a,b) &= \gamma_1\big(\beta_1(x_1^2+y_1^2) + \beta_2(x_2^2+y_2^2) + \beta_3(x_3^2+y_3^2)\big) \\ &\quad + (\beta_1x_1y_1 + \beta_2x_2y_2 + \beta_3x_3y_3)\tfrac{2\gamma_1\gamma_2}{2\gamma_1 - r^2\gamma_3} \\ &=: A\big(\beta_1(x_1^2+y_1^2) + \beta_2(x_2^2+y_2^2) + \beta_3(x_3^2+y_3^2)\big) \\ &\quad + B(\beta_1x_1y_1 + \beta_2x_2y_2 + \beta_3x_3y_3), \end{aligned}$$

where $a \cdot b = Br^2/(2A)$. Further note that if $A = 0$ then $B = 0$ and there is nothing to prove.

For simplicity, we restrict to the case where $\beta_1 = 1 - \lambda$, $\beta_2 = \lambda$ and $\beta_3 = 0$, i.e. a one-parameter family of pair densities where the parameters come from the simultaneous rotation of the two coordinates, as in the Oxygen case.

Applying the method of Lagrange multipliers leads to $\nabla f = \mu_1\nabla g_1 + \mu_2\nabla g_2 + \mu_3\nabla g_3$ where $g_1 = a_1^2 + a_2^2 + a_3^2 - r$, $g_2 = b_1^2 + b_2^2 + b_3^2 - r$ and $g_3 = a_1b_1 + a_2b_2 + a_3b_3 - Br^2/(2A)$:

Lemma 6.3.3. *Let $f : S_r^2 \times S_r^2 \to \mathbb{R}$, $r > 0$ be a symmetric function given by $f(x,y) = A\big((1-\lambda)(x_1^2+y_1^2) + \lambda(x_2^2+y_2^2)\big) + B\big((1-\lambda)x_1y_1 + \lambda x_2y_2\big)$ where $A \in \mathbb{R}\setminus\{0\}$, $B \in \mathbb{R}$, and $\lambda \in [0,1/2]$. Further restrict the values of f to those where $x \cdot y = r^2B/(2A)$. Then, for any critical value of f, we may choose a representative point (a,b) where $a_i = \pm b_i$, $i = 1,2,3$.*

Proof The Lagrange multiplier equations for f with the three restrictions at a critical point (a, b) give the six equations

$$2A(1-\lambda)a_1 + B(1-\lambda)b_1 = 2\mu_1 a_1 + \mu_3 b_1 \tag{6.25}$$

$$2A\lambda a_2 + B\lambda b_2 = 2\mu_1 a_2 + \mu_3 b_2 \tag{6.26}$$

$$0 = 2\mu_1 a_3 + \mu_3 b_3 \tag{6.27}$$

$$2A(1-\lambda)b_1 + B(1-\lambda)a_1 = 2\mu_2 b_1 + \mu_3 a_1 \tag{6.28}$$

$$2A\lambda b_2 + B\lambda a_2 = 2\mu_2 b_2 + \mu_3 a_2 \tag{6.29}$$

$$0 = 2\mu_2 b_3 + \mu_3 a_3. \tag{6.30}$$

Forming the sums (6.25)a_1 + (6.26)a_2 + (6.27)a_3 and (6.28)b_1 + (6.29)b_2 + (6.27)b_3, and making use of the three restrictions, gives the two equations

$$2A[(1-\lambda)a_1^2 + \lambda a_2^2] + B[(1-\lambda)a_1 b_1 + \lambda a_2 b_2] = 2\mu_1 r^2 + \mu_3 \tfrac{r^2 B}{2A} \tag{6.31}$$

$$2A[(1-\lambda)b_1^2 + \lambda b_2^2] + B[(1-\lambda)a_1 b_1 + \lambda a_2 b_2] = 2\mu_2 r^2 + \mu_3 \tfrac{r^2 B}{2A}, \tag{6.32}$$

Suppose now that $B \neq 0$, and taking (6.31) – (6.32), we have

$$2A[(1-\lambda)(a_1^2 - b_1^2) + \lambda(a_2^2 - b_2^2)] = 2(\mu_1 - \mu_2)r^2. \tag{6.33}$$

Similarly, the sums (6.25)b_1 + (6.26)b_2 + (6.27)b_3 and (6.28)a_1 + (6.29)a_2 + (6.27)a_3 give

$$2A[(1-\lambda)a_1 b_1 + \lambda a_2 b_2] + B[(1-\lambda)b_1^2 + \lambda b_2^2] = 2\mu_1 \tfrac{r^2 B}{2A} + \mu_3 r^2 \tag{6.34}$$

$$2A[(1-\lambda)a_1 b_1 + \lambda a_2 b_2] + B[(1-\lambda)a_1^2 + \lambda a_2^2] = 2\mu_2 \tfrac{r^2 B}{2A} + \mu_3 r^2, \tag{6.35}$$

and (6.34) – (6.35) gives

$$B[(1-\lambda)(b_1^2 - a_1^2) + \lambda(b_2^2 - a_2^2)] = 2(\mu_1 - \mu_2)\tfrac{r^2 B}{2A}. \tag{6.36}$$

Rearranging (6.33) and (6.36) we see that

$$(1-\lambda)(a_1^2 - b_1^2) + \lambda(a_2^2 - b_2^2) = (\mu_1 - \mu_2)\tfrac{r^2}{A} = (1-\lambda)(b_1^2 - a_1^2) + \lambda(b_2^2 - a_2^2)$$

and hence $\mu_1 = \mu_2$.

Setting $\mu_1 = \mu_2 = \mu$, the three equations given by (6.25)a_1–(6.28)b_1, (6.26)a_2–

(6.29)b_2 and (6.27)a_3 − (6.30)b_3 are

$$\begin{aligned} 2A(1-\lambda)(a_1^2-b_1^2) &= 2\mu(a_1^2-b_1^2)\\ 2A\lambda(a_2^2-b_2^2) &= 2\mu(a_2^2-b_2^2)\\ 0 &= 2\mu(a_3^2-b_3^2). \end{aligned}$$

It follows that $a_3^2 = b_3^2$ and, when $\lambda \neq 1/2$, at most one of $A(1-\lambda)-\mu$ and $A\lambda-\mu$ can be zero and hence at least one of $a_1^2 = b_1^2$ and $a_2^2 = b_2^2$ hold. The result follows from g_1 and g_2.

For the case when $\lambda = 1/2$, f is given by $\frac{A}{2}(a_1^2+b_1^2+a_2^2+b_2^2)+\frac{B}{2}(a_1b_1+a_2b_2)$ and using the three restrictions gives that $f(a,b) = \frac{A}{2}(2-a_3^2-b_3^2)+\frac{B}{2}(r^2B/(2A)-a_3b_3)$. This leads to the six equations

$$\begin{aligned} 0 &= 2\mu_1 a_1 + \mu_3 b_1, & 0 &= 2\mu_2 b_1 + \mu_3 a_1,\\ 0 &= 2\mu_1 a_2 + \mu_3 b_2, & 0 &= 2\mu_2 b_2 + \mu_3 a_2,\\ -Aa_3 - \tfrac{B}{2}b_3 &= 2\mu_1 a_3 + \mu_3 b_3, & -Ab_3 - \tfrac{B}{2}a_3 &= 2\mu_2 b_3 + \mu_3 a_3, \end{aligned}$$

and an analogous argument to the above gives

$$a_3^2 - b_3^2 = \tfrac{2r^2}{A}(\mu_1-\mu_2) = b_3^2 - a_3^2,$$

from which it follows that $\mu_1 = \mu_2$ and $a_3^2 = b_3^2$. Again following the above argument we see that $2\mu(a_1^2-b_1^2) = 0$ and $2\mu(a_2^2-b_2^2) = 0$ and the result follows for $B \neq 0$.

If $B = 0$ then forming the sum (6.25)a_1 + (6.26)a_2 + (6.27)a_3 + (6.28)b_1 + (6.29)b_2 + (6.30)b_3 and recalling that $|a|^2 = r^2$, $|b|^2 = r^2$ and $a \cdot b = 0$ gives $f(a,b) = (\mu_1+\mu_2)r^2$.

We next note that $a_i = 0$ if and only if $b_i = 0$. For example, if $a_1 = 0$ then (6.25) gives $\mu_3 b_1 = 0$ and hence $b_1 = 0$. Suppose first that $a_1 \neq 0$ and hence (6.25) and (6.28) give

$$\frac{a_1}{b_1} = \frac{\mu_3}{2(A(1-\lambda)-\mu_1)} = \frac{2(A(1-\lambda)-\mu_2)}{\mu_3}$$

and hence $\mu_3^2 = 4(A(1-\lambda)-\mu_1)(A(1-\lambda)-\mu_2)$. Similarly, for $a_2 \neq 0$ and $a_3 \neq 0$ we find $\mu_3^2 = 4(A\lambda-\mu_1)(A\lambda-\mu_2)$ and $\mu_3^2 = 4\mu_1\mu_2$.

Since $a \cdot b = 0$ it is clear that at least four of the a_i and b_i must be non-zero. If $a_1 \neq 0, a_2 \neq 0$ we find that $\lambda = 1/2$ or $\mu_1 + \mu_2 = A$. Similarly, the other two cases give $\lambda = 0$ or $\mu_1 + \mu_2 = A\lambda$ and $\mu_1 + \mu_2 = A(1-\lambda)$ respectively.

If $f(a,b) = Ar^2$ we may choose $a_1 = a_2 = b_1 = -b_2 = r/\sqrt{2}$, $a_3 = b_3 = 0$, if $f(a,b) = A(1-\lambda)r^2$ we may choose $a_1 = a_3 = b_1 = -b_3 = r/\sqrt{2}$, $a_2 = b_2 = 0$ and if $f(a,b) = A\lambda r^2$ we may choose $a_2 = a_3 = b_2 = -b_3 = r/\sqrt{2}$, $a_1 = b_1 = 0$. Hence it remains to consider the cases where $\lambda \in \{0, 1/2\}$.

If $\lambda = 1/2$ we find that $f = \frac{1}{2}(x_1^2 + y_1^2 + x_2^2 + y_2^2) = \frac{1}{2}(2r^2 - x_3^2 - y_3^2)$ and so, by symmetry of the restrictions, we need only consider $f = A(x_1^2 + y_1^2)$, which is equivalent to the $\lambda = 0$ case. We still have that $a_1 = 0$ if and only if $b_i = 0$. If $a_1 = b_1 = 0$ then we may choose $a_2 = a_3 = b_2 = -b_3 = r^2/\sqrt{2}$ and are done. Suppose that $a_1, b_1 \neq 0$ then by the above we have $\mu_3^2 = 4(A - \mu_1)(A - \mu_2)$ and $\mu_3^2 = 4\mu_1\mu_2$ (since at least one of a_2 and a_3 must be non-zero), from which it follows that $\mu_1 + \mu_2 = A$ and hence $f(a,b) = Ar^2$. Setting $a_1 = a_2 = b_1 = -b_2 = r/\sqrt{2}$ and $a_3 = b_3 = 0$ gives the result ☐

From Lemmas 6.3.2 and 6.3.3 we see that

Corollary 6.3.4 (Symmetry of Pair Density Critical Points). *Let* $f : S_r^2 \times S_r^2 \to \mathbb{R}$, $r > 0$, *be a symmetric function given by*

$$\begin{aligned} f(x,y) = {} & \alpha_1(x_1y_1 + x_2y_2 + x_3y_3) + \alpha_2(x_1y_1 + x_2y_2 + x_3y_3)^2 \\ & + \gamma_1\big((1-\lambda)(x_1^2 + y_1^2) + \lambda(x_2^2 + y_2^2)\big) \\ & + \big((1-\lambda)x_1y_1 + \lambda x_2y_2\big)\big(\gamma_2 + \gamma_3(x_1y_1 + x_2y_2 + x_3y_3)\big) \end{aligned} \tag{6.37}$$

where $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, $\alpha_i = \alpha_i(r)$, $\gamma_i = \gamma_i(r)$ *and* $\lambda \in [0, 1/2]$. *Then every isolated critical point* $(a,b) = ((a_1,a_2,a_3),(b_1,b_2,b_3))$ *satisfies* $a_i^2 = b_i^2$, $i = 1,2,3$. *Furthermore, for any continuum of critical points, a representative can be chosen that satisfies* $a_i^2 = b_i^2$, $i = 1,2,3$.

The importance of this result is that it reduces the six-variable problem to one of three variables and three parameters (which determine whether $a_i = \pm b_i$). The restriction to a sphere of radius r then allows us to use spherical polar coordinates and hence reduce the problem further to one of two variables which, as will be seen in the next section, can be solved exactly.

6.3.6 Critical Points of a General Pair Density

Using Corollary 6.3.4, changing to spherical polar coordinates, and setting $\cos\theta_1 = \epsilon_{c\theta}\cos\theta_2 = \cos\theta$, $\sin\theta_1\sin\phi_1 = \epsilon_{s\phi}, \sin\theta_2\sin\phi_2 = \sin\theta\sin\phi$, and $\sin\theta_1\cos\phi_1 = \epsilon_{c\phi}\sin\theta_2\cos\phi_2 = \sin\theta\cos\phi$ where $\epsilon_\alpha \in \{-1, 1\}$, the general form of the pair density (6.37) is given by

$$\rho_2 := \alpha_1 \sin^4\theta \sin^4\phi + \alpha_2 \sin^4\theta \sin^2\phi + \alpha_3 \sin^4\theta + \alpha_4 \sin^2\theta \sin^2\phi + \alpha_5 \sin^2\theta + \alpha_6, \tag{6.38}$$

where the α_i depend on $\epsilon_{c\theta}$, $\epsilon_{c\phi}$, $\epsilon_{s\phi}$, λ and r. This follows by noting that all powers of $\sin\theta$, $\cos\theta$, $\sin\phi$ and $\cos\phi$ are even and using the usual identity $\cos^2\varphi = 1 - \sin^2\varphi$.

We wish to find the general form for the critical points, into which we can insert the specific values of the α_i for our pair densities. The dependence on the ϵ's results in eight possible cases for each critical point.

We begin by finding the first derivatives with respect to θ and ϕ:

$$\begin{aligned} \frac{\partial\rho_2}{\partial\theta} &= 2\sin\theta\cos\theta\big[2\alpha_1\sin^2\theta\sin^4\phi + 2\alpha_2\sin^2\theta\sin^2\phi + 2\alpha_3\sin^2\theta \\ &\qquad + \alpha_4\sin^2\phi + \alpha_5\big], & (6.39) \\ \frac{\partial\rho_2}{\partial\theta} &= 2\sin\phi\cos\phi\big[2\alpha_1\sin^4\theta\sin^2\phi + \alpha_2\sin^4\theta + \alpha_4\sin^2\theta\big]. & (6.40) \end{aligned}$$

We do not calculate the second derivatives at this point, one reason for this is that they will depend strongly on the exact values of α_i and it may well be easier to compare the resulting densities directly. Another reason for using this method instead of, say, the Hessian, is that in order to find the global maximum we would still have to directly compare the densities of the resulting maxima. Consider first (6.39), which is clearly zero if and only if

$$\cos\theta = 0, \quad \sin\theta = 0, \quad \text{or } \sin^2\theta = -\frac{\alpha_4\sin^2\phi + \alpha_5}{2(\alpha_1\sin^4\phi + \alpha_2\sin^2\phi + \alpha_3)}.$$

The first two cases are always possible solutions whereas the third case is only a solution if α_i are such that the value lies in $[0, 1]$. Similarly, for (6.40) we find that the critical points lie at

$$\cos\phi = 0, \quad \sin\phi = 0, \quad \text{and } \sin^2\phi = -\frac{\alpha_2\sin^2\theta + \alpha_4}{2\alpha_1\sin^2\theta},$$

where once again the first two cases are always valid and there is a restriction on the α_i for the third to be valid, i.e. to lie in $[0, 1]$.

It remains to calculate the pair densities at each of the critical points (of which there are clearly nine, although the three cases with $\sin^2\theta = 0$ are degenerate, leading to seven cases). These are given in Table 6.2. The only non-trivial case to evaluate is when the third case for each of θ and ϕ are chosen but the pair of

$\sin^2\theta$	$\sin^2\phi$	ρ_2
0	-	α_6
1	0	$\alpha_3+\alpha_5+\alpha_6$
1	1	$\alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6$
$-\frac{\alpha_5}{2\alpha_3}$	0	$-\frac{\alpha_5^2}{4\alpha_3}+\alpha_6$
$-\frac{\alpha_4+\alpha_5}{2(\alpha_1+\alpha_2+\alpha_3)}$	1	$-\frac{(\alpha_4+\alpha_5)^2}{4(\alpha_1+\alpha_2+\alpha_3)}+\alpha_6$
1	$-\frac{\alpha_2+\alpha_4}{2\alpha_1}$	$-\frac{(\alpha_2+\alpha_4)^2}{4\alpha_1}+\alpha_3+\alpha_5+\alpha_6$
$\frac{\alpha_2\alpha_4-2\alpha_1\alpha_5}{4\alpha_1\alpha_3-\alpha_2^2}$	$\frac{2\alpha_3\alpha_4-\alpha_2\alpha_5}{2\alpha_1\alpha_5-\alpha_2\alpha_4}$	$\frac{-\alpha_1\alpha_5^2-\alpha_3\alpha_4^2+\alpha_2\alpha_4\alpha_5}{4\alpha_1\alpha_3-\alpha_2^2}+\alpha_6$

Table 6.2: The critical points of a general pair density as given in (6.38) along with the value of ρ_2 at those points.

equations is easily solved to give the simultaneous values in the table.

We now have a list of all possible maxima for the pair density and it remains to insert the specific values for the α_i for the ground state of Oxygen.

6.3.7 Global Maximum of the Oxygen Pair Density

Using Corollary 6.3.4 and (6.15) we have

$$\begin{aligned}
\rho_2^O(x,y) = A_0(r)& \\
&+ A_1(r)\big[2\cos^2\theta - 2\epsilon_{c\theta}\cos^2\theta - \sin^2\theta(\epsilon_{c\phi}\cos^2\phi + \epsilon_{s\phi}\sin^2\phi) \\
&\qquad + \lambda(-2\cos^2\theta + \epsilon_{c\theta}\cos^2\theta + 2\sin^2\theta\sin^2\phi - \epsilon_{s\phi}\sin^2\theta\sin^2\phi)\big] \\
&+ A_2(r)\big[2\cos^2\theta - \cos^4\theta - 2\epsilon_{c\theta}\cos^2\theta\sin^2\theta(\epsilon_{c\phi}\cos^2\phi + \epsilon_{s\phi}\sin^2\phi) \\
&\qquad - \sin^4\theta(\epsilon_{c\phi}\cos^2\phi + \epsilon_{s\phi}\sin^2\phi)^2 + \lambda(-2\cos^2\theta + 2\sin^2\theta\sin^2\phi)\big],
\end{aligned}$$

where $A_0(r) := \tilde{F}(r)$ (given by (6.11)), $A_1(r) := r^2R_3^2(r)(R_1^2(r) + R_2^2(r))$, and $A_2(r) := \frac{1}{2}r^4R_3^4(r)$. Rewriting this in terms of only $\sin\theta$ and $\sin\phi$, as in (6.38), we have

$$\begin{aligned}
&\rho_2^O(x,y) = A_0(r) \\
&+ A_1(r)\big[2(1-\epsilon_{c\theta}) + 2(\epsilon_{c\theta}-1)\sin^2\theta - \sin^2\theta(\epsilon_{c\phi} + (\epsilon_{s\phi}-\epsilon_{c\phi})\sin^2\phi) \\
&\qquad + \lambda(-2 + 2\sin^2\theta + \epsilon_{c\theta} - \epsilon_{c\theta}\sin^2\theta + (2-\epsilon_{s\phi})\sin^2\theta\sin^2\phi)\big] \\
&+ A_2(r)\big[1 - \sin^4\theta - 2\epsilon_{c\theta}\sin^2\theta(1-\sin^2\theta)(\epsilon_{c\phi} + (\epsilon_{s\phi}-\epsilon_{c\phi})\sin^2\phi) \\
&\qquad - \sin^4\theta(\epsilon_{c\phi} + (\epsilon_{s\phi}-\epsilon_{c\phi})\sin^2\phi)^2 + 2\lambda(-1 + \sin^2\theta + \sin^2\theta\sin^2\phi)\big].
\end{aligned}$$

Ignoring $A_0(r)$, which is identical for each critical point, this gives

$$\begin{aligned}
\alpha_1 &= -A_2(\epsilon_{s\phi} - \epsilon_{c\phi})^2, \\
\alpha_2 &= 2A_2(\epsilon_{s\phi} - \epsilon_{c\phi})(\epsilon_{c\theta} - \epsilon_{c\phi}), \\
\alpha_3 &= 2A_2(\epsilon_{c\theta}\epsilon_{c\phi} - 1), \\
\alpha_4 &= A_1\big(- (\epsilon_{s\phi} - \epsilon_{c\phi}) + \lambda(2 - \epsilon_{s\phi})\big) + A_2\big(- 2\epsilon_{c\theta}(\epsilon_{s\phi} - \epsilon_{c\phi}) + 2\lambda\big), \\
\alpha_5 &= A_1\big(- 2 + 2\epsilon_{c\theta} - \epsilon_{c\phi} + \lambda(2 - \epsilon_{c\theta})\big) + 2A_2(-\epsilon_{c\theta}\epsilon_{c\phi} + \lambda), \\
\alpha_6 &= A_1\big(2 - 2\epsilon_{c\theta} + \lambda(-2 + \epsilon_{c\theta})\big) + A_2(1 - 2\lambda),
\end{aligned}$$

and inserting these values into the critical points of Table 6.2 we obtain the following critical values:

- $\rho_2^O = \alpha_6 = A_1(r)[2 - 2\epsilon_{c\theta} + \lambda(-2 + \epsilon_{c\theta})] + A_2(r)[1 - 2\lambda].$

This is clearly maximized when $\epsilon_{c\theta} = -1$ (since $\theta = 0$, the pair density is independent of ϕ and hence also of $\epsilon_{c\phi}$ and $\epsilon_{s\phi}$) and gives a pair density of

$$\rho_2^{O,1} := A_1(r)[4 - 3\lambda] + A_2(r)[1 - 2\lambda]. \tag{6.41}$$

- $\rho_2^O = \alpha_3 + \alpha_5 + \alpha_6 = -A_1(r)\epsilon_{c\phi} - A_2(r).$

This is clearly maximized when $\epsilon_{c\phi} = -1$ (and is independent of both $\epsilon_{c\theta}$ and $\epsilon_{s\phi}$) and gives a pair density of

$$\rho_2^{O,2} := A_1(r) - A_2(r).$$

- $\rho_2^O = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 = A_1(r)[2\lambda - \epsilon_{s\phi}(1 + \lambda)] + A_2(r)[2\lambda - 1],$

where we have used the fact that $\epsilon_\alpha^2 = 1$.

This is clearly maximized when $\epsilon_{s\phi} = -1$ (and is independent of $\epsilon_{c\theta}$ and $\epsilon_{c\phi}$) and gives

$$\rho_2^{O,3} := A_1(r)[1 + 3\lambda] + A_2(r)[-1 + 2\lambda].$$

- $\rho_2^O = -\dfrac{\alpha_5^2}{4\alpha_3} + \alpha_6.$

After inserting the 8 possible combinations for the ϵ's we find that there are two possible pair densities:

$$\begin{aligned}\rho_2^{O,4,1} &= \tfrac{1}{16A_2(r)}\left[A_1(r)(-5+3\lambda)+A_2(r)(2+2\lambda)\right]^2 \\ &\quad + A_1(r)[4-3\lambda]+A_2(r)[1-2\lambda], \qquad (6.42)\\ \rho_2^{O,4,2} &= \tfrac{1}{16A_2(r)}\left[A_1(r)(1+\lambda)+A_2(r)(2+2\lambda)\right]^2 \\ &\quad + A_1(r)[-\lambda]+A_2(r)[1-2\lambda], \qquad (6.43)\end{aligned}$$

which are valid for $f_{4,i} \in [0,1]$ where

$$\begin{aligned} f_{4,1} &= \frac{1+\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{5-3\lambda}{8}, \qquad (6.44)\\ f_{4,2} &= \frac{1+\lambda}{4} + \frac{A_1(r)}{A_2(r)}\frac{1+\lambda}{8}. \qquad (6.45)\end{aligned}$$

The $f_{4,i}$ correspond to $\sin^2\theta$ in Table 6.2. The first case corresponds to $(\epsilon_{c\theta},\epsilon_{c\phi},\epsilon_{s\phi}) = (-1,1,\pm1)$ and the second to $(1,-1,\pm1)$, the other four possibilities giving $\alpha_3 = 0$.

- $\rho_2^O = -\dfrac{(\alpha_4+\alpha_5)^2}{4(\alpha_1+\alpha_2+\alpha_3)} + \alpha_6.$

After inserting the 8 possible combinations for the ϵ's we find that there are again two possible pair densities:

$$\begin{aligned}\rho_2^{O,5,1} &= \tfrac{1}{16A_2(r)}\left[A_1(r)(-5+4\lambda)+A_2(r)(2+4\lambda)\right]^2 \\ &\quad + A_1(r)[4-3\lambda]+A_2(r)[1-2\lambda], \qquad (6.46)\\ \rho_2^{O,5,2} &= \tfrac{1}{16A_2(r)}\left[A_1(r)(1+4\lambda)+A_2(r)(2+4\lambda)\right]^2 \\ &\quad + A_1(r)[-\lambda]+A_2(r)[1-2\lambda], \qquad (6.47)\end{aligned}$$

which are valid for $f_{5,i} \in [0,1]$ where

$$\begin{aligned} f_{5,1} &= \frac{1+2\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{5-4\lambda}{8}, \qquad (6.48)\\ f_{5,2} &= \frac{1+2\lambda}{4} + \frac{A_1(r)}{A_2(r)}\frac{1+4\lambda}{8}. \qquad (6.49)\end{aligned}$$

The first case corresponds to $(\epsilon_{c\theta},\epsilon_{c\phi},\epsilon_{s\phi}) = (-1,\pm1,1)$ and the second to $(1,\pm1,-1)$, the other four possibilities giving $\alpha_1+\alpha_2+\alpha_3 = 0$.

- $\rho_2^O = -\dfrac{(\alpha_2+\alpha_4)^2}{2\alpha_1} + \alpha_3+\alpha_5+\alpha_6.$

Once again inserting the 8 possible combinations for the ϵ's results in two possible pair densities:

$$\rho_2^{O,6,1} = \tfrac{1}{16A_2(r)}\left[A_1(r)(-2+\lambda)+A_2(r)(4+2\lambda)\right]^2 + A_1(r) - A_2(r), \tag{6.50}$$

$$\rho_2^{O,6,2} = \tfrac{1}{16A_2(r)}\left[A_1(r)(2+3\lambda)+A_2(r)(4+2\lambda)\right]^2 - A_1(r) - A_2(r), \tag{6.51}$$

which are valid for $f_{6,i} \in [0,1]$ where

$$f_{6,1} = \frac{2+\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{(2-\lambda)}{8}, \tag{6.52}$$

$$f_{6,2} = \frac{2+\lambda}{4} + \frac{A_1(r)}{A_2(r)}\frac{2+3\lambda}{8}. \tag{6.53}$$

Note that for this case $f_{6,i}$ correspond to $\sin^2\phi$. The first case corresponds to $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (\pm 1, -1, 1)$ and the second to $(\pm 1, 1, -1)$, the other four possibilities giving $\alpha_1 = 0$.

- $\sin^2\theta = \dfrac{\alpha_2\alpha_4 - 2\alpha_1\alpha_5}{4\alpha_1\alpha_3 - \alpha_2^2}, \quad \sin^2\phi = \dfrac{2\alpha_3\alpha_4 - \alpha_2\alpha_5}{2\alpha_1\alpha_5 - \alpha_2\alpha_4}.$

We note that there are no choices of ϵ_α which give a valid pair density in this case. This can easily be seen by considering the denominator of $\sin^2\theta$ which is identically zero.

We now wish to prove the following result on the global maximum of the Oxygen pair density:

Theorem 6.3.5 (Maximum of Oxygen Pair Density). *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define* $A_0(r) := R_1^4(r) + 4R_1^2(r)R_2^2(r) + R_2^4(r) + 2(R_1^2(r) + R_2^2(r))R_3^2(r) + 1/2\, R_3^4(r)$, $A_1(r) := (R_1^2(r) + R_2^2(r))R_3^2(r)$, $A_2(r) := 1/2\, R_3^4(r)$. *Further denote*

$$\begin{aligned}\rho_2^{O,1}(r,\lambda) &:= A_0(r) + A_1(r)[4-3\lambda] + A_2(r)[1-2\lambda],\\ \rho_2^{O,5,1}(r,\lambda) &:= A_0(r) + \tfrac{1}{16A_2(r)}\left[A_1(r)(-5+4\lambda)+A_2(r)(2+4\lambda)\right]^2\\ &\quad + A_1(r)[4-3\lambda] + A_2(r)[1-2\lambda],\end{aligned}$$

and $f(r,\lambda) := \frac{1+2\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{5-4\lambda}{8}$. *Then the global maximum of the Oxygen pair density is given by*

$$\rho_2^{O,max}(r,\lambda) = \begin{cases} \rho_2^{O,5,1} & \text{if } f(r,\lambda) \in (0,1] \\ \rho_2^{O,1} & \text{else.}\end{cases}$$

In order to prove the theorem it is sufficient to determine the maximum of the critical points $\rho_2^{O,1}$-$\rho_2^{O,6,2}$ as given above. This then gives the maximum of the part of the pair density which has angular dependence, and the purely radial part is given by $A_0(r)$.

We now move on to compare these critical points in order to find the global maximum. Since $\rho_2^{O,5,1} - \rho_2^{O,1} = \frac{1}{16A_2(r)}\left[A_1(r)(-5+4\lambda) + A_2(r)(2+4\lambda)\right]^2 = 4A_2(r)f^2(r,\lambda)$ it is clear that, whenever it exists, $\rho_2^{O,5,1} > \rho_2^{O,1}$.

We have that $f_{5,1} \leq (1+2\lambda)/4 \leq 1$ and hence for $\rho_2^{O,5,1}$ not to exist we must have $f_{5,1} < 0$, which gives the case $\rho_2^{O,max} = \rho_2^{O,1} = A_1(r)[4-3\lambda] + A_2(r)[1-2\lambda]$.

Comparing the critical points that always exist gives

$$\rho_2^{O,1} - \rho_2^{O,2} = A_1(r)(3-3\lambda) + A_2(r)(2-2\lambda) > 0, \tag{6.54}$$

$$\rho_2^{O,1} - \rho_2^{O,3} = A_1(r)(3-6\lambda) + A_2(r)(2-4\lambda) \geq 0, \tag{6.55}$$

and hence, except when $\lambda = 1/2$, neither $\rho_2^{O,2}$ nor $\rho_2^{O,3}$ are the global maximum. The degeneracy of $\rho_2^{O,1}$ and $\rho_2^{O,3}$ when $\lambda = 1/2$ is due to the invariance of the pair density under rotation about the e_2 axis. Since, for $\rho_2^{O,1}$ we have $\epsilon_{c\theta} = -1$ and for $\rho_2^{O,3}$ we have $\epsilon_{s\phi} = -1$, both cases have the electrons opposite each other in the e_1-e_3 plane and hence give the same pair density.

In order to compare $\rho_2^{O,1}$ with the critical points that have an angular dependence we use the following lemma:

Lemma 6.3.6. *Suppose that* $A_i(r) \geq 0$ *and let*

$$\begin{aligned}
\rho_2^{(1)} &= e_1 A_1(r) + e_2 A_2(r),\\
\rho_2^{(2)} &= -\frac{\big(\beta_2 A_1(r) + \gamma_2 A_2(r)\big)^2}{2\alpha_2 A_2(r)} + g_1 A_1(r) + g_2 A_2(r),\\
f_2 &= -\frac{\beta_2 A_1(r) + \gamma_2 A_2(r)}{\alpha_2 A_2(r)}, \quad \textit{and}\\
f_3 &= -\frac{\beta_3 A_1(r) + \gamma_3 A_2(r)}{\alpha_3 A_2(r)}.
\end{aligned}$$

Suppose further that $-\alpha_2 A_2(r) > 0$*,* $-\alpha_3 A_2(r) > 0$*,* $\gamma_3 > 0$*,* $f_2 \in [0,1]$*,* $f_3 < 0$*, and* $2e_2 - (\gamma_2 + 2g_2) \leq 0$*. If*

$$h := \big(2e_1 - (\beta_2 + 2g_1)\big)\gamma_3 - \big(2e_2 - (\gamma_2 + 2g_2)\big)\beta_3 \geq 0$$

it follows that $\rho_2^{(1)} > \rho_2^{(2)}$.

Proof Since $f_2 \geq 0$ and $-\alpha_2 A_3(r) > 0$ it follows that $\beta_2 A_1(r) + \gamma_2 A_2(r) \geq 0$.

Hence, since $f_2 \leq 1$, $\rho_2^{(2)}$ is bounded above by

$$\rho_2^{(2)} \leq \tfrac{1}{2}\big((\beta_2 + 2g_1)A_1(r) + (\gamma_2 + 2g_2)A_2(r)\big).$$

We then have

$$2(\rho_2^{(1)} - \rho_2^{(2)}) \geq \big(2e_1 - (\beta_2 + 2g_1)\big)A_1(r) + \big(2e_2 - (\gamma_2 + 2g_2)\big)A_2(r). \tag{6.56}$$

Now, since $f_3 < 0$, $-\alpha_3 A_2(r) > 0$, and $\gamma_3 > 0$, we have

$$A_2(r) < -\frac{\beta_3 A_1(r)}{\gamma_3},$$

and since $2e_2 - (\gamma_2 + 2g_2) < 0$ it follows that

$$\big(2e_2 - (\gamma_2 + 2g_2)\big)A_2(r) > -\big(2e_2 - (\gamma_2 + 2g_2)\big)\frac{\beta_3 A_1(r)}{\gamma_3}.$$

Inserting this lower bound into γ_3 times (6.56), and recalling that $\gamma_3 > 0$, gives that

$$2\gamma_3(\rho_2^{(1)} - \rho_2^{(2)}) > hA_1(r),$$

where h is as in the statement of the lemma, and the result follows. □

When applying this lemma we will have that $\rho_2^{(1)} = \rho_2^{O,1}$, $f_3 = f_{5,1}$ and $\rho_2^{(2)}$ and f_2 the critical point to be tested. Hence, from (6.41), we have that $e_1 = 4 - 3\lambda$ and $e_2 = 1 - 2\lambda$. It is clear from (6.46) and (6.48) that $\gamma_3 = 2 + 4\lambda > 0$, $-\alpha_2 A_2(r) = 8A_2(r)$, which is positive whenever $A_2(r)$ is positive, and $\beta_3 = -5 + 4\lambda$. It remains to check the conditions that depend on the exact form of the critical point to be tested.

For $\rho_2^{O,4,1}$ and $\rho_2^{O,4,2}$, equations (6.42)-(6.45) give for both cases $\alpha_2 = -8$, $\gamma_2 = 2 + 2\lambda$, and $g_2 = 1 - 2\lambda$. This gives that $-\alpha_2 A_2(r) > 0$ and $2e_2 - (\gamma_2 + 2g_2) = -2 - 2\lambda < 0$. For $\rho_2^{O,4,1}$ we have $\beta_2 = -5 + 3\lambda$ and $g_1 = 4 - 3\lambda$, giving $h = 4\lambda(3 - \lambda) \geq 0$ and hence $\rho_2^{O,1} > \rho_2^{O,4,1}$. For $\rho_2^{O,4,2}$ we find $\beta_2 = 1 + 2\lambda$ and $g_1 = -\lambda$, giving $h = 4 + 4\lambda(4 - 3\lambda) > 0$ and hence $\rho_2^{O,1} > \rho_2^{O,4,2}$.

Moving on to $\rho_2^{O,5,2}$ we have from (6.46) and (6.48) that $\alpha_2 = -8$, $\beta_2 = 1 + 4\lambda$, $\gamma_2 = 2 + 4\lambda$, $g_1 = -\lambda$ and $g_2 = 1 - 2\lambda$. This gives that $-\alpha_2 A_2(r) > 0$ and $2e_2 - (\gamma_2 + 2g_2) = -2 - 4\lambda < 0$ and $h = 4 - 16\lambda^2 \geq 0$ and hence $\rho_2^{O,1} > \rho_2^{O,5,2}$.

Finally, for $\rho_2^{O,6,1}$ and $\rho_2^{O,6,2}$, equations (6.50)-(6.53) give for both cases $\alpha_2 = -8$, $\gamma_2 = 4+2\lambda$, and $g_2 = -1$. This gives that $-\alpha_2 A_2(r) > 0$ and $2e_2-(\gamma_2+2g_2) = -6\lambda \leq 0$. For $\rho_2^{O,6,1}$ we have $\beta_2 = -2+\lambda$ and $g_1 = 1$, giving $h = 4(4-3\lambda-\lambda^2) > 0$ and hence $\rho_2^{O,1} > \rho_2^{O,6,1}$. For $\rho_2^{O,6,2}$ we find $\beta_2 = 2+3\lambda$ and $g_1 = -1$, giving $h = 4(4-4\lambda-3\lambda^2) > 0$ and hence $\rho_2^{O,1} > \rho_2^{O,6,2}$.

We have therefore shown that, when $\rho_2^{O,max} = \rho_2^{O,1}$, this is indeed the global maximum. It remains to consider the case when $\rho_2^{O,max} = \rho_2^{O,5,1}$. Since in this case $\rho_2^{O,5,1} > \rho_2^{O,1}$ it follows from (6.54) and (6.55) that $\rho_2^{O,5,1} > \rho_2^{O,2}$ and $\rho_2^{O,5,1} > \rho_2^{O,3}$.

The final part of the analysis is to compare $\rho_2^{O,5,1}$ with the five cases which have an existence condition. The first stage of this is to note that, since $A_2(r) > 0$, $16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,\alpha})$ has the same sign as $\rho_2^{O,5,1} - \rho_2^{O,\alpha}$. For the five cases we find

$$\begin{aligned}
16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,4,1}) &= [A_1(r)(-5+4\lambda) + A_2(r)(2+4\lambda)]^2 \\
&\quad - [A_1(r)(-5+3\lambda) + A_2(r)(2+2\lambda)]^2, \\
16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,4,2}) &= (24 - 42\lambda + 15\lambda^2)A_1^2(r) + 4\lambda(2+3\lambda)A_2^2(r) \\
&\quad + (40 - 64\lambda + 28\lambda^2)A_1(r)A_2(r) > 0, \\
16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,5,2}) &= 24(1-2\lambda)A_1^2(r) + 40(1-2\lambda)A_1(r)A_2(r) \geq 0, \\
16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,6,1}) &= (21 - 36\lambda + 15\lambda^2)A_1^2(r) + 4(5 - 8\lambda + 3\lambda^2)A_2^2(r) \\
&\quad + 4(11 - 18\lambda + 7\lambda^2)A_1(r)A_2(r) > 0, \\
16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,6,2}) &= (21 - 52\lambda + 7\lambda^2)A_1^2(r) + 4(5 - 8\lambda + 3\lambda^2)A_2^2(r) \\
&\quad + 4(11 - 26\lambda + 5\lambda^2)A_1(r)A_2(r).
\end{aligned}$$

The degeneracy between $\rho_2^{O,5,1}$ and $\rho_2^{O,5,2}$ when $\lambda = 1/2$ is due to the invariance of the pair density under simultaneous rotation about the e_2 axis. Since $\sin^2\phi = 1$ for each case, it is clear that all points lie in the e_1-e_3 plane. Furthermore the angle between the points is the same in each case, leading to the same pair density.

The fact that the angle is the same follows by considering that, when $\lambda = 0$, we have $\sin^2\theta_{5,1} = 1/2 - 3A_1(r)/(8A_2(r))$ and $\sin^2\theta_{5,2} = 1/2 - 3A_1(r)/(8A_2(r))$. However, since for $\rho_2^{O,5,1}$ we have $\epsilon_{c\theta} = -1$ and $\epsilon_{s\phi} = 1$ and for $\rho_2^{O,5,2}$ we have $\epsilon_{c\theta} = 1$ and $\epsilon_{s\phi} = -1$, this means that the two angles are given by $\Theta_1 = \pi - 2\theta_{5,1}$ and $\Theta_2 = 2\theta_{5,2}$, where $\theta_{5,i} \in [0, \pi/2]$.

We then have that, for $\alpha = 3A_1(r)/(8A_2(r))$,

$$\begin{aligned}\cos(\theta_{5,1} + \theta_{5,2}) &= \cos\theta_{5,1}\cos\theta_{5,2} - \sin\theta_{5,1}\sin\theta_{5,2} \\ &= \sqrt{1/2 - \alpha}\sqrt{1/2 + \alpha} - \sqrt{1/2 + \alpha}\sqrt{1/2 - \alpha} = 0,\end{aligned}$$

and hence $\theta_{5,1}+\theta_{5,2} = \pi/2$, from which it follows that $\Theta_1 - \Theta_2 = \pi - 2(\theta_{5,1}+\theta_{5,2}) = 0$. This degeneracy actually extends further since, as previously noted, the pair density is invariant under rotation about the e_2 axis and hence there is an S^1 of global maxima. When later plotting the pair densities we will use the unique continuation with respect to λ of the global maximum.

We now need only consider $\rho_2^{O,4,1}$ and $\rho_2^{O,6,2}$. For the case of $\rho_2^{O,4,1}$ we see that $\rho_2^{O,5,1} - \rho_2^{O,4,1} > 0$ if and only if

$$\left(\frac{A_1(r)(-5+4\lambda) + A_2(r)(2+4\lambda)}{A_1(r)(-5+3\lambda) + A_2(r)(2+2\lambda)}\right)^2 \geq 1.$$

The identity

$$\frac{A_1(r)(-5+4\lambda) + A_2(r)(2+4\lambda)}{A_1(r)(-5+3\lambda) + A_2(r)(2+2\lambda)} = 1 + \lambda\frac{A_1(r) + 2A_2(r)}{A_1(r)(-5+3\lambda) + A_2(r)(2+2\lambda)}$$

shows that this holds when $A_1(r)(-5+3\lambda)+A_2(r)(2+2\lambda) > 0$, which is precisely the condition required for $\rho_2^{O,4,1}$ to exist. Thus $\rho_2^{O,5,1} \geq \rho_2^{O,4,1}$ whenever they both exist, with equality only when $\lambda = 0$.

The degeneracy in this case is due to the fact that, when $\lambda = 0$, the pair density is invariant under simultaneous rotation about the e_3 axis. Both cases have $\epsilon_{c\theta} = -1$ and hence the value of α_6 in each pair density is the same, and furthermore both have the same symmetry of ϕ. The values of $\sin^2\theta$ are the same (or the pair densities would be different) and hence they are equivalent up to a simultaneous rotation by $\pi/2$ around the e_3 axis. Once again there is actually an S^1 of global maxima in this case.

The remaining case is $\rho_2^{O,6,2}$. We first note that $-A_1(r)A_2(r) \geq -\frac{1}{2}(A_1^2(r) + A_2^2(r))$ and hence

$$16A_2(r)(\rho_2^{O,5,1} - \rho_2^{O,6,2}) \geq (43 - 104\lambda + 17\lambda^2)A_1^2(r) + (42 - 84\lambda + 22\lambda^2)A_2^2(r).$$

From (6.53) we see that, in order for $\rho_2^{O,6,2}$ to exist we require that $A_2(r) \geq A_1(r)\frac{2+3\lambda}{4-2\lambda}$. Noting that the coefficient of the $A_2^2(r)$ term in the above is positive and inserting this lower bound gives

$$\begin{aligned}16A_2(r)(4-2\lambda)^2(\rho_2^{O,5,1}-\rho_2^{O,6,2})\\ \geq 2(428-1092\lambda+783\lambda^2-590\lambda^2+133\lambda^4)A_1^2(r),\end{aligned}$$

and it remains to determine the sign of

$$h := 428-1092\lambda+783\lambda^2-590\lambda^2+133\lambda^4.$$

Differentiating with respect to λ gives that

$$\begin{aligned}\frac{dh}{d\lambda} &= -1092+1566\lambda-1770\lambda^2+523\lambda^3\\ &\leq -1092+1566/2+523/8=-1949/8<0,\end{aligned}$$

and hence the minimum value of h is at $\lambda = 1/2$ and is given by $197/16 > 0$. It therefore follows that $\rho_2^{O,5,1} > \rho_2^{O,6,2}$, completing the proof of Theorem 6.3.5.

6.3.8 H$_2$O Bond Angle

We now wish to use Proposition 6.1.1 to predict the bond angle of the water molecule. We see that both cases of $\rho_2^{O,max}$ of Theorem 6.3.5 have $\epsilon_{c\theta} = -1$ and further, $\rho_2^{O,1}$ is independent of ϕ so the values of $\epsilon_{c\phi}$ and $\epsilon_{s\phi}$ are irrelevant in this case. We recall that $\rho_2^{O,5,1}$ has $\cos\phi = 0$ and thus the pair density is independent of the choice of $\epsilon_{c\phi}$, and we have that $\epsilon_{s\phi} = 1$. It follows that the angle between the electrons is given, in both cases, by $\pi - 2\theta$ where θ is obtained from

$$\sin^2\theta := \max\Big\{0, \tfrac{1+2\lambda}{4} - \tfrac{A_1(r)}{A_2(r)}\tfrac{5-4\lambda}{8}\Big\},$$

and hence the bond angle is given by

$$\Theta = \pi - 2\arcsin\Big(\max\Big\{0, \tfrac{1+2\lambda}{4} - \tfrac{A_1(r)}{A_2(r)}\tfrac{5-4\lambda}{8}\Big\}\Big)^{1/2}. \tag{6.57}$$

For fixed values of r, the angle between the electrons, as given by (6.57), is monotonically decreasing with λ. To prove this, note that, by symmetry, we may restrict to $\theta_1 \in [0,\pi]$ and then, since $\sin\theta_1 = \sin\theta_2$ and $\cos\theta_1 = -\cos\theta_2$, we must have $\theta_i \in [0, \frac{\pi}{2}]$ and $\theta_j \in [\frac{\pi}{2}, \pi]$. Therefore, without loss of generality, $\theta \in [0, \frac{\pi}{2}]$. It is clear that $\frac{1+2\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{5-4\lambda}{8}$ is monotonically increasing with λ, hence $\sin^2\theta$ is also monotonically increasing with λ, and by the restriction to $[0, \frac{\pi}{2}]$, so is θ. Therefore $\Theta = \pi - 2\theta$ is monotonically decreasing with λ.

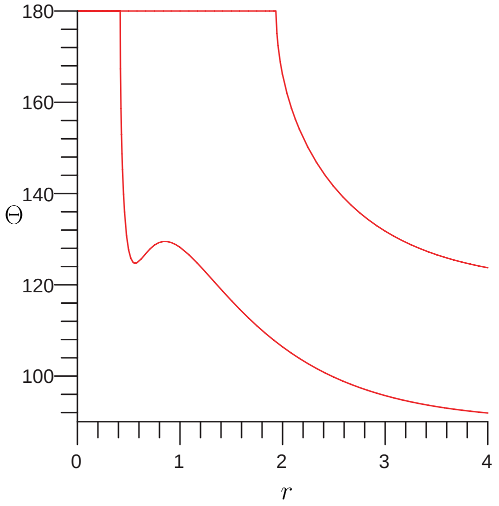


Figure 6.5: Bond angle range for Oxygen, the upper curve shows the bond angle for $\lambda = 0$, the lower curve for $\lambda = 1/2$ and thus, by the monotonicity of the bond angle, the bond range at a given r is the interval between the two curves.

The main use of this result is that we may easily determine the range over which the bond angle may freely vary with no penalty to the energy (which is independent of λ). The bond angle range is plotted in Figure 6.5. One interesting aspect of this plot is that, for suitable fixed λ, the bond angle is non-monotone with respect to r. This is analogous to the Nitrogen case in Figure 6.3 and again is due to the non-monotonicity of the radial functions $R_i(r)$.

Further we note that, since $A_2(r)$ decays exponentially slower than $A_1(r)$, the bond angle range as $r \to \infty$ is given by

$$\Theta = \pi - 2 \arcsin \left(\max \left\{ 0, \tfrac{1+2\lambda}{4} \right\} \right)^{1/2}.$$

By the monotonicity we see that the extremal values are at $\lambda = 0$ and $\lambda = 1/2$, giving $\Theta \in [90°, 120°]$.

6.3.9 Comparison with experimental data

We wish to compare our calculated values for the bond angle in water to the experimentally measured values. As discussed earlier we cannot give a precise value of r at which to test the bond angle range. However, we will use the position of the outer maximum of the radial distribution as an indicative value. These values are given in Table 5.2. The experimental bond lengths and bond angles are shown in Table 6.1.

We wish to find the predicted range of angles for the water molecule and so use (6.57) with $\lambda \in \{0, 1/2\}$ and $r = 0.897$ giving an angle range of $\theta \in [129.4^\circ, 180^\circ]$. The lower bound is obviously still quite a way from the experimental result. However, we have already noted that the choice of r is in no way expected to be ideal, and we have not taken into account the interaction of the bonded atoms.

One interesting thing of note is that the traditional method of predicting bond angles finds a maximum at 90° and then justifies the fact that it is more than 90° by claiming that the Hydrogen atoms repel each other. (See, for example, [AdP01]). As can be seen from the above, our method predicts a bond angle which is larger than the experimental data. In order to understand this it would be necessary to perform a full multi-centre calculation and determine the attraction or repulsion between the Hydrogen atoms.

In section 6.1.2, we noted that, whilst H_2O is strongly bent, Li_2O is a linear molecule. Whilst the bond length is different to that of water (experimentally it is 3.035au [BB01]), we would expect the position of the electrons to be reasonably well modelled by taking half the bond length. (In fact, due to the large positive dipole towards the Li atoms, it could be argued that we should take a position at less than half the bond length). Figure 6.5 shows that, at these sort of distances, a suitable ground state wavefunction exists such that the pair density has its global maximum with an angle of 180°.

6.3.10 Plots of the Oxygen Pair Density

We now use the bond angle information from (6.57) to plot some pair densities for Oxygen. As with the Nitrogen case, we may fix the position of one of the electrons to be at a maximizing point and then plot the pair density as a function of the position of the other electron. In the plots below we follow the convention that the black dot is the position of the fixed electron (in its optimal position) and the pair density is denoted by the colour of the sphere, red being the highest and blue being the lowest.

We begin by plotting the two extremal cases for a range of values of r. The values have been chosen so that the lowest value of r gives a point close to the nucleus where we would expect the s-orbitals to be dominant, the middle value of r gives a 'bonding' distance, the final value shows the pair density at very long range and should be a good approximation to the limit as $r \to \infty$. Three views of each pair density are shown, the 'side' view aims to show all the maxima and minima as well as the position of the fixed electron, the 'front' view is looking along the positive x-axis (the side on which the fixed electron is chosen to sit) and the 'back' view shows the view along the negative x-axis.

From Figure 6.5 and the associated monotonicity argument, we know that the bond angle for $\lambda = 0$ is greater than that for $\lambda = 1/2$. This is clearly visible when $r = 0.9$ where the $\lambda = 0$ case in Figure 6.6 clearly shows a single maximum opposite the fixed electron, whereas the $\lambda = 1/2$ case in Figure 6.7 has two maxima, one on the same side of the sphere as the fixed electron, and one on the opposite side.

The $\lambda = 1/2$ case is slightly different to all other cases in that the second maxima on the 'back' of the sphere is as strong as the one on the front. This is directly related to our discussion above concerning the degeneracy when $\lambda = 1/2$.

Whilst the two cases for high values of r look similar, both having two maxima and two minima and a bond angle approaching 90°, there are some subtle differences. One is that the minima on the 'front' in the $\lambda = 1/2$ case is centred on the fixed electron, whilst in the $\lambda = 0$ case it is clearly not. This may be explained by the fact that the density in the $\lambda = 0$ case is strongly polarized in the z-direction, meaning that there is an enhanced probability of finding electrons near the poles, forcing the minimum towards the equator.

Finally, Figure 6.8 shows a pair density at a value of λ and distance r chosen to demonstrate that the global maximum for the non-degenerate case is on the 'front'.

We also plot the Oxygen single-particle densities in order to show the discrepancy between naive intuition about where the pair density should be maximized and the actual analytic maxima. These are shown in Figures 6.9 and 6.10.

It is clear that it is very difficult to predict the position of the electrons in the maximizing pair density simply from looking at the single particle density. For example, Figure 6.9 shows that, for $\lambda = 0$, the single particle density is strongly polarised and intuitively one may expect the most likely pair density to have one electron at each pole. While this is true for small values of r, for larger values it is clearly not the case. (It is worth noting that for $\lambda = 0$ there is in fact an S^1 of possible positions for the first electron, obtained by rotating around the z axis so the chosen position is not unique).

Similarly, intuition and Figure 6.10 would suggest that the optimal pair density would have electrons at opposite poles of the sphere, this time having an S^1 rotated around the y axis. However this is again clearly not the case for even bond-length values r. It is true (as we would suspect from intuition) that the maximizing pair density occurs when the electrons are located on the S^1 of maximum single particle density.

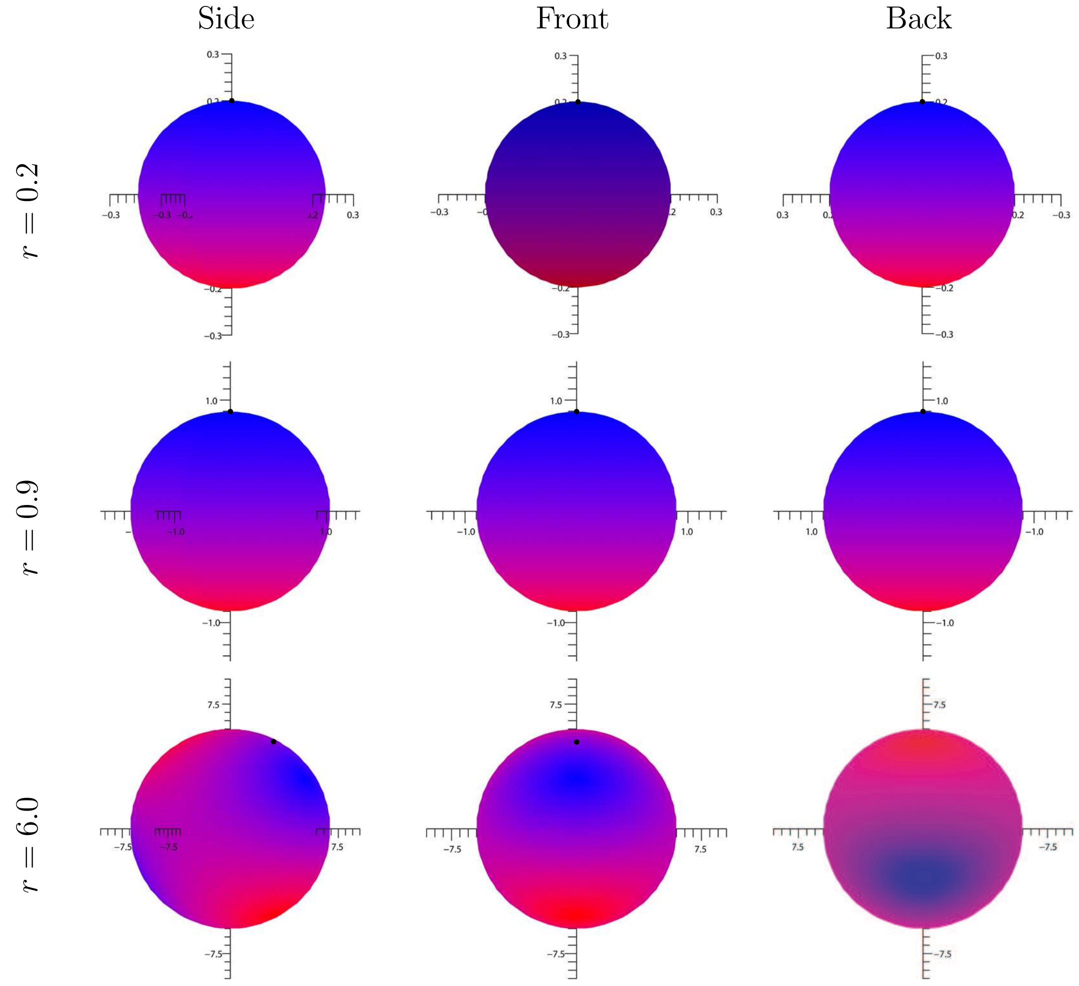


Figure 6.6: Oxygen pair densities for various distances with $\lambda = 0$. Front denotes the view along the positive e_1 axis, Back the view along the negative e_1 axis, and Side the view with the e_1 axis out of the plane and the e_3 axis vertical. The black dot denotes the position of one electron for the maximum pair density.

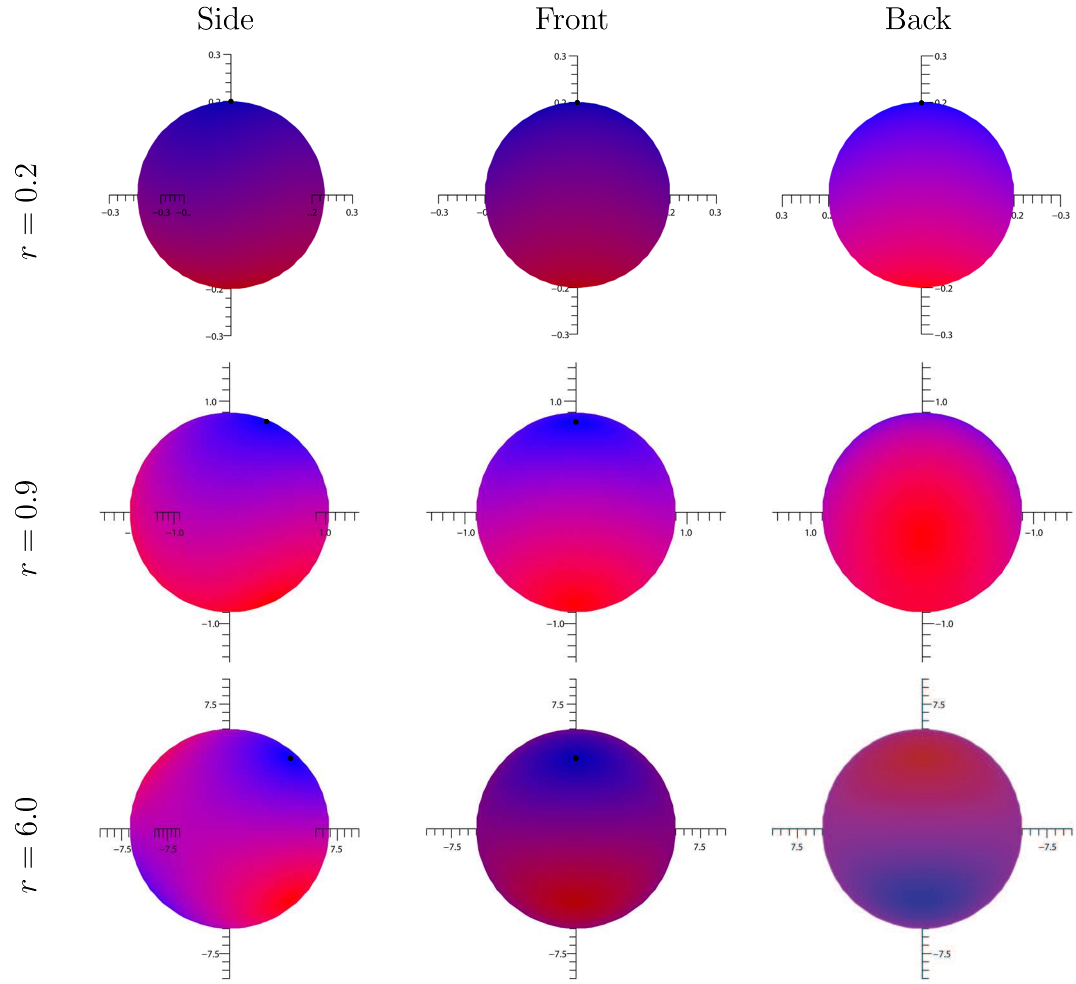


Figure 6.7: Oxygen pair densities for various distances with $\lambda = 1/2$, details as in Figure 6.6.

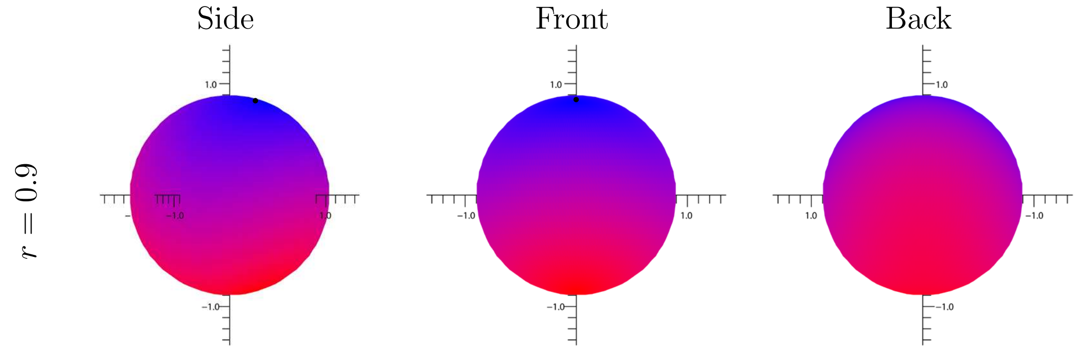


Figure 6.8: Oxygen pair densities with $\lambda = 0.4$, details as in Figure 6.6.

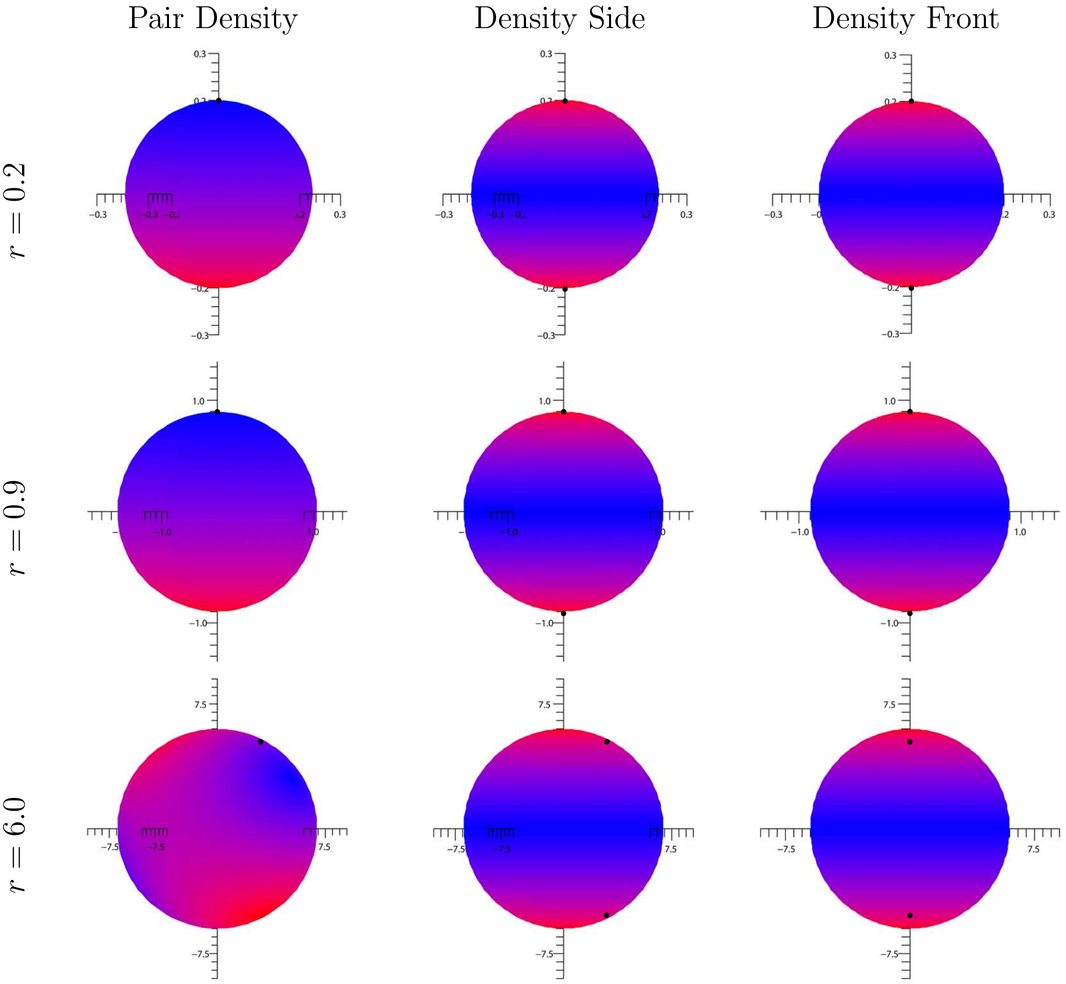


Figure 6.9: Oxygen pair density and single particle density with $\lambda = 0$, orientations as in Figure 6.6. The black dots denote the positions of the electrons in the maximum pair density.

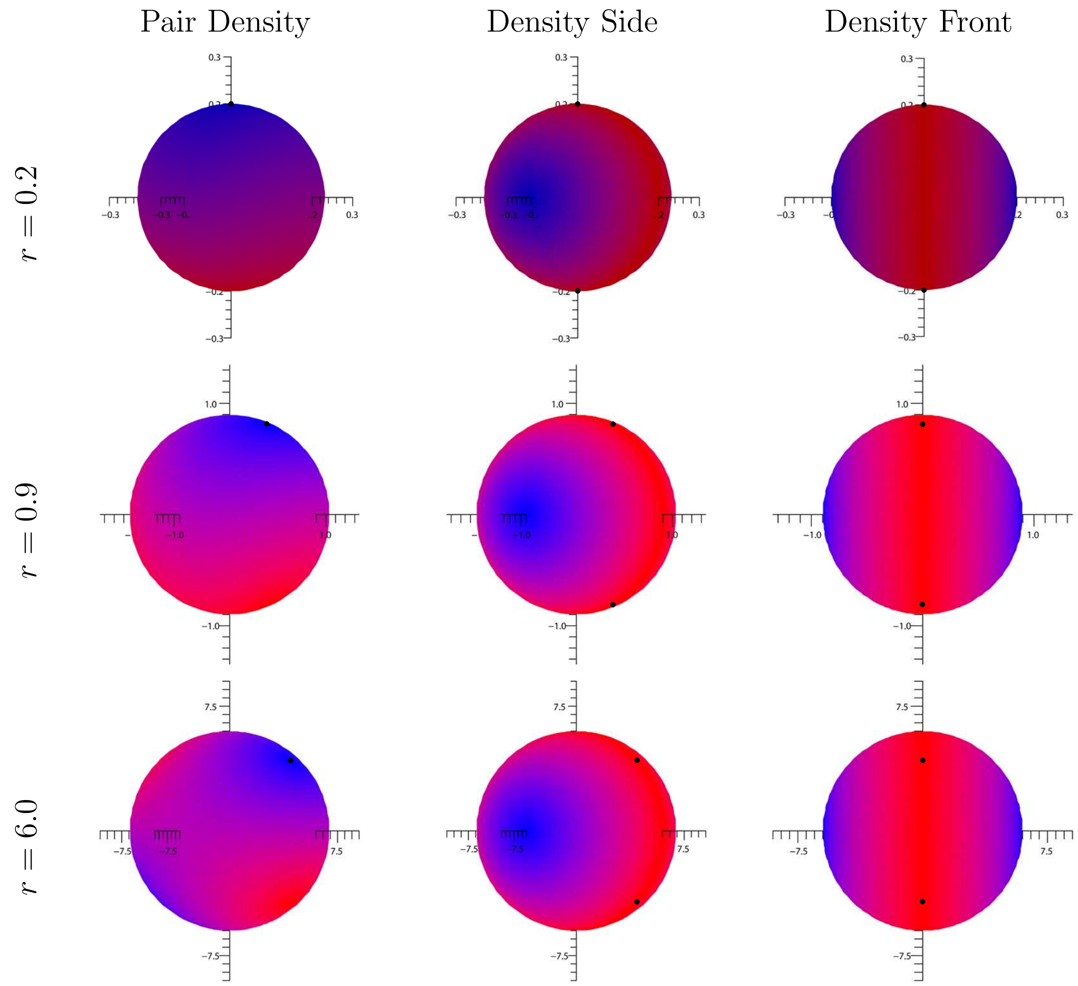


Figure 6.10: Oxygen pair density and single particle density with $\lambda = 1/2$, details as in Figure 6.9.

6.3.11 Advantages of Analytic Results

Our results show that a reasonably good approximation to the bond angle in water can be obtained with an easily-computable wavefunction formed from a minimal basis of one-electron orbitals, along with the simple assumption that the two bonds with the Hydrogen atoms will pass through the maximum of the pair density.

As can be seen from the equation for the bond angle (6.57), one of the most important factors affecting the bond angle at a given distance from the nucleus is the relative rates of decay of the $1s$, $2s$ and $2p$ orbitals, which are represented by $A_1(r)$ and $A_2(r)$. Hence it is very important to accurately model the effects of screening, which is done via the variational model in Section 4.2.

Furthermore, since Theorem 6.3.5 is independent of the exact forms of the radial components $R_i(r)$, it would be interesting to perform a similar calculation to that of Chapters 2-4 (enforcing the canonical form of the wavefunction), but in a near-complete basis. The resulting radial functions should then give a more accurate prediction of the pair density and thus of the bond angle.

Another point of note is the relative magnitudes of the pair density at various distances from the nucleus, along with the difference between the maxima and minima at a fixed distance. For example, at short distances the difference between the maxima and minima is of order 10 ($r = 0.1$) whilst at long distances it is of order 10^{-18} ($r = 6$). Similarly, the total pair density at $r = 0.1$ is of order 10^3 whilst at $r = 6$ it is of order 10^{-18}. This shows that numerical approximations could come into difficulty when trying to find the critical points, and also how useful an analytic expression is when trying to determine such physical properties.

Another advantage of an analytic expression for the pair density comes when plotting figures such as those in Figures 6.5-6.10 where no interpolative approximations are required. In other words, once the pair density has been calculated explicitly, using it to plot figures and retrieve other such data is relatively simple and needs little further computation, unlike with a numerical approximation where only values at a discrete number of points are known.

6.4 Neon

We now move onto the remaining atoms in the second row. We expect that the Fluorine ground state is similar to the Oxygen ground state in that instead of having the Nitrogen ground state ‘plus’ a p-orbital, we have the Neon ground state ‘minus’ a p-orbital. The idea of subtracting a p-orbital from a wavefunction is less well defined than the idea of adding one but this will be discussed later. Hence we will need the density and pair density for Neon. We begin by noting that $\Psi^{Ne} = |1s\overline{1s}2s\overline{2s}p_1\overline{p_1}p_2\overline{p_2}p_3\overline{p_3}\rangle$ and hence

$$\begin{aligned}\rho^{Ne}(x) &= 2|1s(x)|^2 + 2|2s(x)|^2 + 2|p_1(x)|^2 + 2|p_2(x)|^2 + 2|p_3(x)|^2 \\ &= 2R_1^2(r) + 2R_2^2(r) + 2r^2R_3^2(r), \qquad (6.58)\end{aligned}$$

where the $R_i(r)$ are as before.

We expect the pair density to have a similar form to that of Nitrogen, in that it should only depend on the angle between the electrons. Using the same method with Slater's rules, we have

$$\begin{aligned}\rho_2^{Ne}(x,y) = \tfrac{1}{2}\Big[&|1s(x)|^2|1s(y)|^2 + 4|1s(x)|^2|2s(y)|^2 + |2s(x)|^2|2s(y)|^2 \\ &+ 4\big(|1s(x)|^2 + |2s(x)|^2\big)\sum_{i=1}^{3}|p_i(y)|^2 + \sum_{i=1}^{3}|p_i(x)|^2|p_i(y)|^2 \\ &+ 2\sum_{1\le i<j\le 3}\big(|p_i(x)|^2|p_j(y)|^2 + |p_j(x)|^2|p_i(y)|^2\big) + x \leftrightarrow y\Big] \\ - 2\Big[&1s(x)2s(x)1s(y)2s(y) + \sum_{1\le i<j\le 3} p_i(x)p_i(y)p_j(x)p_j(y) \\ &+ \big(1s(x)1s(y) + 2s(x)2s(y)\big)\sum_{i=1}^{3} p_i(x)p_i(y)\Big].\end{aligned}$$

Writing the p-orbitals in Cartesian coordinates we have

$$\begin{aligned}\rho_2^{Ne}(x,y) &= F(r) - 2R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)(x_1y_1 + x_2y_2 + x_3y_3) \\ &\quad + R_3^4(r)\big[x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 + 2x_1^2y_2^2 + 2x_2^2y_1^2 + 2x_1^2y_3^2 + 2x_3^2y_1^2 \\ &\qquad + 2x_2^2y_3^2 + 2x_3^2y_2^2 - 2x_1y_1x_2y_2 - 2x_1y_1x_3y_3 - 2x_2y_2x_3y_3\big] \\ &= F(r) - 2R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)\sum_{i=1}^{3} x_iy_i \\ &\quad + R_3^4(r)\Big[2\sum_{i=1}^{3}x_i^2\sum_{j=1}^{3}y_j^2 - \big|\sum_{i=1}^{3}x_iy_i\big|^2\Big] \\ &= \tilde{F}(r) - 2R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)r^2\cos\Theta - R_3^4(r)r^4\cos^2\Theta, \qquad (6.59)\end{aligned}$$

where

$$\tilde{F} = R_1^4(r) + 2R_1^2(r)R_2^2(r) + R_2^2(r) + 4\big(R_1^2(r) + R_2^2(r)\big)r^2R_3^2(r) + 2r^4R_3^4(r).$$

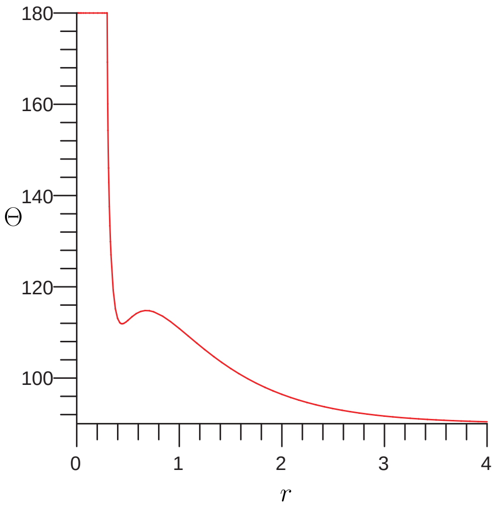


Figure 6.11: Bond angle range for Neon

It is now clear that the Neon density and pair density are invariant under simultaneous rotation of both coordinates. We denote the pair density in (6.59) by

$$\rho_2^{Ne} =: A_0(r) - 2A_1(r)\cos\Theta - A_2(r)\cos^2\Theta$$

where $A_0(r) := \tilde{F}(r)$, $A_1(r) := r^2 R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)$, $A_2(r) := r^4 R_3^4(r)$, and in particular, $A_i(r) \geq 0$. We then have, following the analogous calculation for Nitrogen, that the maximum is at $\cos\Theta = \max\big\{-1, -\frac{A_1(r)}{A_2(r)}\big\}$:

Theorem 6.4.1 (Maximum of Neon Pair Density)**.** *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define*

$$A_0(r) := R_1^4(r) + 2R_1^2(r)R_2^2(r) + R_2^2(r) + 4\big(R_1^2(r) + R_2^2(r)\big)R_3^2(r) + 2R_3^4(r),$$

$A_1(r) := \big(R_1^2(r) + R_2^2(r)\big)R_3^2(r)$, $A_2(r) := R_3^4(r)$. *Then the maximum of the Neon ground state pair density*

$$\rho_2^{Ne} = A_0(r) - 2A_1(r)\cos\Theta - A_2(r)\cos^2\Theta$$

is at $\cos\Theta = \max\left\{-1, -\frac{A_1(r)}{A_2(r)}\right\}$.

This maximizing angle for Neon is shown in Figure 6.11, whilst the pair density plots are very similar to those of Nitrogen and are therefore not included here. We see that, as $r \to \infty$, the fact that $A_2(r)$ decays exponentially slower than $A_1(r)$ gives a limit of $\cos\Theta = 0$, or $\Theta = 90°$.

Unlike in the previous cases there is no experimental data for the NeH_2

molecule and hence we cannot test the accuracy of our predictions in this case. This lack of experimental data leads us to expect that, for a multi-centre calculation, the NeH_2 molecule should be energetically unstable.

6.5 Fluorine

6.5.1 Canonical Ground State Pair Density

We begin by noting that the Fluorine case is slightly more complicated than that of Oxygen as there is no simple way to write a Slater determinant as another which is ‘missing’ some orbitals. By Lemma 6.2.1 we consider the $S_3\Psi = \frac{1}{2}\Psi$ case:

$$\begin{aligned}\Psi^F &= \alpha_1|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_2p_3}\rangle + \alpha_2|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_1p_3}\rangle \\ &\quad + \alpha_3|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_1p_2}\rangle \\ &=: \alpha_1\Psi_1 + \alpha_2\Psi_2 + \alpha_3\Psi_3\end{aligned}$$

where $\alpha_i \in \mathbb{C}, \sum|\alpha_i|^2 = 1$, and the labelling has been chosen to emphasize the idea of ‘missing’ orbitals from a spherically symmetric wavefunction. The corresponding pair density is given by

$$\begin{aligned}\rho_2^F &= |\alpha_1|^2\rho_2^{\Psi_1} + |\alpha_2|^2\rho_2^{\Psi_2} + |\alpha_3|^2\rho_2^{\Psi_3} + 2\mathrm{Re}(\alpha_1\alpha_2^*)\langle\Psi_1|V_{x,y}|\Psi_2\rangle \\ &\quad + 2\mathrm{Re}(\alpha_1\alpha_3^*)\langle\Psi_1|V_{x,y}|\Psi_3\rangle + 2\mathrm{Re}(\alpha_2\alpha_3^*)\langle\Psi_2|V_{x,y}|\Psi_3\rangle.\end{aligned}$$

We begin by finding the pair density for one of the Ψ_i:

$$\begin{aligned}\rho_2^{\Psi_3}(x,y) &= \rho_2^{Ne} - \tfrac{1}{2}\big[p_3^2(x)\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_1p_2}\rangle}(y) + x \leftrightarrow y\big] \\ &\qquad + p_3(x)p_3(y)\big(1s(x)1s(y) + 2s(x)2s(y) + p_1(x)p_1(y) + p_2(x)p_2(y)\big) \\ &= \rho_2^{Ne} - \tfrac{1}{2}\big[p_3^2(x)\rho^{Ne}(y) + x \leftrightarrow y\big] \\ &\qquad + p_3(x)p_3(y)\Big(1s(x)1s(y) + 2s(x)2s(y) + \sum_{i=1}^{3}p_i(x)p_i(y)\Big).\end{aligned}$$

Hence, by symmetry, we see that

$$\begin{aligned}\rho_2^{\Psi_i}(x,y) &= \rho_2^{Ne} - \tfrac{1}{2}\big[p_i^2(x)\rho^{Ne}(y) + x \leftrightarrow y\big] \\ &\qquad + p_i(x)p_i(y)\Big(1s(x)1s(y) + 2s(x)2s(y) + \sum_{j=1}^{3}p_j(x)p_j(y)\Big).\end{aligned}$$

The next stage is to derive a form for the cross terms in the pair density, for which we need to use Slater's rules for a pair of Slater determinants which differ by one orbital. For example, consider

$$\begin{aligned}
\langle\Psi_2|V_{x,y}|\Psi_3\rangle &= \tfrac{1}{2}p_2(x)p_3(x)\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_1}\rangle}(y) \\
&\quad - \tfrac{1}{2}p_2(x)p_3(y)\big(1s(x)1s(y)+2s(x)2s(y)+p_1(x)p_1(y)\big) + x \leftrightarrow y \\
&= \tfrac{1}{2}p_2(x)p_3(x)\rho^{Ne}(y) - \tfrac{1}{2}p_2(x)p_3(x)\big(p_2^2(y)+p_3^2(y)\big) \\
&\quad - \tfrac{1}{2}p_2(x)p_3(y)\big(1s(x)1s(y)+2s(x)2s(y)+p_1(x)p_1(y)\big) + x \leftrightarrow y \\
&= \tfrac{1}{2}p_2(x)p_3(x)\Big[\rho^{Ne}(y) - 1s(x)1s(y) - 2s(x)2s(y) - \sum_{i=1}^{3} p_i(x)p_i(y)\Big] \\
&\quad + x \leftrightarrow y,
\end{aligned}$$

and the other cases are analogous.

We therefore see that the pair density for a general Fluorine ground state wavefunction is given by

$$\begin{aligned}
\rho_2^F(x,y) &= \rho_2^{Ne} - \tfrac{1}{2}\Big(\sum_{i=1}^{3}|\alpha_i|^2p_i^2(x)\Big)\rho^{Ne}(y) \\
&\quad + \tfrac{1}{2}\Big(\sum_{i=1}^{3}|\alpha_1|^2p_i(x)p_i(y)\Big)\Big(1s(x)1s(y)+2s(x)2s(y)+\sum_{i=1}^{3}p_i(x)p_i(y)\Big) \\
&\quad + \big(\mathrm{Re}(\alpha_1\alpha_2^*)p_1(x)p_2(x) - \mathrm{Re}(\alpha_1\alpha_3^*)p_1(x)p_3(x) \\
&\qquad\qquad + \mathrm{Re}(\alpha_2\alpha_3^*)p_2(x)p_3(x)\big)\rho^{Ne}(y) \\
&\quad - \big(\mathrm{Re}(\alpha_1\alpha_2^*)p_1(x)p_2(y) - \mathrm{Re}(\alpha_1\alpha_3^*)p_1(x)p_3(y) + \mathrm{Re}(\alpha_2\alpha_3^*)p_2(x)p_3(y)\big) \\
&\qquad \times \big(1s(x)1s(y)+2s(x)2s(y)+\sum_{i=1}^{3}p_i(x)p_i(y)\big) + x \leftrightarrow y \\
&= \rho_2^{Ne} - \tfrac{1}{2}\rho^{Ne}(y)|\alpha_1p_1(x)-\alpha_2p_2(x)+\alpha_3p_3(x)|^2 \\
&\quad + \tfrac{1}{2}\big(\alpha_1p_1(x)-\alpha_2p_2(x)+\alpha_3p_3(x)\big)\big(\alpha_1^*p_1(y)-\alpha_2^*p_2(y)+\alpha_3^*p_3(y)\big) \\
&\qquad \times \Big(1s(x)1s(y)+2s(x)2s(y)+\sum_{i=1}^{3}p_i(x)p_i(y)\Big) + x \leftrightarrow y.
\end{aligned}$$

The minus sign with α_2 comes from the fact that the p_2 orbital is at an odd position in the Neon Slater determinant whilst the p_1 and p_3 orbitals are at even positions, leading to a change of sign in order to apply the annihilation operator to the p_2 as compared to the other two cases. We could equally well have defined Ψ^F with a minus sign on the second term.

Similarly to the Oxygen case, we set $ke^{-Z|x|/2}(a+ib)\cdot x := \alpha_1p_1 - \alpha_2p_2 + \alpha_3p_3$

and we now have

$$\begin{aligned}\rho_2^F(x,y) = \rho_2^{Ne}(x,y) &- \tfrac{1}{2}\Big(\rho^{Ne}(y)k^2e^{-Z|x|}\big(x\cdot(a\otimes a+b\otimes b)x\big) + x\leftrightarrow y\Big)\\ &+ \Big(1s(x)1s(y)+2s(x)2s(y)+\sum_{i=1}^{3}p_i(x)p_i(y)\Big)\\ &\qquad\times k^2e^{-Z(|x|+|y|)/2}\big(x\cdot(a\otimes a+b\otimes b)y\big).\end{aligned}$$

We now note that the terms $\rho_2^{Ne}(x,y)$, $\rho^{Ne}(\cdot)$, $1s(x)1s(y)+2s(x)2s(y)$ and $\sum_{i=1}^{3}p_i(x)p_i(y)$ are invariant under simultaneous rotation of x and y, which leads to the one parameter family of pair densities

$$\begin{aligned}\rho_2^{\Psi(R)}(x,y) = \rho_2^{Ne}(x,y) &- \tfrac{1}{2}\Big(\rho^{Ne}(y)k^2e^{-Z|x|}\big((1-\lambda)x_3^2+\lambda x_1^2\big) + x\leftrightarrow y\Big)\\ &+ \big(1s(x)1s(y)+2s(x)2s(y)+\sum_{i=1}^{3}p_i(x)p_i(y)\big)\\ &\qquad\times k^2e^{-Z(|x|+|y|)/2}\big((1-\lambda)x_3y_3+\lambda x_1y_1\big),\end{aligned}$$

with a corresponding one-parameter wavefunction given by

$$\Psi^F = \sqrt{1-\lambda}|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_1p_2}\rangle + i\sqrt{\lambda}|1s\overline{1s}2s\overline{2s}p_1p_2p_3\overline{p_2p_3}\rangle.$$

We now insert the expressions for ρ^{Ne} and ρ_2^{Ne} from (6.58) and (6.59) and change to spherical polar coordinates to obtain

$$\begin{aligned}\rho_2^F = \tilde{F}(r) &- 2r^2R_3^2(r)\big(R_1^2(r)+R_2^2(r)\big)\\ &\qquad\times\big(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2(\sin\phi_1\sin\phi_2+\cos\phi_1\cos\phi_2)\big)\\ &- r^4R_3^4(r)\big(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2(\sin\phi_1\sin\phi_2+\cos\phi_1\cos\phi_2)\big)^2\\ &- \tfrac{1}{2}r^2R_3^2\big(2R_1^2(r)+2R_2^2(r)+2r^2R_3^2(r)\big)\\ &\qquad\times\big((1-\lambda)(\cos^2\theta_1+\cos^2\theta_2)+\lambda(\sin^2\theta_1\sin^2\phi_1+\sin^2\theta_2\sin^2\phi_2)\big)\\ &+ r^2R_3^2(r)\big((1-\lambda)\cos\theta_1\cos\theta_2+\lambda\sin\theta_1\sin\phi_1\sin\theta_2\sin\phi_2\big)\\ &\qquad\times\big(R_1^2(r)+R_2^2(r)+r^2R_3^2(r)(\cos\theta_1\cos\theta_2\\ &\qquad\qquad+\sin\theta_1\sin\theta_2(\sin\phi_1\sin\phi_{,2}+\cos\phi_1\cos\phi_2)\big).\end{aligned}$$

Denoting $A_0(r) := \tilde{F}(r)$, $A_1(r) := r^2R_3^2(r)\big(R_1^2(r)+R_2^2(r)\big)$ and $A_2(r) := r^4R_3^4(r)$

and using Corollary 6.3.4, we have

$$\begin{aligned}
\rho_2^F &= A_0(r) \\
&\quad + A_1(r)\big[-2\big(\epsilon_{c\theta}\cos^2\theta + \sin^2\theta(\epsilon_{s\phi}\sin^2\phi + \epsilon_{c\phi}\cos^2\phi)\big) + (1-\lambda)\epsilon_{c\theta}\cos^2\theta \\
&\qquad\qquad - 2(1-\lambda)\cos^2\theta - 2\lambda\sin^2\theta\sin^2\phi + \lambda\epsilon_{s\phi}\sin^2\theta\sin^2\phi\big] \\
&\quad + A_2(r)\big[-\big(\epsilon_{c\theta}\cos^2\theta + \sin^2\theta(\epsilon_{s\phi}\sin^2\phi + \epsilon_{c\phi}\cos^2\phi)\big)^2 \\
&\qquad - 2(1-\lambda)\cos^2\theta - 2\lambda\sin^2\theta\sin^2\phi \\
&\qquad + \big((1-\lambda)\epsilon_{c\theta}\cos^2\theta + \lambda\epsilon_{s\phi}\sin^2\theta\sin^2\phi\big) \\
&\qquad\qquad \times \big(\epsilon_{c\theta}\cos^2\theta + \sin^2\theta(\epsilon_{s\phi}\sin^2\phi + \epsilon_{c\phi}\cos^2\phi)\big)\big].
\end{aligned}$$

From this expression and the identity $\cos^2\varphi = 1 - \sin^2\varphi$ it is simple to find the corresponding expressions for the α_i (once again ignoring the contribution of $A_0(r)$ to α_6):

$$\begin{aligned}
\alpha_1 &:= A_2(r)(\epsilon_{s\phi} - \epsilon_{c\phi})\big[-(\epsilon_{s\phi} - \epsilon_{c\phi}) + \lambda\epsilon_{s\phi}\big], \\
\alpha_2 &:= A_2(r)\big[(\epsilon_{s\phi} - \epsilon_{c\phi})\big(2(\epsilon_{c\theta} - \epsilon_{c\phi}) - (1-\lambda)\epsilon_{c\theta}\big) - \lambda\epsilon_{s\phi}(\epsilon_{c\theta} - \epsilon_{c\phi})\big], \\
\alpha_3 &:= A_2(r)[2(\epsilon_{c\theta}\epsilon_{c\phi} - 1) - (1-\lambda)\epsilon_{c\theta}(\epsilon_{c\phi} - \epsilon_{c\theta})], \\
\alpha_4 &:= A_1(r)[-2(\epsilon_{s\phi} - \epsilon_{c\phi}) - 2\lambda + \epsilon_{s\phi}\lambda] \\
&\qquad + A_2(r)[-2\epsilon_{c\theta}(\epsilon_{s\phi} - \epsilon_{c\phi}) - 2\lambda + (1-\lambda)\epsilon_{c\theta}(\epsilon_{s\phi} - \epsilon_{c\phi}) + \lambda\epsilon_{c\theta}\epsilon_{s\phi}] \\
\alpha_5 &:= A_1(r)[\epsilon_{c\theta} - 2\epsilon_{c\phi} + 2 - 2\lambda + \lambda\epsilon_{c\theta}] \\
&\qquad + A_2(r)[3 - 2\epsilon_{c\theta}\epsilon_{c\phi} - \lambda + (1-\lambda)\epsilon_{c\theta}(\epsilon_{c\theta} - \epsilon_{c\phi})], \\
\alpha_6 &:= A_1(r)[-\epsilon_{c\theta} + 2\lambda - \lambda\epsilon_{c\theta} - 2] + A_2(r)[-2 + \lambda].
\end{aligned}$$

We may now use these values of the α_i to compute the global maximum pair density for Fluorine. The critical points in Table 6.2 have the following forms:

- $\rho_2^F = \alpha_6 = A_1(r)[-\epsilon_{c\theta} + 2\lambda - \lambda\epsilon_{c\theta} - 2] + A_2(r)[-2 + \lambda].$

This is clearly maximized when $\epsilon_{c\theta} = -1$ (and is independent of $\epsilon_{c\phi}$ and $\epsilon_{s\phi}$) and gives a value of

$$\rho_2^{F,1} := A_1(r)[-1 + 3\lambda] + A_2(r)[-2 + \lambda].$$

- $\rho_2^F = \alpha_3 + \alpha_5 + \alpha_6 = -2A_1(r)\epsilon_{c\phi} - A_2(r).$

This is clearly maximized when $\epsilon_{c\phi} = -1$ (and is independent of $\epsilon_{c\theta}$ and $\epsilon_{s\phi}$) and has a maximum value of

$$\rho_2^{F,2} := 2A_1(r) - A_2(r).$$

- $$\begin{aligned}\rho_2^F &= \alpha_1+\alpha_2+\alpha_3+\alpha_4+\alpha_5+\alpha_6\\ &= A_1(r)[-2\lambda+\lambda\epsilon_{s\phi}-2\epsilon_{s\phi}]+A_2(r)[-1-\lambda].\end{aligned}$$

This is clearly maximized when $\epsilon_{s\phi} = -1$ (and is independent of $\epsilon_{c\theta}$ and $\epsilon_{c\phi}$) which gives a maximum value of

$$\rho_2^{F,3} := A_1(r)[2-3\lambda] + A_2(r)[-1-\lambda].$$

- $$\rho_2^F = -\frac{\alpha_5^2}{4\alpha_3} + \alpha_6.$$

After inserting the 8 possible combinations for the ϵ's we find that there are two possible pair densities:

$$\begin{aligned}\rho_2^{F,4,1} &:= \tfrac{1}{4A_2(r)(2+2\lambda)}\left[A_1(r)(-1-3\lambda)+A_2(r)(3+\lambda)\right]^2\\ &\quad + A_1(r)[-1+3\lambda]+A_2(r)[-2+\lambda], \qquad (6.60)\\ \rho_2^{F,4,2} &:= \tfrac{1}{4A_2(r)(2+2\lambda)}\left[A_1(r)(5-\lambda)+A_2(r)(3+\lambda)\right]^2\\ &\quad + A_1(r)[-3+\lambda]+A_2(r)[-2+\lambda], \qquad (6.61)\end{aligned}$$

which are valid for $f_{4,i} \in [0,1]$, where

$$f_{4,1} := \frac{3+\lambda}{4(1+\lambda)} - \frac{A_1(r)}{A_2(r)}\frac{1+3\lambda}{2(2+2\lambda)}, \qquad (6.62)$$

$$f_{4,2} := \frac{3+\lambda}{4(1+\lambda)} + \frac{A_1(r)}{A_2(r)}\frac{5-\lambda}{2(2+2\lambda)}. \qquad (6.63)$$

The first case corresponds to $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (-1, 1, \pm1)$ and the second to $(1, -1, \pm1)$, the other four possibilities giving $\alpha_3 = 0$.

- $$\rho_2^F = -\frac{(\alpha_4+\alpha_5)^2}{4(\alpha_1+\alpha_2+\alpha_3)} + \alpha_6.$$

Again inserting the 8 possible combinations for the ϵ's we find that there are two possible pair densities:

$$\begin{aligned}
\rho_2^{F,5,1} &:= \tfrac{1}{8A_2(r)}\left[A_1(r)(-1-4\lambda)+A_2(r)(3-2\lambda)\right]^2 \\
&\quad + A_1(r)[-1+3\lambda]+A_2(r)[-2+\lambda], \qquad (6.64)\\
\rho_2^{F,5,2} &:= \tfrac{1}{8A_2(r)}\left[A_1(r)(5-4\lambda)+A_2(r)(3-2\lambda)\right]^2 \\
&\quad + A_1(r)[-3+\lambda]+A_2(r)[-2+\lambda], \qquad (6.65)
\end{aligned}$$

which are valid for $f_{5,i} \in [0,1]$, where

$$\begin{aligned}
f_{5,1} &:= \frac{3-2\lambda}{4} - \frac{A_1(r)}{A_2(r)}\frac{1+4\lambda}{4}, \qquad (6.66)\\
f_{5,2} &:= \frac{3-2\lambda}{4} + \frac{A_1(r)}{A_2(r)}\frac{5-4\lambda}{4}. \qquad (6.67)
\end{aligned}$$

The first case corresponds to $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (-1, \pm 1, 1)$ and the second to $(1, \pm 1, -1)$, the other four possibilities giving $\alpha_1 + \alpha_2 + \alpha_3 = 0$.

- $\rho_2^F = -\dfrac{(\alpha_2+\alpha_4)^2}{2\alpha_1} + \alpha_3 + \alpha_5 + \alpha_6.$

Inserting the 8 possible combinations for the ϵ's we find once again that there are two possible pair densities:

$$\begin{aligned}
\rho_2^{F,6,1} &:= \tfrac{1}{8A_2(r)(2-\lambda)}\left[A_1(r)(-4-\lambda)+A_2(r)(4-3\lambda)\right]^2 \\
&\quad + 2A_1(r) - A_2(r), \qquad (6.68)\\
\rho_2^{F,6,2} &:= \tfrac{1}{8A_2(r)(2-\lambda)}\left[A_1(r)(4-3\lambda)+A_2(r)(4-3\lambda)\right]^2 \\
&\quad - 2A_1(r) - A_2(r), \qquad (6.69)
\end{aligned}$$

which are valid for $f_{6,i} \in [0,1]$, where

$$\begin{aligned}
f_{6,1} &:= \frac{4-3\lambda}{4(2-\lambda)} - \frac{A_1(r)}{A_2(r)}\frac{4+\lambda}{4(2-\lambda)}, \qquad (6.70)\\
f_{6,2} &:= \frac{4-3\lambda}{4(2-\lambda)} + \frac{A_1(r)}{A_2(r)}\frac{4-3\lambda}{4(2-\lambda)}. \qquad (6.71)
\end{aligned}$$

The first case corresponds to $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (\pm 1, -1, 1)$ and the second to $(\pm 1, 1, -1)$, the other four possibilities giving $\alpha_1 = 0$.

The final case is that in which both $\sin^2\theta$ and $\sin^2\phi$ are non-trivial and are given by

- $\sin^2\theta = \dfrac{\alpha_2\alpha_4 - 2\alpha_1\alpha_5}{4\alpha_1\alpha_3 - \alpha_2^2}, \quad \sin^2\phi = \dfrac{2\alpha_3\alpha_4 - \alpha_2\alpha_5}{2\alpha_1\alpha_5 - \alpha_2\alpha_4}.$

We note that, out of the 8 possible combinations of the ϵ's, there are 6 cases where $\sin^2\theta$ is well defined, giving the following possible values:

- $\frac{3}{2} + \frac{A_1(r)}{2A_2(r)} > 1$,
- $\frac{3}{2} + \frac{3A_1(r)}{2A_2(r)} > 1$,
- $\frac{A_1(r)(16-34\lambda+4\lambda^2)+A_2(r)(12-22\lambda-4\lambda^2)}{A_2(-4+16\lambda-16\lambda^2)} < 0$,
- $\frac{A_1(r)(-16+30\lambda-14\lambda^2)+A_2(r)(-12+18\lambda-6\lambda^2)}{A_2(2-2\lambda)^2} < 0$,
- $\frac{A_1(r)(-16+26\lambda-10\lambda^2)+A_2(r)(-12+18\lambda-6\lambda^2)}{A_2(2-2\lambda)^2} < 0$,
- $\frac{A_1(r)(16-38\lambda+12\lambda^2)+A_2(r)(12-22\lambda-4\lambda^2)}{A_2(-4+16\lambda-16\lambda^2)} < 0$.

Hence none of these cases are valid choices for $\sin^2\theta$ and the final critical point never exists.

Theorem 6.5.1 (Maximum of Fluorine Pair Density)**.** *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define* $A_0(r) := R_1^4(r) + 2R_1^2(r)R_2^2(r) + R_2^4(r) + 4\big(R_1^2(r) + R_2^2(r)\big)R_3^2(r) + 2R_3^4(r)$, $A_1(r) := \big(R_1^2(r) + R_2^2(r)\big)R_3^2(r)$, $A_2(r) := R_3^4(r)$. *Further denote*

$$\begin{aligned} \rho_2^{F,2}(r,\lambda) &:= 2A_1(r) - A_2(r), \\ \rho_2^{F,6,1}(r,\lambda) &:= \tfrac{1}{4A_2(r)(4-2\lambda)}\left[A_1(r)(-4-\lambda) + A_2(r)(4-3\lambda)\right]^2 \\ &\quad + 2A_1(r) - A_2(r), \end{aligned}$$

and $f(r,\lambda) = \frac{4-3\lambda}{2(4-2\lambda)} - \frac{A_1(r)}{A_2(r)}\frac{4+\lambda}{2(4-2\lambda)}$. *Then the global maximum of the Fluorine pair density is given by*

$$\rho_2^{F,max}(r,\lambda) = \begin{cases} \rho_2^{F,6,1} & \text{if } f(r,\lambda) \in (0,1] \\ \rho_2^{F,2} & \text{else.} \end{cases}$$

As with the Oxygen case, it suffices to consider the part of the pair density with angular dependence, the critical points of which are given by $\rho_2^{F,1}$ to $\rho_2^{F,6,2}$.

We begin with the case from Theorem 6.5.1 in which $\rho_2^{F,max} = \rho_2^{F,2}$ and $f < 0$. Comparing to the cases which have no existence condition we find that

$$\begin{aligned} \rho_2^{F,2} - \rho_2^{F,1} &= 3(1-\lambda)A_1(r) + (1-\lambda)A_2(r) > 0, \\ \rho_2^{F,2} - \rho_2^{F,3} &= 3\lambda A_1(r) + \lambda A_2(r) \geq 0, \end{aligned}$$

where equality in the second case occurs only when $\lambda = 0$.

This degeneracy is similar to that in the Oxygen case. When $\lambda = 0$, the pair density is invariant under rotation around the e_3 axis and hence, since x and y lie opposite each other in the e_1-e_2 plane for both $\rho_2^{F,2}$ and $\rho_2^{F,3}$, the values of these critical points are the same.

In order to compare $\rho_2^{F,2}$ to the critical points with existence conditions we use Lemma 6.3.6 with $\rho_2^{(1)} = \rho_2^{F,2}$ and $f_3 = f_{6,1}$. We then set $\rho_2^{(2)}$ and f_2 to be the critical point to be tested. We therefore have $e_1 = 2$, $e_1 = -1$, $\alpha_3 = -4(2 - \lambda)$, $\beta_3 = -4 - \lambda$ and $\gamma_3 = 4 - 3\lambda$, giving that $-\alpha_3 A_2(r) > 0$ whenever $A_2(r) > 0$ and $\gamma_3 > 0$. It therefore remains to check the two conditions $2e_2 - (\gamma_2 + 2g_2) \leq 0$ and $h \geq 0$ in each case.

From (6.60)-(6.63) we see that $\alpha_{4,i} = -4(1 + \lambda)$, $\beta_{4,1} = -1 - 3\lambda$, $\beta_{4,2} = 5 - \lambda$, $\gamma_{4,i} = 3 + \lambda$, $g_{(4,1),1} = -1 + 3\lambda$, $g_{(4,2),1}) = -3 + \lambda$, and $g_{(4,i),2} = -2 + \lambda$. Hence $2e_2 - (\gamma_{4,i} + 2g_{(4,i),2} = -1 - 3\lambda < 0$ and

$$h_{4,1} = 20 - 6(5\lambda + \lambda^2) > 0, \quad h_{4,2} = 14 - 10(2\lambda + \lambda^2) > 0$$

and thus, when $f_{6,1} < 0$, $\rho_2^{F,2} > \rho_2^{F,4,i}$ for $i = 1, 2$.

From (6.64)-(6.67) we see that $\alpha_{4,i} = -4$, $\beta_{5,1} = -1 - 4\lambda$, $\beta_{5,2} = 5 - 4\lambda$, $\gamma_{5,i} = 3 - 2\lambda$, $g_{(5,1),1} = -1 + 3\lambda$, $g_{(5,2),1} = -3 + \lambda$, and $g_{(5,i),2} = -2 + \lambda$. Hence $2e_2 - (\gamma_{5,i} + 2g_{(5,i),2}) = -1 < 0$ and

$$h_{4,1} = 20 - 4(6\lambda - \lambda^2) > 0, \quad h_{4,2} = 14 - 4(2\lambda + \lambda^2) > 0,$$

and thus, when $f_{6,1} < 0$, $\rho_2^{F,2} > \rho_2^{F,5,i}$ for $i = 1, 2$.

From (6.69) and (6.71) we have $\alpha_{6,2} = -4(2 - \lambda)$, $\beta_{6,2} = 4 - 3\lambda$, $\gamma_{6,2} = 4 - 3\lambda$, $g_{(6,2),1} = -2$ and $g_{(6,2),2} = -1$. These give $2e_2 - (\gamma_{6,2} + 2g_{(6,2),2}) = -4 + 3\lambda < 0$ and $h_{6,2} = 8 - 6(2\lambda - \lambda^2) > 0$, and so, when $f_{6,1} < 0$, $\rho_2^{F,2} > \rho_2^{F,6,1}$.

This completes the analysis for the case when $\rho_2^{F,max} = \rho_2^{F,2}$ and we move on to the case when $\rho_2^{F,max} = \rho_2^{F,6,1}$. It is clear that, when it is valid, $\rho_2^{F,6,1} > \rho_2^{F,2}$ and so, by the above analysis, $\rho_2^{F,6,1}$ is larger than both $\rho_2^{F,1}$ and $\rho_2^{F,3}$.

From (6.60) and (6.68) we have

$$\begin{aligned} 8(1+\lambda)(2-\lambda)A_2(r)(\rho_2^{F,6,1}-\rho_2^{F,4,1}) &= (14+13\lambda+10\lambda^3-3\lambda^2)A_1^2(r) \\ + (28-6\lambda-34\lambda^2+24\lambda^3)&A_1(r)A_2(r) + (14-19\lambda-27\lambda^2+18\lambda^3)A_2^2(r) \\ &=: a_{11}A_1^2(r)+a_{12}A_1(r)A_2(r)+a_{22}A_2^2(r), \end{aligned}$$

and differentiating the a_{ij} with respect to λ we have

$$\begin{aligned} \frac{da_{11}}{d\lambda} &= 13-6\lambda+30\lambda^2>0, \\ \frac{da_{12}}{d\lambda} &= -6-68\lambda+72\lambda^2<0, \\ \frac{da_{22}}{d\lambda} &= -19-54\lambda+54\lambda^2<0. \end{aligned}$$

Thus the minimum values on $\lambda \in [0,1/2]$ are given by $\tilde{a}_{11}=14$, $\tilde{a}_{12}=39/2$ and $\tilde{a}_{22}=0$. Thus we see that $\rho_2^{F,6,1}>\rho_2^{F,4,1}$.

Using (6.61) and (6.68) gives

$$\begin{aligned} 8(1+\lambda)(2-\lambda)A_2(r)(\rho_2^{F,6,1}-\rho_2^{F,4,2}) &= (-34+69\lambda-3\lambda^2+2\lambda^3)A_1^2(r) \\ &\quad + (-12+30\lambda-18\lambda^2+12\lambda^3)A_1(r)A_2(r) \\ &\quad + (14-19\lambda-27\lambda^2+18\lambda^3)A_2^2(r) \\ &=: a_{11}A_1^2(r)+a_{12}A_1(r)A_2(r)+a_{22}A_2^2(r), \end{aligned}$$

We note that a_{22} is the same as in the previous case and hence is non-negative. Differentiating with respect to lambda gives

$$\frac{da_{11}}{d\lambda}=69-6\lambda+6\lambda^2>0, \text{ and } \frac{da_{12}}{d\lambda}=30-36\lambda+36\lambda^2>0,$$

and so both have their maximum values at $\lambda=1/2$, both of which are zero and hence a_{11} and a_{12} are both non-positive. In order to prove that $\rho_2^{F,6,1}-\rho_2^{F,4,2}>0$ we use the following lemma:

Lemma 6.5.2. *Let* $A_i(r) \geq 0$ *for* $i = 1, 2$ *and define*

$$f := a_{11}A_1^2(r) + a_{12}A_1(r)A_2(r) + a_{22}A_2^2(r)$$

where $a_{11} \leq 0$*,* $a_{12} \leq 0$ *and* $a_{22} \geq 0$*. Further define*

$$g := -\frac{\beta A_1(r) + \gamma A_2(r)}{\alpha A_2(r)}$$

and suppose $-\alpha A_2(r) > 0$*,* $\beta > 0$*,* $\alpha + \gamma < 0$ *and* $g \leq 1$*. If* $h \geq 0$ *where*

$$h := a_{11}(\alpha + \gamma)^2 - a_{12}(\alpha + \gamma)\beta + a_{22}\beta^2,$$

then $f \geq 0$*.*

Proof Since $g \leq 1$, $-\alpha A_2(r) > 0$ and $\beta > 0$, it follows that

$$A_1(r) \leq -\frac{\alpha + \gamma}{\beta} A_2(r)$$

where both sides are non-negative, and hence we also have that

$$A_1^2(r) \leq \left(\frac{\alpha + \gamma}{\beta}\right)^2 A_2^2(r).$$

Recalling that a_{11} and a_{12} are both non-positive and inserting the two bounds above gives that

$$\beta^2 f \geq [a_{11}(\alpha + \gamma)^2 - a_{12}(\alpha + \gamma)\beta + a_{22}\beta^2]A_2^2(r).$$

Noting that $\beta^2 > 0$ and $A_2^2(r) \geq 0$ gives the result. $\square$

Setting $g = f_{4,2}$ from (6.63) gives $\alpha = -4(1 + \lambda)$, $\beta = 5 - \lambda$, $\gamma = 3 + \lambda$, and so $-\alpha A_2(r) > 0$, $\beta > 0$, $\alpha + \gamma = -1 - 3\lambda < 0$ and from the above, $a_{11}, a_{12} \leq 0$ and $a_{22} \geq 0$. These values give that $h = 256 - 768\lambda + 1024\lambda^3$, and so

$$\frac{dh}{d\lambda} = 3072\lambda^2 - 768 \leq 0,$$

giving a minimum value for h at $\lambda = 1/2$ with value zero. Hence $h \geq 0$ and applying Lemma 6.5.2 shows that $\rho_2^{F,6,1} > \rho_2^{F,4,2}$.

Moving on to consider $\rho_2^{F,5,1}$, we have from (6.64) and (6.68) that

$$\begin{aligned}8(2-\lambda)A_2(r)(\rho_2^{F,6,1}-\rho_2^{F,5,1}) &= (14-7\lambda-23\lambda^2+16\lambda^3)A_1^2(r)\\ &\quad+(28-22\lambda-22\lambda^2+16\lambda^3)A_1(r)A_2(r)\\ &\quad+(14-15\lambda-3\lambda^2+4\lambda^3)A_2^2(r)\\ &=: a_{11}A_1^2(r)+a_{12}A_1(r)A_2(r)+a_{22}A_2^2(r),\end{aligned}$$

from which we find

$$\begin{aligned}\frac{da_{11}}{d\lambda} &= -7-46\lambda+48\lambda^2<0,\\ \frac{da_{12}}{d\lambda} &= -22-44\lambda+48\lambda^2<0,\\ \frac{da_{22}}{d\lambda} &= -15-6\lambda+12\lambda^2<0.\end{aligned}$$

Hence the minimum values of a_{11}, a_{12} and a_{22} on $\lambda \in [0,1/2]$ are all at $\lambda = 1/2$ and are respectively 27/4, 27/2 and 25/4, showing that all three coefficients are positive and $\rho_2^{F,6,1} > \rho_2^{F,5,1}$.

The next case is $\rho_2^{F,5,2}$ and by (6.65) and (6.68) we have

$$\begin{aligned}8(2-\lambda)A_2(r)(\rho_2^{F,6,1}-\rho_2^{F,5,2}) &= (-34+113\lambda-71\lambda^2+16\lambda^3)A_1^2(r)\\ &\quad+(-12+78\lambda-62\lambda^2+16\lambda^3)A_1(r)A_2(r)\\ &\quad+(14-15\lambda-3\lambda^2+4\lambda^3)A_2^2(r)\\ &=: a_{11}A_1^2(r)+a_{12}A_1(r)A_2(r)+a_{22}A_2^2(r).\end{aligned}$$

Once again a_{22} is the same as in the previous case and is always positive. However, a_{11} and a_{12} do not have a fixed sign so we bound them from below by functions which do. Let $b_{11} = -34+68\lambda \le 0$, $b_{12} = -12+24\lambda \le 0$, and so

$$\begin{aligned}a_{11}-b_{11} &= \lambda(45-71\lambda+16\lambda^2) \ge 0,\\ a_{12}-b_{12} &= \lambda(54-62\lambda+16\lambda^2) \ge 0,\end{aligned}$$

which, setting $b_{22} := a_{22}$ gives that $8(2-\lambda)A_2(r)(\rho_2^{F,6,1}-\rho_2^{F,5,2}) \ge b_{11}A_1^2(r) + b_{12}A_1(r)A_2(r) + b_{22}A_2^2(r)$.

We now apply Lemma 6.5.2 with the a_{ij} given by b_{ij} and $g = f_{5,2}$ from (6.67), giving $\alpha = -4$, $\beta = 5-\lambda$, $\gamma = 3-2\lambda$, and so $-\alpha A_2(r) > 0$, $\beta > 0$, $\alpha+\gamma = -1-2\lambda < 0$ and from the above, $b_{11}, b_{12} \le 0$ and $b_{22} \ge 0$.

These values give that

$$\begin{aligned} h &= 256 - 955\lambda + 1125\lambda^2 + 60\lambda^3 - 208\lambda^4 + 64\lambda^5 \\ &\geq 256 - 955\lambda + 917\lambda^2 =: \tilde{h}, \end{aligned}$$

where $\frac{d\tilde{h}}{d\lambda} = 1834\lambda^2 - 955 < 0$, giving a minimum value for $\tilde{h}$ at $\lambda = 1/2$ with value -38. Hence $h \geq \tilde{h} > 0$ and applying Lemma 6.5.2 shows that $\rho_2^{F,6,1} > \rho_2^{F,5,2}$.

The final case is that of $\rho_2^{F,6,2}$ where from (6.68) and (6.69) we get

$$8(2-\lambda)A_2(r)(\rho_2^{F,6,1} - \rho_2^{F,6,2}) = 8\lambda(4-\lambda)A_1^2(r) + 4\lambda(8-3\lambda)A_1(r)A_2(r),$$

from which it is clear that both coefficients are non-negative and hence $\rho_2^{F,6,1} \geq \rho_2^{F,6,2}$ with equality only when $\lambda = 0$. This degeneracy when $\lambda = 0$ is because the pair density is invariant under rotation about the e_3 axis and all points lie in the e_1-e_2 plane. We have $f_{6,1} = 1/2 - \alpha$ and $f_{6,2} = 1/2 + \alpha$, where $\alpha = A_1(r)/(2A_2(r))$ and, analogously to the Oxygen case, this gives the same angle of separation and hence the same value at the two critical points.

6.5.2 Bond Angle

We are interested in the angle between the two electrons and note that $\rho_2^{F,6,1}$ corresponds to $\epsilon_{c\theta} = \pm 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = 1$, $\rho_2^{F,2}$ corresponds to $\epsilon_{c\theta} = \pm 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = \pm 1$ and both have $\sin^2\theta = 1$ (hence the lack of dependence on $\epsilon_{c\theta}$). We then see that

$$\sin^2\phi = \max\left\{0, \frac{4-3\lambda}{4(2-\lambda)} - \frac{A_1(r)}{A_2(r)}\frac{4+\lambda}{4(2-\lambda)}\right\}. \tag{6.72}$$

By assumption, $\sin\theta_1 = \sin\theta_2$ and so we have $\theta_1 = \theta_2 = \frac{\pi}{2}$. We also note that if $\sin^2\phi = 0$ then the choice of $\epsilon_{s\phi}$ is arbitrary and we can therefore take $\phi_2 = \pi - \phi_1$ in both cases, where ϕ_1 is given by (6.72). The angle between the electrons is then given by $\Phi = \pi - 2\phi_1$, where ϕ_1 is given by (6.72).

As with the Oxygen case, we may consider the effect of changing λ at a fixed value of r. Considering the function inside the max in (6.72) we see that

$$\frac{d}{d\lambda}\left[\frac{4-3\lambda}{4(2-\lambda)} - \frac{A_1(r)}{A_2(r)}\frac{4+\lambda}{4(2-\lambda)}\right] = -\frac{\big(1 + 3A_1(r)/A_2(r)\big)}{2(2-\lambda)^2} < 0,$$

and by a similar argument to the Oxygen case we see that we may choose ϕ_1 so that $\arcsin\phi_1$ is monotonically increasing with λ, and hence the bond angle is also

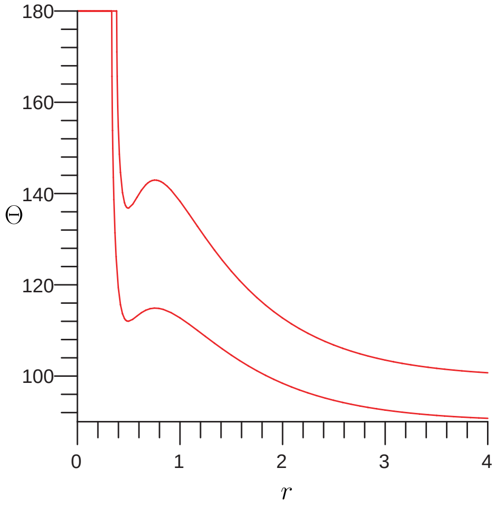


Figure 6.12: Bond angle range for Fluorine, the upper curve is for $\lambda = 1/2$ and the lower curve for $\lambda = 0$. The area between the two curves gives the range in which the angle may vary with no energy penalty.

monotonically increasing with λ. The bond angle range is shown in Figure 6.12.

As with the Neon case, there is no experimental data for the FH_2 molecule and hence we cannot test the accuracy of our predictions in this case. Again, for a multi-centre calculation, the FH_2 molecule should be energetically unstable.

As with the Oxygen case, the pair densities for the two extremal values of $\lambda = 0, 1/2$ are shown for a range of radial distances in Figures 6.13 and 6.14. The small value of r demonstrates the linearity at short range, the large value the $r \to \infty$ limits, and the middle value of $r = 0.8$ has been chosen as it is approximately the position of the outer maximum of the radial distribution.

The limiting values of $r \to \infty$ are once again given by noting that $A_2(r)$ decays exponentially slower than $A_1(r)$ and hence $\sin^2\phi = \max\{0, \frac{4-3\lambda}{4(2-\lambda)}\}$. For $\lambda = 0$, this once again gives $\Phi = 90°$ and when $\lambda = 1/2$ it gives $\Phi = \pi - 2\arcsin(\sqrt{5/12}) \sim 99.6°$.

It is worth noting that the specific values of r chosen are somewhat smaller than those for the Oxygen case, which is an effect of the contraction of the atoms across the period. In contrast to the Oxygen case, the $\lambda = 0$ case has the smallest angle, which can be seen from the relative direction of the monotonicities of the angles.

Figures 6.15 and 6.16 show the pair density along with the single particle density for $\lambda = 0$ and $\lambda = 1/2$ respectively. In general, the density is dominated by the p_2-orbital which is always present. For the $\lambda = 0$ case, the p_1 orbital is given full weighting and hence the density is invariant under rotation about the

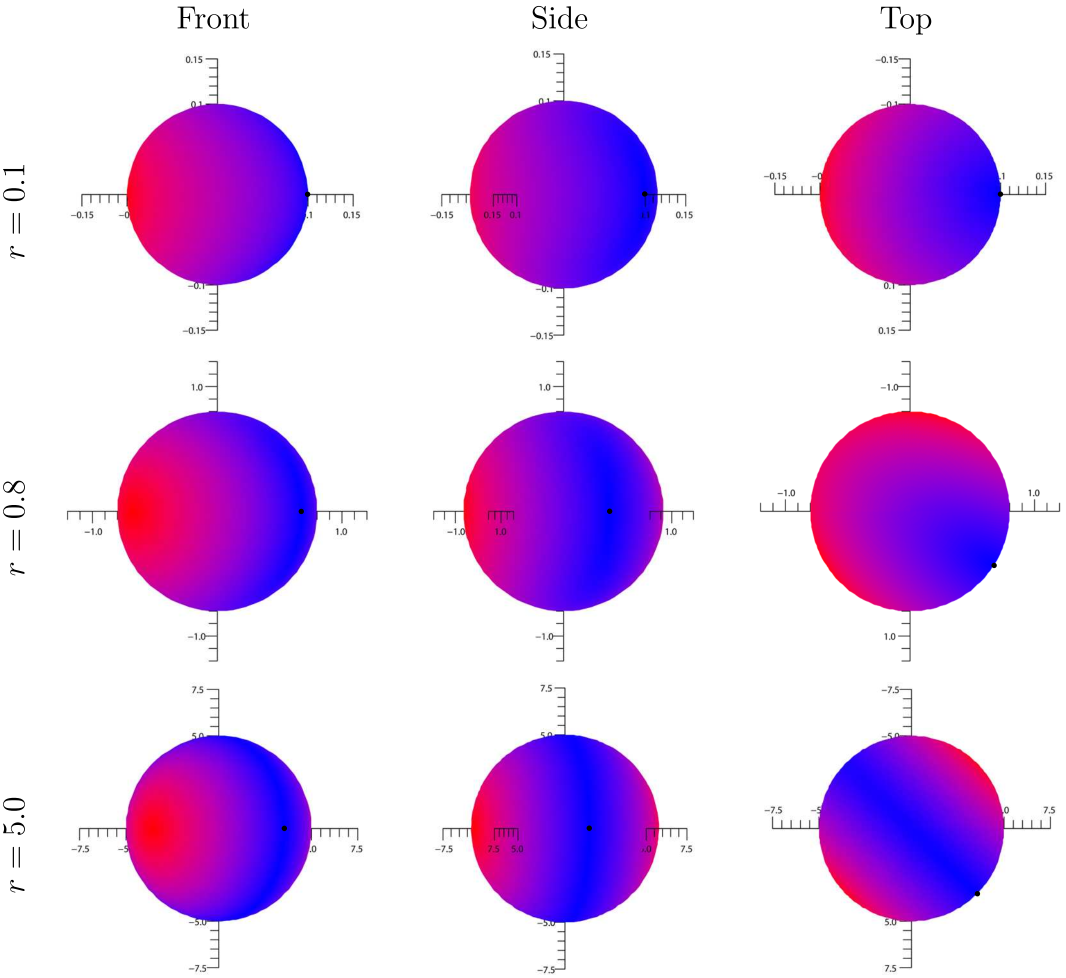


Figure 6.13: Fluorine pair densities for various distances with $\lambda = 0$. Front denotes the view along the positive e_1 axis, Top the view along the positive e_3 axis and Side the view with the e_1 axis out of the plane and the e_3 axis vertical. The black dot denotes the maximizing position of one of the electrons.

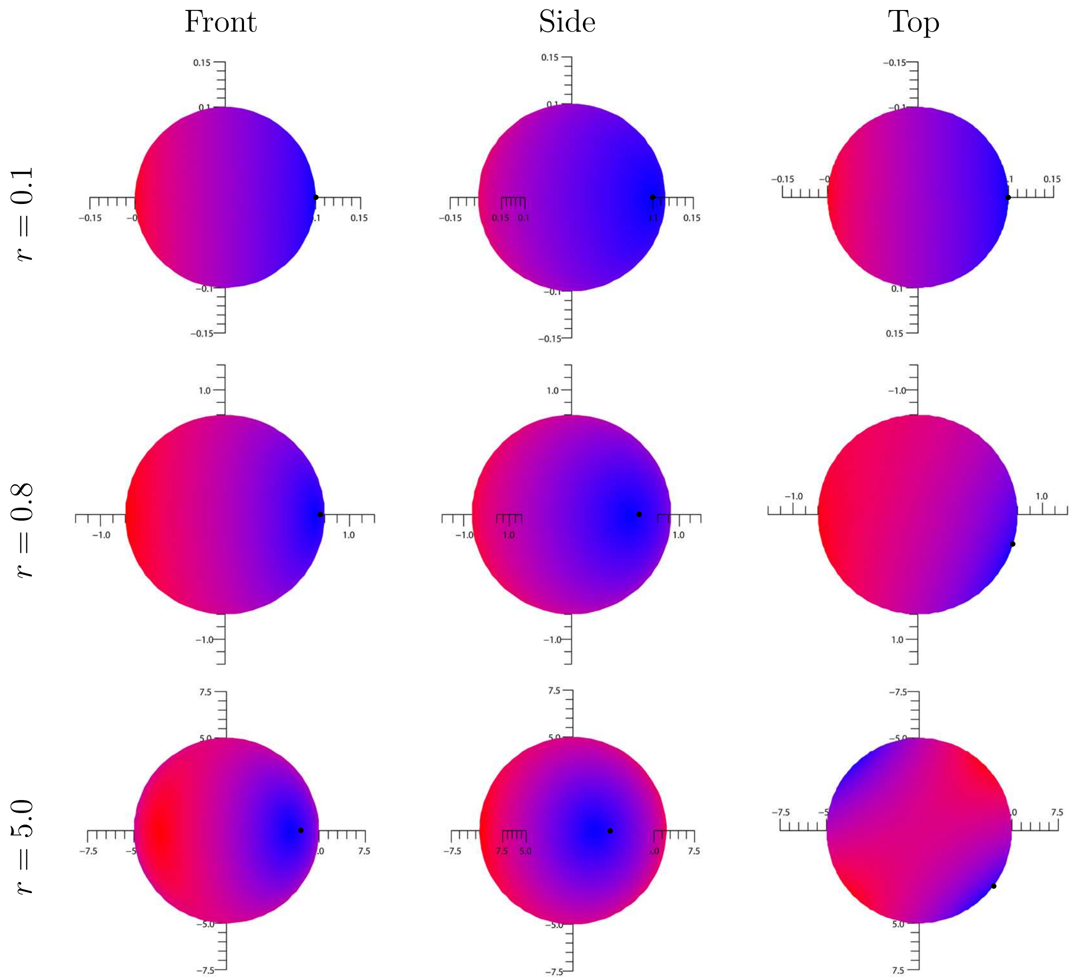


Figure 6.14: Fluorine pair densities for various distances with $\lambda = 1/2$, orientations and description as in Figure 6.13.

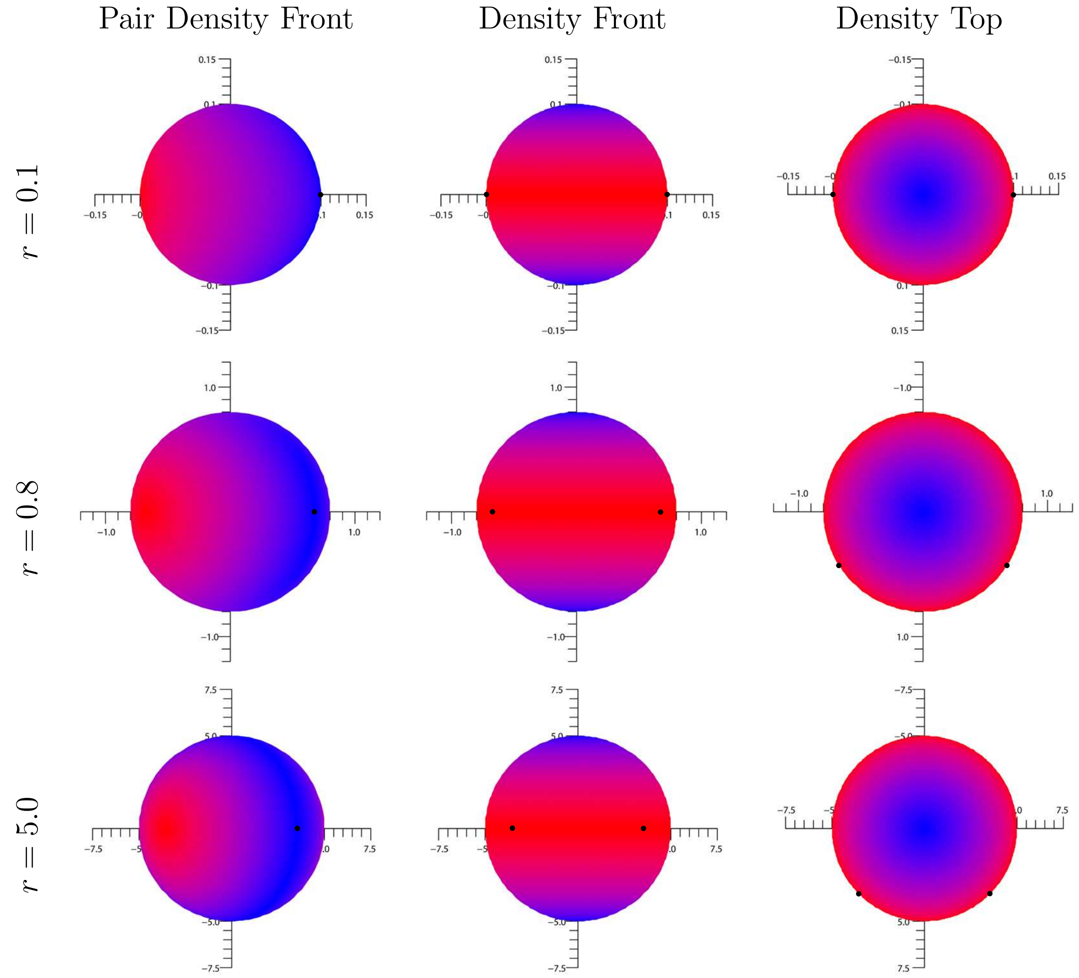


Figure 6.15: Fluorine pair density and single particle density with $\lambda = 0$, orientations as in Figure 6.13. The black dots denote the positions of the electrons in the maximum pair density.

e_3 axis. In contrast, in the $\lambda = 1/2$ case the p_1 and p_3 orbitals have the same weighting and the density is invariant under rotation about the e_2 axis.

Similar effects to the Oxygen case can be seen, in that for the $\lambda = 0$ case the two maximizing points of the pair density lie in the plane of maximum density. In the $\lambda = 1/2$ case we see the same effect as with the $\lambda = 0$ case of Oxygen where the maximizing points move away from the maximum of the density.

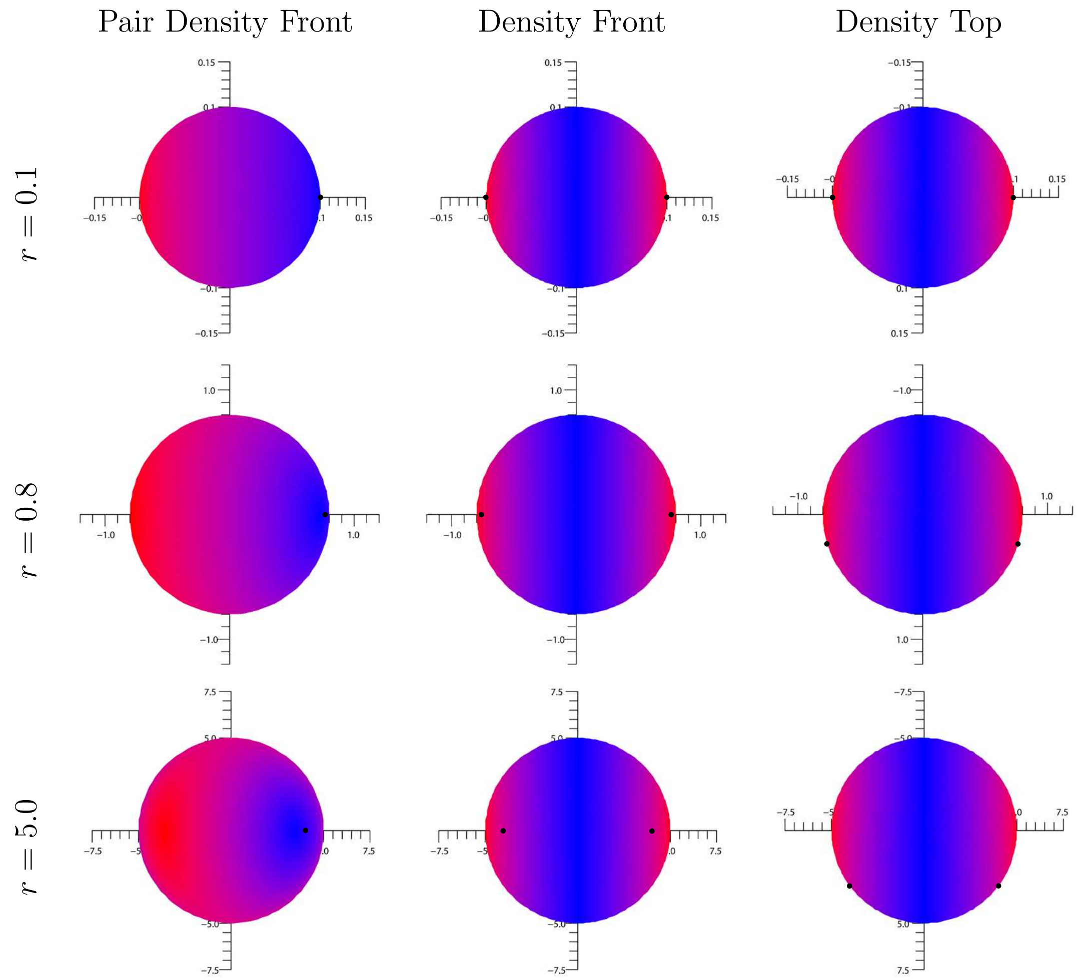


Figure 6.16: Fluorine pair density and single particle density with $\lambda = 1/2$, details as in Figure 6.15.

6.6 Beryllium

We now move on to consider atoms which have a correlated ground state. In these cases we have an extra parameter which we may vary - the relative phase between the two correlated wavefunctions. However, we claim that in order to attain the minimum energy, this relative phase is in fact fixed.

6.6.1 Uniqueness of Correlated Wavefunctions

Lemma 6.6.1. *Let Ψ_1 and Ψ_2 be two wavefunctions which are coupled by H, i.e. $\langle\Psi_1|H|\Psi_2\rangle \neq 0$. Then the normalized correlated wavefunction Ψ which minimizes the energy is unique, up to an overall phase factor $e^{i\mu}$, $\mu \in \mathbb{R}$. Furthermore, the expected value of any linear operator on Ψ is independent of this phase factor.*

Proof Consider a general interaction matrix

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$$

where $a, c \in \mathbb{R}$ and $b \in \mathbb{C}$. In our case we have $a = \langle H\rangle_{\Psi_1}, c = \langle H\rangle_{\Psi_2}$ and $b = \langle\Psi_1|H|\Psi_2\rangle$. We see that the normalized eigenvectors of the form $\Psi = \frac{1}{\|\Psi_1+y\Psi_2\|}(\Psi_1 + y\Psi_2)$ satisfy

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix}\begin{pmatrix} 1 \\ y \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ y \end{pmatrix}$$

and hence we have

$$a + by = \lambda \text{ and } b^* + cy = \lambda y.$$

It follows that

$$y = \tfrac{\lambda-a}{b} = \tfrac{b^*}{\lambda-c} \tag{6.73}$$

and we recover the energies

$$\lambda_\pm = \tfrac{a+c}{2} \pm \sqrt{\left(\tfrac{a-c}{2}\right)^2 + |b|^2}.$$

Note in particular that, for a fixed energy λ, (6.73) uniquely determines y.

It is clear that, for any linear operator $\mathcal{O}$, we have

$$\langle\mathcal{O}\rangle_\Psi = \frac{1}{\|\Psi_1 + y\Psi_2\|^2}\Big(\langle\mathcal{O}\rangle_{\Psi_1} + \langle\mathcal{O}\rangle_{\Psi_2} + 2\mathrm{Re}\big(\langle\Psi_1|\mathcal{O}|\Psi_2\rangle\big)\Big),$$

all three terms of which are trivially invariant under multiplying both Ψ_1 and Ψ_2 by a phase factor $e^{i\mu}$. Noting that we may choose $\mathcal{O} = H$ gives the result. $\square$

In particular, it follows trivially that the pair density is invariant under choice of overall phase of the wavefunction. Further, we may choose the coefficient y to be real and absorb any phase factor into a new wavefunction $\tilde{\Psi_2} := e^{i\nu}\Psi_2$, i.e. $\Psi = c\Psi_1 + \sqrt{1-c^2}\tilde{\Psi_2}$. In addition (6.73) shows that, if both Ψ_1 and Ψ_2 may be made real by multiplying by suitable phase factors, then we may take y to be real.

6.6.2 Canonical Ground State Pair Density

We now consider the case of Beryllium which has a general ground state wavefunction of the form

$$\Psi^{Be} = c|1s\overline{1s}2s\overline{2s}\rangle - \frac{\sqrt{1-c^2}}{\sqrt{3}}\big(|1s\overline{1s}p_1\overline{p_1}\rangle + |1s\overline{1s}p_2\overline{p_2}\rangle + |1s\overline{1s}p_3\overline{p_3}\rangle\big) \tag{6.74}$$

where $c \in \mathbb{R}$ is a variationally determined parameter. Note that both parts of the wavefunction are real so this restriction to $c \in \mathbb{R}$ is allowed by Lemma 6.6.1. Since c is close to one (see Table 4.6), we expect the first Slater determinant to dominate at short ranges (where $1s$- and $2s$-orbitals dominate) but the other part of the wavefunction will dominate at longer ranges due to the slower decay rate of the p-orbitals.

We first rewrite the wavefunction as

$$\Psi^{Be} = c\Psi_1 - \sqrt{1-c^2}\Psi_2,$$

where Ψ_1 and Ψ_2 are the natural choices from (6.74), and note that the pair density is given by

$$\rho_2^{Be} = c^2\rho_2^{\Psi_1} + (1-c^2)\rho_2^{\Psi_2} - 2c\sqrt{1-c^2}\mathrm{Re}\langle\Psi_1|V_{x,y}|\Psi_2\rangle.$$

It is clear that $\rho_2^{\Psi_1}$ has no angular dependence but we will compute it for completeness. For a general Slater determinant $|1s\overline{1s}\psi\overline{\psi}\rangle$ the pair density is given by

$$\begin{aligned}\rho_2^{|1s\overline{1s}\psi\overline{\psi}\rangle} &= |1s(x)|^2|1s(y)|^2 + 2|1s(x)|^2|\psi(y)|^2 + 2|\psi(x)|^2|1s(y)|^2 \\ &\quad + |\psi(x)|^2|\psi(y)|^2 - 2\cdot 1s(x)1s(y)\psi(x)\psi(y).\end{aligned}$$

The other terms in the pair density of Ψ^{Be} are of the form

$$\langle 1s\overline{1s}\psi\overline{\psi}|V_{x,y}|1s\overline{1s}\varphi\overline{\varphi}\rangle = \psi(x)\psi(y)\varphi(x)\varphi(y),$$

where $\psi \neq \varphi$. Hence we see that

$$\begin{aligned}
\rho_2^{\Psi_1} &= |1s(x)|^2|1s(y)|^2 + 2|1s(x)|^2|2s(y)|^2 + 2|2s(x)|^2|1s(y)|^2 \\
&\quad + |2s(x)|^2|2s(y)|^2 - 2\cdot 1s(x)1s(y)2s(x)2s(y), \\
\rho_2^{\Psi_2} &= |1s(x)|^2|1s(y)|^2 + \frac{2}{3}\sum_{1\le i<j\le 3} p_i(x)p_i(y)p_j(x)p_j(y) \\
&\quad + \frac{1}{3}\Big[2|1s(x)|^2\sum_{i=1}^{3}|p_i(y)|^2 + 2|1s(y)|^2\sum_{i=1}^{3}|p_i(x)|^2 \\
&\qquad + \sum_{i=1}^{3}|p_i(x)|^2|p_i(y)|^2 - 2\cdot 1s(x)1s(y)\sum_{i=1}^{3}p_i(x)p_i(y)\Big],
\end{aligned}$$

and the final term is given by

$$\langle\Psi_1|V_{x,y}|\Psi_2\rangle = \frac{1}{\sqrt{3}}2s(x)2s(y)\sum_{i=1}^{3}p_i(x)p_i(y).$$

Using the definitions of $R_i(r)$ from (6.8), this gives

$$\begin{aligned}
\rho_2^{Be} &= F(r) + \tfrac{1-c^2}{3}\big[R_3^4(r)(x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 + 2x_1y_1x_2y_2 + 2x_1y_1x_3y_3 \\
&\qquad + 2x_2y_2x_3y_3) - 2R_1^2(r)R_3^2(r)(x_1y_1 + x_2y_2 + x_3y_3)\big] \\
&\quad - \tfrac{2c\sqrt{1-c^2}}{\sqrt{3}}R_2^2(r)R_3^2(r)(x_1y_1 + x_2y_2 + x_3y_3) \\
&= F(r) + \tfrac{1-c^2}{3}\left[R_3^4(r)r^4\cos^2\Theta - 2R_1^2(r)R_3^2(r)r^2\cos\Theta\right] \\
&\quad - \tfrac{2c\sqrt{1-c^2}}{\sqrt{3}}R_2^2(r)R_3^2(r)r^2\cos\Theta \\
&=: A_0(r) - A_1(r)\cos\Theta + A_2(r)\cos^2\Theta,
\end{aligned}$$

where

$$A_0(r) := F(r) = R_1^4(r) + c^2\big(2R_1^2(r)R_2^2(r) + R_2^2(r)\big) + \tfrac{4}{3}(1-c^2)R_1^2(r)r^2R_3^2(r),$$

$A_1(r) := 2\frac{1-c^2}{3}R_1^2(r)R_3^2(r)r^2 + \frac{2c\sqrt{1-c^2}}{\sqrt{3}}R_2^2(r)R_3^2(r)r^2$, and $A_2(r) := \frac{1-c^2}{3}R_3^4(r)r^4$. Recall that $A_i(r) > 0$, $i = 1, 2$.

Consider $\rho_2^{Be} = A_0(r) - A_1(r)\cos\Theta + A_2(r)\cos^2\Theta$, giving

$$\begin{aligned}\frac{\partial}{\partial\Theta}\rho_2^{Be} &= A_1(r)\sin\Theta - 2A_2(r)\sin\Theta\cos\Theta,\\ \frac{\partial^2}{\partial\Theta^2}\rho_2^{Be} &= A_1(r)\cos\Theta - 2A_2(r)(\cos^2\Theta - \sin^2\Theta).\end{aligned}$$

It is clear that the critical points lie at $\sin\Theta = 0$ and $\cos\Theta = \frac{A_1(r)}{2A_2(r)}$.

If $\sin\Theta = 0$ then we have $\cos\Theta = \pm 1$ and hence $\frac{\partial^2}{\partial\Theta^2}\rho_2^{Be} = \pm A_1(r) - 2A_2(r)$. The case for $\cos\Theta = -1$ gives $-A_1(r) - 2A_2(r)$ and is clearly a maximum; if $\cos\Theta = +1$, we have $A_1(r) - 2A_2(r)$, which is negative (and hence a maximum) when $\frac{A_1(r)}{2A_2(r)} < 1$. However, comparing the two densities which are respectively $A_0(r) + A_1(r) + A_2(r)$ and $A_0(r) - A_1(r) + A_2(r)$, it is clear that the $\cos\Theta = -1$ case is always the global maximum. It follows trivially that, when it exists, $\cos\Theta = \frac{A_1(r)}{2A_2(r)}$ must be a minimum and hence the predicted bond angle is 180° for all r. As can be seen from Table 6.1, this agrees with the experimental data. We have shown:

Theorem 6.6.2 (Maximum of Beryllium Pair Density)**.** *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define*

$$A_0(r) := R_1^4(r) + c^2\big(2R_1^2(r)R_2^2(r) + R_2^2(r)\big) + \tfrac{4}{3}(1-c^2)R_1^2(r)R_3^2(r),$$

$A_1(r) := 2\frac{1-c^2}{3}R_1^2(r)R_3^2(r) + \frac{2c\sqrt{1-c^2}}{\sqrt{3}}R_2^2(r)R_3^2(r)$, *and* $A_2(r) := \frac{1-c^2}{3}R_3^4(r)$, *for* $c \in [0,1]$. *Then the maximum of the Beryllium ground state pair density*

$$\rho_2^{Be} = A_0(r) - A_1(r)\cos\Theta + A_2(r)\cos^2\Theta$$

is at $\Theta = 180°$ *for all* r.

6.7 Carbon

6.7.1 Canonical Ground State Pair Density

We wish to find a general form for the Carbon ground state wavefunction and by Lemma 6.2.1 we may choose the case $S_3\Psi = \Psi$. The form used to calculate the energy was

$$\Psi = c|1s\overline{1s}2s\overline{2s}p_1p_2\rangle - \sqrt{1-c^2}|1s\overline{1s}p_1p_2p_3\overline{p_3}\rangle,$$

which is an eigenfunction of L_3 with eigenvalue zero. By an almost identical argument as for Fluorine in Section 6.5 we know that we may write a general form of the first Slater determinant as

$$\Psi_1 := \alpha_1|1s\overline{1s}2s\overline{2s}p_2p_3\rangle + \alpha_2|1s\overline{1s}2s\overline{2s}p_1p_3\rangle + \alpha_3|1s\overline{1s}2s\overline{2s}p_1p_2\rangle$$

where $\alpha_i \in \mathbb{C}$, $\sum|\alpha_i|^2 = 1$, which has a pair density of the form

$$\begin{aligned}\rho_2^{\Psi_1}(x,y) = \rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x,y)&\\ - \tfrac{1}{2}\big[\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)x\big) + x\leftrightarrow y\big]&\\ + \Big(1s(x)1s(y) + 2s(x)2s(y) + \sum_{i=1}^{3}p_i(x)p_i(y)\Big)&\\ \times k^2e^{-Z(|x|+|y|)/2}\big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)y\big),&\end{aligned}$$

where $k\big((a_1+ib_1)\cdot x\big)e^{-Z|x|/2} := \alpha_1p_1(x) - \alpha_2p_2(x) + \alpha_3p_3(x)$. We also note that by Lemma 6.6.1 we may choose c to be real and choose Ψ_2 with a suitable phase factor.

Similarly, following the argument from the Oxygen case in Section 6.3.1, we may write the general form of the second Slater determinant as

$$\begin{aligned}\Psi_2 :&= \beta_1|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle + \beta_2|1s\overline{1s}p_2\overline{p_2}p_1p_3\rangle + \beta_3|1s\overline{1s}p_3\overline{p_3}p_1p_2\rangle\\ &= \beta_1|1s\overline{1s}p_1p_2p_3\overline{p_1}\rangle - \beta_2|1s\overline{1s}p_1p_2p_3\overline{p_2}\rangle + \beta_3|1s\overline{1s}p_1p_2p_3\overline{p_3}\rangle\end{aligned}$$

where $\beta_i \in \mathbb{C}$, $\sum|\beta_i|^2 = 1$, which gives the pair density

$$\begin{aligned}\rho_2^{\Psi_2}(x,y) = \rho_2^{|1s\overline{1s}p_1p_2p_3\rangle}(x,y)&\\ + \tfrac{1}{2}\big[\rho^{|1s\overline{1s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big(x\cdot(a_2\otimes a_2 + b_2\otimes b_2)x\big) + x\leftrightarrow y\big]&\\ - 1s(x)1s(y)k^2e^{-Z(|x|+|y|)/2}\big(x\cdot(a_2\otimes a_2 + b_2\otimes b_2)y\big),&\end{aligned}$$

where $k\big((a_2+ib_2)\cdot x\big)e^{-Z|x|/2} := \beta_1p_1(x) - \beta_2p_2(x) + \beta_3p_3(x)$.

The remaining term in the pair density we need to calculate is $\langle\Psi_1|V_{x,y}|\Psi_2\rangle$. We begin with $\langle\Psi_1|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle$ and, since all required Slater determinants differ by two orbitals, we have

$$\begin{aligned}&\langle\Psi_1|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle\\ &= \alpha_1\langle 1s\overline{1s}2s\overline{2s}p_2p_3|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle + \alpha_2\langle 1s\overline{1s}2s\overline{2s}p_1p_3|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle\\ &\quad + \alpha_3\langle 1s\overline{1s}2s\overline{2s}p_1p_2|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle\\ &= \alpha_1\langle 1s\overline{1s}2s\overline{2s}p_2p_3|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle - \alpha_2\langle 1s\overline{1s}2s\overline{2s}p_1p_3|V_{x,y}|1s\overline{1s}p_2\overline{p_1}p_1p_3\rangle\\ &\quad + \alpha_3\langle 1s\overline{1s}2s\overline{2s}p_1p_2|V_{x,y}|1s\overline{1s}p_3\overline{p_1}p_1p_2\rangle\\ &= \tfrac{1}{2}\cdot 2s(x)2s(y)\big[\alpha_1p_1(x)p_1(y) - \alpha_2p_2(x)p_1(y) + \alpha_3p_3(x)p_1(y) + x\leftrightarrow y\big],\end{aligned}$$

and similarly

$$\begin{aligned}&\langle\Psi_1|V_{x,y}|1s\overline{1s}p_2\overline{p_2}p_1p_3\rangle\\ &= -\tfrac{1}{2}\cdot 2s(x)2s(y)\big[\alpha_1 p_1(x)p_2(y) - \alpha_2 p_2(x)p_2(y) + \alpha_3 p_3(x)p_2(y) + x\leftrightarrow y\big]\end{aligned}$$

$$\begin{aligned}&\langle\Psi_1|V_{x,y}|1s\overline{1s}p_3\overline{p_3}p_1p_2\rangle\\ &= \tfrac{1}{2}\cdot 2s(x)2s(y)\big[\alpha_1 p_1(x)p_3(y) - \alpha_2 p_2(x)p_3(y) + \alpha_3 p_3(x)p_3(y) + x\leftrightarrow y\big].\end{aligned}$$

It is now clear that

$$\begin{aligned}\langle\Psi_1|V_{x,y}|\Psi_2\rangle = \tfrac{1}{2}\cdot 2s(x)2s(y)\big[&\alpha_1\beta_1^* p_1(x)p_1(y) - \alpha_2\beta_1^* p_2(x)p_1(y)\\ &+ \alpha_3\beta_1^* p_3(x)p_1(y) - \alpha_1\beta_2^* p_1(x)p_2(y) + \alpha_2\beta_2^* p_2(x)p_2(y)\\ &- \alpha_3\beta_2^* p_3(x)p_2(y) + \alpha_1\beta_3^* p_1(x)p_3(y) - \alpha_2\beta_3^* p_2(x)p_3(y)\\ &+ \alpha_3\beta_3^* p_3(x)p_3(y) + x\leftrightarrow y\big],\end{aligned}$$

which we rewrite as

$$\begin{aligned}\langle\Psi_1|V_{x,y}|\Psi_2\rangle = \tfrac{1}{2}\cdot 2s(x)2s(y)\big[&\big(\alpha_1 p_1(x) - \alpha_2 p_2(x) + \alpha_3 p_3(x)\big)\\ &\times\big(\beta_1^* p_1(y) - \beta_2^* p_2(y) + \beta_3^* p_3(y)\big) + x\leftrightarrow y\big].\end{aligned}$$

Using the definitions of a_i and b_i above, this is equivalent to

$$\tfrac{1}{2}\cdot 2s(x)2s(y)\big[k^2 e^{-\frac{|x|}{2}} e^{-\frac{|y|}{2}}\big((a_1+ib_1)\cdot x\big)\big((a_2-ib_2)\cdot y\big) + x\leftrightarrow y\big].$$

Hence the canonical Carbon pair density for wavefunctions is

$$\begin{aligned}&\rho_2^{\Psi}(x,y) \qquad (6.75)\\ &= c^2\Big[\rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x,y)\\ &\qquad - \tfrac{1}{2}\big[\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)x\big) + x\leftrightarrow y\big]\\ &\qquad + \Big(1s(x)1s(y) + 2s(x)2s(y) + \sum_{i=1}^{3} p_i(x)p_i(y)\Big)\\ &\qquad\qquad \times k^2e^{-Z(|x|+|y|)/2}\big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)y\big)\Big]\\ &\quad + (1-c^2)\Big[\rho_2^{|1s\overline{1s}p_1p_2p_3\rangle}(x,y)\\ &\qquad + \tfrac{1}{2}\big[\rho^{|1s\overline{1s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big(x\cdot(a_2\otimes a_2 + b_2\otimes b_2)x\big) + x\leftrightarrow y\big]\\ &\qquad - 1s(x)1s(y)k^2e^{-Z(|x|+|y|)/2}\big(x\cdot(a_2\otimes a_2 + b_2\otimes b_2)y\big)\Big]\end{aligned}$$

$$
\begin{aligned}
- c\sqrt{1-c^2}\mathrm{Re}\Big[&2s(x)2s(y)k^2e^{-Z(|x|+|y|)/2} \\
&\qquad\qquad\qquad \times \big((a_1+ib_1)\cdot x\big)\big((a_2-ib_2)\cdot y\big) + x \leftrightarrow y\Big].
\end{aligned}
$$

We now simultaneously rotate x and y to bring Ψ_1 to the canonical form

$$\Psi_1 = \sqrt{1-\lambda}|1s\overline{1s}2s\overline{2s}p_2p_3\rangle + i\sqrt{\lambda}|1s\overline{1s}2s\overline{2s}p_1p_3\rangle$$

where λ takes some value in $[0,1/2]$. We wish to use the fact that Ψ must be a minimizer of the energy to determine restrictions on the β_i. Suppose we have two general wavefunctions $\Phi = \sum\alpha_i|\Phi_i\rangle$ with $\sum|\alpha_i|^2 = 1$ and $\tilde{\Phi} = \sum\tilde{\alpha}_i|\Phi_i\rangle$ with $\sum|\tilde{\alpha}_i|^2 = 1$. Suppose further that the Slater determinants forming Φ and $\tilde{\Phi}$ satisfy $\langle\Phi_i|V_{ee}|\Phi_i\rangle = \langle\Phi_j|V_{ee}|\Phi_j\rangle\ \forall i,j$ and $\langle\Phi_i|V_{ee}|\Phi_j\rangle = 0\ \forall i \neq j$, then we clearly have $\langle\Phi|V_{ee}|\Phi\rangle = \langle\tilde{\Phi}|V_{ee}|\tilde{\Phi}\rangle$, i.e. the choice of coefficients does not affect the individual energies.

It is clear that the Slater determinants in each of Ψ_1 and Ψ_2 satisfy these conditions (due to rotational symmetry and orthogonality of the angular terms). It follows that only the term $2c\sqrt{1-c^2}\mathrm{Re}\langle\Psi_1|V_{ee}|\Psi_2\rangle$ (where we have used that $c\in\mathbb{R}$), will affect the energy. For the canonical Ψ_1 and a general Ψ_2 this gives

$$
\begin{aligned}
&\langle\Psi_1|V_{ee}|\Psi_2\rangle = \sqrt{1-\lambda}\big(\beta_1^*\langle 1s\overline{1s}2s\overline{2s}p_2p_3|V_{ee}|1s\overline{1s}p_1\overline{p_1}p_2p_3\rangle \\
&- \beta_2^*\langle 1s\overline{1s}2s\overline{2s}p_2p_3|V_{ee}|1s\overline{1s}p_1\overline{p_2}p_2p_3\rangle + \beta_3^*\langle 1s\overline{1s}2s\overline{2s}p_2p_3|V_{ee}|1s\overline{1s}p_1\overline{p_3}p_2p_3\rangle\big) \\
&+ i\sqrt{\lambda}\big(- \beta_1^*\langle 1s\overline{1s}2s\overline{2s}p_1p_3|V_{ee}|1s\overline{1s}p_2\overline{p_1}p_1p_3\rangle \\
&+ \beta_2^*\langle 1s\overline{1s}2s\overline{2s}p_1p_3|V_{ee}|1s\overline{1s}p_2\overline{p_2}p_1p_3\rangle - \beta_3^*\langle 1s\overline{1s}2s\overline{2s}p_1p_3|V_{ee}|1s\overline{1s}p_2\overline{p_3}p_1p_3\rangle\big).
\end{aligned}
$$

Noting that each integral has Slater determinants that differ by two orbitals, we have

$$
\begin{aligned}
\langle\Psi_1|V_{ee}|\Psi_2\rangle = \sqrt{1-\lambda}&\big(\beta_1^*(2sp_1|p_12s) - \beta_2^*(2sp_1|p_22s) + \beta_3^*(2sp_1|p_32s)\big) \\
&+ i\sqrt{\lambda}\big(- \beta_1^*(2sp_2|2sp_1) + \beta_2^*(2sp_2|p_22s) - \beta_3^*(2sp_2|p_32s)\big).
\end{aligned}
$$

Now note that for $i \neq j$, $(2sp_i|p_j2s) = 0$ and $(2sp_i|p_i2s) = (2sp_j|p_j2s)$, which follow from orthogonality of the angular parts and rotational symmetry respectively. It is therefore clear that

$$\langle\Psi_1|V_{ee}|\Psi_2\rangle = \big(\sqrt{1-\lambda}\beta_1^* + i\sqrt{\lambda}\beta_2^*\big)(2sp_1|p_12s).$$

For the ground state energy we require $\langle \Psi_1 | V_{ee} | \Psi_2 \rangle = (2sp_1|p_1 2s)$, giving

$$E = \tfrac{a+c}{2} \pm \sqrt{\left(\tfrac{a-c}{2}\right)^2 + |b|^2}.$$

where $a := \langle \Psi_1 | V_{ee} | \Psi_1 \rangle, c := \langle \Psi_2 | V_{ee} | \Psi_2 \rangle$ and $b := \langle \Psi_1 | V_{ee} | \Psi_2 \rangle$. Hence, for Ψ to be a minimizing wavefunction it must follow that

$$\sqrt{1-\lambda}\beta_1^* + i\sqrt{\lambda}\beta_2^* = 1, \quad |\beta_1|^2 + |\beta_2|^2 + |\beta_3|^2 = 1.$$

It is trivial to show (by considering the equations for $\beta_1 = \sqrt{1-\lambda}(1+\gamma_1)$ and $\beta_2 = i\sqrt{\lambda}(1+\gamma_2)$, $\gamma_i \in \mathbb{C}$) that the only solutions are

$$\beta_1 = \sqrt{1-\lambda}, \quad \beta_2 = i\sqrt{\lambda}, \quad \beta_3 = 0,$$

which gives a canonical one-parameter wavefunction with corresponding one-parameter family of pair densities.

As in previous cases, for ease of notation, choosing $\alpha_3 = \sqrt{1-\lambda}$ and $\alpha_1 = i\sqrt{\lambda}$, (and hence $\beta_3 = \sqrt{1-\lambda}$ and $\beta_1 = i\sqrt{\lambda}$), we have a canonical pair density

$$\begin{aligned}
&\rho_2^{\Psi}(x,y) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.76)\\
&= c^2\Big[\rho_2^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(x,y)\\
&\qquad - \tfrac{1}{2}\big[\rho^{|1s\overline{1s}2s\overline{2s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big((1-\lambda)x_3^2 + \lambda x_1^2\big) + x \leftrightarrow y\big]\\
&\qquad + \Big(1s(x)1s(y) + 2s(x)2s(y) + \sum_{i=1}^{3} p_i(x)p_i(y)\Big)\\
&\qquad\qquad\qquad \times k^2 e^{-Z(|x|+|y|)/2}\big((1-\lambda)x_3y_3 + \lambda x_1y_1\big)\Big]\\
&\quad + (1-c^2)\Big[\rho_2^{|1s\overline{1s}p_1p_2p_3\rangle}(x,y)\\
&\qquad\qquad + \tfrac{1}{2}\big[\rho^{|1s\overline{1s}p_1p_2p_3\rangle}(y)k^2e^{-Z|x|}\big((1-\lambda)x_3^2 + \lambda x_1^2\big) + x \leftrightarrow y\big]\\
&\qquad\qquad - 1s(x)1s(y)k^2e^{-Z(|x|+|y|)/2}\big((1-\lambda)x_3y_3 + \lambda x_1y_1\big)\Big]\\
&\quad - 2c\sqrt{1-c^2}\Big[2s(x)2s(y)k^2e^{-Z(|x|+|y|)/2}\big((1-\lambda)x_3y_3 + \lambda x_1y_1\big)\Big],
\end{aligned}$$

given by the wavefunction

$$\begin{aligned}
\Psi^C &= c\big(\sqrt{1-\lambda}|1s\overline{1s}2s\overline{2s}p_1p_2\rangle + i\sqrt{\lambda}|1s\overline{1s}2s\overline{2s}p_2p_3\rangle\big)\\
&\quad + \sqrt{1-c^2}\big(\sqrt{1-\lambda}|1s\overline{1s}p_1p_2p_3\overline{p_3}\rangle + i\sqrt{\lambda}|1s\overline{1s}p_1p_2p_3\overline{p_1}\rangle\big).
\end{aligned}$$

By Corollary 6.3.4 we set $\cos\theta_1 = \epsilon_{c\theta}\cos\theta_2, \sin\phi_1 = \epsilon_{s\phi}\sin\phi_2$ and $\cos\phi_1 =$

$\epsilon_{c\phi}\cos\phi_2$ where $\epsilon_i = \pm 1$. Recalling the notation from (6.8) and denoting the angle between x and y by Θ, where $\cos\Theta = \epsilon_{c\theta}\cos^2\theta + \sin^2\theta(\epsilon_{c\phi}\cos^2\phi + \epsilon_{s\phi}\sin^2\phi)$, we see that

$$\begin{aligned}
\rho^{|1s\overline{1s}2s\overline{2s}sp_1p_2p_3\rangle} &= 2R_1^2(r) + 2R_2^2(r) + r^2R_3^2(r),\\
\rho^{|1s\overline{1s}sp_1p_2p_3\rangle} &= 2R_1^2(r) + r^2R_3^2(r),\\
\rho_2^{|1s\overline{1s}2s\overline{2s}sp_1p_2p_3\rangle} &= F_1(r) - r^2R_3^2(r)\big(R_1^2(r) + R_2^2(r)\big)\cos\Theta - \tfrac{1}{2}r^4R_3^4(r)\cos^2\Theta,\\
\rho_2^{|1s\overline{1s}sp_1p_2p_3\rangle} &= F_2(r) - r^2R_3^2(r)R_1^2(r)\cos\Theta - \tfrac{1}{2}r^4R_3^4(r)\cos^2\Theta,
\end{aligned}$$

where $F_1(r)$ is given by (6.11) and $F_2(r) = R_1^4(r) + 2R_1^2(r)r^2R_3^2(r) + \frac{1}{2}r^4R_3^4(r)$.

Using the identity $\cos^2\varphi = 1 - \sin^2\varphi$, this gives a pair density of the form

$$\begin{aligned}
\rho_2^C = F(r) &+ \big[\epsilon_{c\theta} + \sin^2\theta(\epsilon_{c\phi} - \epsilon_{c\theta}) + \sin^2\theta\sin^2\phi(\epsilon_{s\phi} - \epsilon_{s\phi})\big]\\
&\qquad\qquad \times \big[- c^2r^2R_3^2(r)R_2^2(r) - r^2R_3^2(r)R_1^2(r)\big]\\
&+ \big[1 + 2(\epsilon_{c\theta}\epsilon_{c\phi} - 1)\sin^2\theta + 2\epsilon_{c\theta}(\epsilon_{s\phi} - \epsilon_{c\phi})\sin^2\theta\sin^2\phi + 2(1 - \epsilon_{c\phi}\epsilon_{c\theta})\sin^4\theta\\
&\quad + 2(\epsilon_{c\phi} - \epsilon_{c\theta})(\epsilon_{s\phi} - \epsilon_{c\phi})\sin^4\theta\sin^2\phi + 2(1 - \epsilon_{s\phi}\epsilon_{c\phi})\sin^4\theta\big] \cdot \tfrac{1}{2}r^4R_3^4(r)\\
&+ \big[(1-\lambda)(1 - \sin^2\theta) + \lambda\sin^2\theta\sin^2\phi\big]\\
&\qquad \times \big[(1 - 2c^2)\big(2R_1^2(r) + r^2R_3^2(r)\big)r^2R_3^2(r) - c^2R_2^2(r)r^2R_3^2(r)\big]\\
&+ \big[(1-\lambda)\epsilon_{c\theta}(1 - \sin^2\theta) + \lambda\epsilon_{s\phi}\sin^2\theta\sin^2\phi\big]\\
&\qquad \times \big[(2c^2 - 1)R_1^2(r) + (c^2 - 2c\sqrt{1-c^2})R_2^2(r)\big]r^2R_3^2(r)\\
&+ \big[(1-\lambda)\epsilon_{c\theta}(1 - \sin^2\theta) + \lambda\epsilon_{s\phi}\sin^2\theta\sin^2\phi\big]\\
&\qquad \times \big[\epsilon_{c\theta} + \sin^2\theta(\epsilon_{c\phi} - \epsilon_{c\theta}) + \sin^2\theta\sin^2\phi(\epsilon_{s\phi} - \epsilon_{s\phi})\big]c^2r^4R_3^4(r),
\end{aligned}$$

where $F(r) = c^2F_1(r) + (1-c^2)F_2(r)$.

Following the method in Section 6.3.6, and in particular, recalling the identity

$$\rho_2 := \alpha_1\sin^4\theta\sin^4\phi + \alpha_2\sin^4\theta\sin^2\phi + \alpha_3\sin^4\theta + \alpha_4\sin^2\theta\sin^2\phi + \alpha_5\sin^2\theta + \alpha_6,$$

we find that, with the notation $A_1(r) = R_1^2(r)r^2R_3^2(r)$, $A_2(r) = R_2^2(r)r^2R_3^2(r)$ and $A_3(r) = r^4R_3^4(r)$, and ignoring for the moment the $F(r)$ contribution to α_6,

$$\begin{aligned}
\alpha_1 &= \big(- 1 + \epsilon_{c\phi}\epsilon_{s\phi} + \lambda c^2(1 - \epsilon_{s\phi}\epsilon_{c\phi})\big)A_3(r),\\
\alpha_2 &= \big(1 - \epsilon_{c\phi}\epsilon_{s\phi} + (1 - c^2)\epsilon_{c\theta}(\epsilon_{s\phi} - \epsilon_{c\phi}) + \lambda c^2\epsilon_{c\phi}(\epsilon_{s\phi} - \epsilon_{c\theta})\big)A_3(r),\\
\alpha_3 &= (1 - c^2 + c^2\lambda)(\epsilon_{c\theta}\epsilon_{c\phi} - 1)A_3(r),\\
\alpha_4 &= \big(2\lambda + 2c^2\lambda(\epsilon_{s\phi} - 2) + \epsilon_{c\phi} - \epsilon_{s\phi} - \lambda\epsilon_{s\phi}\big)A_1(r)\\
&\qquad + \big(c^2(\epsilon_{c\phi} - \epsilon_{s\phi}) + \lambda c^2(\epsilon_{s\phi} - 2) - 2c\sqrt{1-c^2}\lambda\epsilon_{s\phi}\big)A_2(r)
\end{aligned}$$

$$\begin{aligned}
&+ \big((1-c^2)\epsilon_{c\theta}(\epsilon_{c\phi}-\epsilon_{s\phi}) + \lambda + \lambda c^2(\epsilon_{c\theta}\epsilon_{c\phi}-2)\big)A_3(r),\\
\alpha_5 = &\big(2(\epsilon_{c\theta}-1+\lambda) - \epsilon_{c\phi} + 4c^2(1-\lambda) - \lambda\epsilon_{c\theta} + 2c^2\lambda(\epsilon_{c\theta}-2)\big)A_1(r)\\
&+ \big(c^2(2-\epsilon_{c\phi}) + \lambda c^2(\epsilon_{c\theta}-2) + 2c\sqrt{1-c^2}\lambda\epsilon_{c\theta}(1-\lambda)\big)A_2(r)\\
&+ \big(\lambda - \epsilon_{c\theta}\epsilon_{c\phi} + c^2(1-\lambda)\epsilon_{c\theta}\epsilon_{c\phi}\big)A_3(r),\\
\alpha_6 = &\big(2(1-\lambda-\epsilon_{c\theta}) + 2c^2\epsilon_{c\theta}(1-\lambda) - 4c^2(1-\lambda) + \lambda\epsilon_{c\theta}\big)A_1(r)\\
&+ \big(-2c^2(1-\lambda) - c^2\lambda\epsilon_{c\theta} - 2c\sqrt{1-c^2}\epsilon_{c\theta}(1-\lambda)\big)A_2(r)\\
&+ \big(\tfrac{1}{2} - \lambda - c^2(1-\lambda)\big)A_3(r).
\end{aligned}$$

One thing that is immediately obvious from the above is that the pair density is much more complicated than in the non-correlated cases previously discussed. For reasons given in Section 6.3.11, we wish to give an analytic proof of the global maximum. However, we need to deal with the extra correlation parameter c, which we believe will have a large effect on this maximum.

6.7.2 Explicit Critical Points

As with previous cases, we begin by finding the possible critical points of the pair density. Inserting the values of the α_i and each of the eight combinations of the ϵ_α into the critical points of Table 6.2, we obtain the following:

- $\rho_2^C = \alpha_6.$

It is clear the α_6 depends only on $\epsilon_{c\theta}$ and hence we have, for $\epsilon_{c\theta} = -1$ and $\epsilon_{c\theta} = 1$ respectively,

$$\begin{aligned}
\rho_2^{C,1,1} :&= (-3\lambda - 6c^2 + 6c^2\lambda + 4)A_1(r)\\
&+ \big(-2c^2 + 2c\sqrt{1-c^2}(1-\lambda) + 3c^2\lambda\big)A_2(r)\\
&+ (-c^2 + \tfrac{1}{2} + c^2\lambda - \lambda)A_3(r)\\
\rho_2^{C,1,2} :&= (-\lambda - 2c^2 + 2c^2\lambda)A_1(r)\\
&+ \big(-2c^2 - 2c\sqrt{1-c^2}(1-\lambda) + c^2\lambda\big)A_2(r)\\
&+ (-c^2 + \tfrac{1}{2} + c^2\lambda - \lambda)A_3(r)
\end{aligned}$$

- $\rho_2^C = \alpha_3 + \alpha_5 + \alpha_6 = -\epsilon_{c\phi}A_1(r) - c^2\epsilon_{c\phi}A_2(r) - \tfrac{1}{2}A_3(r).$

which is clearly maximized when $\epsilon_{c\phi}$ is -1 and gives

$$\rho_2^{C,2} := A_1(r) + c^2A_2(r) - \tfrac{1}{2}A_3(r)$$

- $\rho_2^C = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6.$

This expression depends only on $\epsilon_{s\phi}$ and, for $\epsilon_{s\phi} = 1$ and $\epsilon_{s\phi} = -1$ respectively,

$$\begin{aligned}
\rho_2^{C,3,1} &= (-2c^2\lambda + \lambda - 1)A_1(r) + (-c^2\lambda - c^2 - 2c\sqrt{1-c^2}\lambda)A_2(r) \\
&\quad + (-c^2\lambda - \tfrac{1}{2} + \lambda)A_3(r) \\
\rho_2^{C,3,2} &= (-6c^2\lambda + 3\lambda + 1)A_1(r) + (-3c^2\lambda + c^2 + 2c\sqrt{1-c^2}\lambda)A_2(r) \\
&\quad + (-c^2\lambda - \tfrac{1}{2} + \lambda)A_3(r).
\end{aligned}$$

It is trivial to see that

$$\rho_2^{C,3,2} - \rho_2^{C,3,1} = (-4c^2\lambda + 2\lambda + 2)A_1 + (-2c^2\lambda + 2c^2 + 4c\sqrt{1-c^2}\lambda)A_2$$

and clearly both coefficients are positive. Hence it follows that $\rho_2^{C,3} := \rho_2^{C,3,2} > \rho_2^{C,3,1}$ and we need only consider this case when finding the global maximum.

This is a complete list of critical points which are always possible (i.e. have no dependence on the relative magnitudes of the radial terms). We now move on to enumerate the remaining cases, each of which has an existence condition. For the three cases in Table 6.2 which have one non-trivial angular dependence we denote the term which must lie in $[0, 1]$ by f_i, and it must be of the form

$$f_i =: -\frac{1}{a_i A_3(r)}\big(b_i A_1(r) + c_i A_2(r) + d_i A_3(r)\big), \tag{6.77}$$

where a_i, b_i, c_i and d_i are coefficients depending only on c and λ.

The reason that the denominator of f_i depends only on $A_3(r)$ can be seen from Table 6.2 and (6.38), which show that the denominator contains only coefficients corresponding to terms which contain $\sin^4\theta$. All four orbitals in the integrals $(\psi_1\psi_2|\psi_3\psi_4)$ that contribute to these a_i must have angular dependence, and thus be p-orbitals with radial term $rR_3(r)$.

We further note from Table 6.2 that the resulting pair densities are of the form

$$\rho_i = -\frac{1}{2a_i A_3(r)}\big(b_i A_1(r) + c_i A_2(r) + d_i A_3(r)\big)^2 + g_i(r), \tag{6.78}$$

where $g_i(r)$ is a linear combination of $A_1(r), A_2(r)$,and $A_3(r)$ with coefficients in c and λ. Note that (6.78) uniquely determines the sign of α_i, even though (6.77) does not.

- $\rho_2^C = -\frac{\alpha_5^2}{4\alpha_3} + \alpha_6.$

For this case we find that there are two possible maxima, denoted $\rho_2^{C,4,1}$ (with

$(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (-1, 1, \pm 1))$ and $\rho_2^{C,4,2}$ (with $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (1, -1, \pm 1))$ with corresponding coefficients

$$\begin{aligned}
a_{4,1} = a_{4,2} = a_4 = 4\big(-1 + c^2(1-\lambda)\big), \\
b_{4,1} &= -5 + 6c^2(1-\lambda) + 3\lambda, \\
b_{4,2} &= 1 + 2c^2(1-\lambda) + \lambda, \\
c_{4,1} &= c^2(1-3\lambda) - 2c\sqrt{1-c^2}(1-\lambda), \\
c_{4,2} &= c^2(3-\lambda) + 2c\sqrt{1-c^2}(1-\lambda), \\
d_{4,1} = d_{4,2} = d_4 = 1 - c^2(1-\lambda) + \lambda,
\end{aligned} \tag{6.79}$$

and

$$\begin{aligned}
g_{4,1}(r) &= \big(4 - 6c^2(1-\lambda) - 3\lambda\big)A_1(r) + \big(c^2(3\lambda - 2) + 2c\sqrt{1-c^2}(1-\lambda)\big)A_2(r) \\
&\quad + \big(-c^2(1-\lambda) + \tfrac{1}{2} - \lambda\big)A_3(r), \\
g_{4,2}(r) &= \big(-2c^2(1-\lambda) - \lambda\big)A_1(r) + \big(c^2(-2+\lambda) - 2c\sqrt{1-c^2}(1-\lambda)\big)A_2(r) \\
&\quad + \big(-c^2(1-\lambda) + \tfrac{1}{2} - \lambda\big)A_3(r).
\end{aligned} \tag{6.80}$$

- $\rho_2^C = -\frac{(\alpha_4+\alpha_5)^2}{4(\alpha_1+\alpha_2+\alpha_3)} + \alpha_6.$

Again there are two possibilities with $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (-1, \pm 1, 1)$ and $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (1, \pm 1, -1)$ respectively, with coefficients

$$\begin{aligned}
a_{5,1} = a_{5,2} = a_5 = 4(c^2 - 1), \\
b_{5,1} &= -5 + 6c^2 + 4\lambda - 8c^2\lambda, \\
b_{5,2} &= 1 + 2c^2 - 8c^2\lambda + 4\lambda, \\
c_{5,1} &= c^2 - 4c^2\lambda - 2c\sqrt{1-c^2}, \\
c_{5,2} &= 3c^2 - 4c^2\lambda + 2c\sqrt{1-c^2}, \\
d_{5,1} = d_{5,2} = d_5 = 1 - c^2 - 2c^2\lambda + 2\lambda,
\end{aligned} \tag{6.81}$$

and

$$\begin{aligned}
g_{5,1}(r) &= \big(4 - 6c^2(1-\lambda) - 3\lambda\big)A_1(r) + \big(c^2(-2 + 3\lambda) + 2c\sqrt{1-c^2}(1-\lambda)\big)A_2(r) \\
&\quad + \big(-c^2(1-\lambda) + \tfrac{1}{2} - \lambda\big)A_3(r), \\
g_{5,2}(r) &= \big(-2c^2(1-\lambda) - \lambda\big)A_1(r) + \big(c^2(-2+\lambda) - 2c\sqrt{1-c^2}(1-\lambda)\big)A_2(r) \\
&\quad + \big(-c^2(1-\lambda) + \tfrac{1}{2} - \lambda\big)A_3(r).
\end{aligned} \tag{6.82}$$

The third and final case with at least one of $\sin^2\theta$ and $\sin^2\phi$ having a trivial form is given by

- $\rho_2^C = -\frac{(\alpha_2+\alpha_4)^2}{4\alpha_1} + \alpha_3 + \alpha_5 + \alpha_6,$

which once again leads to two possibilities with $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (\pm 1, -1, 1)$ and $(\epsilon_{c\theta}, \epsilon_{c\phi}, \epsilon_{s\phi}) = (\pm 1, 1, -1)$ respectively, with coefficients

$$\begin{aligned}
a_{6,1} = a_{6,2} = a_6 &= 4(c^2\lambda - 1),\\
b_{6,1} &= -2 + \lambda - 2c^2\lambda,\\
b_{6,2} &= 2 + 3\lambda - 6c^2\lambda,\\
c_{6,1} &= -2c^2 - c^2\lambda - 2c\sqrt{1-c^2}\lambda,\\
c_{6,2} &= 2c^2 - 3c^2\lambda + 2c\sqrt{1-c^2}\lambda,\\
d_{6,1} = d_{6,2} = d_6 &= 2 - 3c^2\lambda + \lambda,
\end{aligned} \tag{6.83}$$

and

$$\begin{aligned}
g_{6,1}(r) &= A_1(r) + c^2 A_2(r) - \tfrac{1}{2}A_3(r),\\
g_{6,2}(r) &= -A_1(r) - c^2 A_2(r) - \tfrac{1}{2}A_3(r).
\end{aligned} \tag{6.84}$$

We are left to consider the final critical point in Table 6.2, where both $\sin^2\theta$ and $\sin^2\phi$ are non-trivial. We wish to show that this critical point does not exist, i.e. that there are no values of r and λ for which both the expressions for $\sin^2\theta$ and $\sin^2\phi$ are in $[0,1]$. This is an easy but lengthy set of calculations and is given in Appendix D, which shows that, for $c \in (1/\sqrt{3}, 1]$, this critical point never exists.

6.7.3 Global Maximum

Theorem 6.7.1 (Maximum of Carbon Pair Density)**.** *Let* $R_1(r) := |1s(r)|$, $R_2(r) := |2s(r)|$, $R_3^2(r) := |p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2$ *and define* $A_0(r) := R_1^4(r) + 2R_1^2(r)R_3^2(r) + \frac{1}{2}R_3^4(r) + c^2\big(2R_1^2(r)R_2^2(r) + R_2^4(r) + 2R_2^2(r)R_3^2(r)\big)$, $A_1(r) := \big(R_1^2(r) + R_2^2(r)\big)R_3^2(r)$, *and* $A_2(r) := R_3^4(r)$. *Further denote*

$$\alpha = 4(c^2\lambda - 1), \quad \beta = -2 + \lambda - 2c^2\lambda,$$

$$\gamma = -2c^2 - c^2\lambda - 2c\sqrt{1-c^2}\lambda, \quad \delta = 2 - 3c^2\lambda + \lambda,$$

$$g(r,\lambda) = A_1(r) + c^2 A_2(r) - \tfrac{1}{2}A_3(r),$$

and

$$f(r,\lambda) = -\frac{\beta A_1(r) + \gamma A_2(r) + \delta A_3(r)}{\alpha A_3(r)}.$$

Then the global maximum of the Carbon ground state pair density is given by

$$\rho_2^{C,max}(r,\lambda) = \begin{cases} A_0(r) - \frac{(\beta A_1(r)+\gamma A_2(r)+\delta A_3(r))^2}{2\alpha A_3(r)} + g(r,\lambda) & \text{if } f(r) \in (0,1], \\ A_0(r) + g(r,\lambda) & \text{else.} \end{cases}$$

We note that the two possibilities for the global maximum correspond to $\rho_2^{C,6,1}$ and $\rho_2^{C,2}$ respectively. Hence to prove the theorem it remains to show that, when it exists, $\rho_2^{C,6,1}$ is larger than all other allowed critical points from the previous section, and if it does not exist then $\rho_2^{C,2}$ is the global maximum.

Note that the constraint in Theorem 6.7.1 that $f(r,\lambda) \notin [0,1]$ is equivalent to $f_6(r,\lambda) < 1$: It is clear from (6.83) that $a_{6,1}$, $b_{6,1}$ and $c_{6,1}$ are all negative and

$$-\frac{d_6}{a_6} = \frac{2 - 3c^2\lambda + \lambda}{4(1-c^2\lambda)} = 1 + \frac{-2 + (1+c^2)\lambda}{4(1-c^2\lambda)} < 1,$$

which, using (6.77) implies that $f_6(r,\lambda) < 1$.

In order to compare the values of the critical points we use the Lemmas of Appendix E, and it is clear that we need to determine the sign of a number of polynomials in c and λ. This is analogous to the requirements of Lemmas 6.3.6 and 6.5.2 but with an extra dependence on c.

It is possible to rigorously determine the sign of these polynomials, however, this leads to a large number of lengthy calculations. Instead we choose to plot the required function, evaluated at our value of c, using Maple and hence determine the sign. This should be very reliable since (as will be seen later) the polynomials are at most quartic in λ and on the interval $[0,1/2]$.

We begin with the case from Theorem 6.7.1 for which $\rho_2^{C,max}(r,\lambda) = A_0(r) + g(r,\lambda) = \rho_2^{C,2}$, i.e. when $f(r,\lambda) < 0$. In order to compare this to $\rho_2^{C,1,1}$, $\rho_2^{C,1,2}$ and $\rho_2^{C,3}$ we apply Lemma E.1 with $\rho_2^{(1)} = \rho_2^{C,2}$, giving

$$e_{1,1} = 1, \quad e_{1,2} = c^2, \text{ and } e_{1,3} = -\tfrac{1}{2}.$$

Further we see that $f = f(r,\lambda)$ from Theorem 6.7.1 has

$$\alpha = 4(c^2\lambda - 1), \quad \beta = -2 + \lambda - 2c^2\lambda, \quad \gamma = 2 + 3\lambda - 6c^2\lambda, \quad \delta = 2 - 3c^2\lambda + \lambda,$$

and it is therefore clear that, for $A_3(r) > 0$, $-\alpha A_3(r) > 0$ and $\delta > 0$. It remains to calculate $\tilde{e}_{i,j}$ for the three cases.

From Section 6.7.2 it is clear that, for $\rho_2^{C,1,1}$ we have

$$\begin{aligned} e_{(1,1),1} &= -3\lambda - 6c^2 + 6c^2\lambda + 4, \\ e_{(1,1),2} &= -2c^2 + 2c\sqrt{1-c^2} + 3c^2\lambda - 2c\sqrt{1-c^2}\lambda, \\ e_{(1,1),3} &= -c^2 + \tfrac{1}{2} + c^2\lambda - \lambda, \end{aligned} \tag{6.85}$$

which leads to

$$\begin{aligned} h_1 &= (2 - 12c^2 + 16c^4)\lambda^2 + 2(3 - 1c^2 - 8c^4)\lambda - 8 + 14c^2, \text{ and} \\ h_2 &= 2(-c^2 + 4c^4 + 2(1-2c^2)c\sqrt{1-c^2})\lambda^2 + 2c^2(-1 - 5c^2 + 4c\sqrt{1-c^2})\lambda \\ &\quad + 4c^2 + 4c^4 - 4c\sqrt{1-c^2}. \end{aligned}$$

The functions are shown in Figure 6.17 and are both clearly positive. Lemma E.1 then gives that $\rho_2^{C,2} > \rho_2^{C,1,1}$.

For $\rho_2^{C,1,2}$ we have

$$\begin{aligned} e_{(1,2),1} &= -\lambda - 2c^2 + 2c^2\lambda, \\ e_{(1,2),2} &= -2c^2 - 2c\sqrt{1-c^2} + c^2\lambda + 2c\sqrt{1-c^2}\lambda, \\ e_{(1,2),3} &= -c^2 + \tfrac{1}{2} + c^2\lambda - \lambda, \end{aligned} \tag{6.86}$$

from which it follows that

$$\begin{aligned} h_1 &= 2c^2(c^2 - 1)\lambda^2 + 2(3 - 5c^2 - 2c^4)\lambda + 6c^2, \text{ and} \\ h_2 &= 2c^2(c^2 + 2c\sqrt{1-c^2})\lambda^2 + 2(c^2 - 5c^4 - 4(1+c^2)c\sqrt{1-c^2})\lambda \\ &\quad + 4c^2 + 2c^4 + 4c\sqrt{1-c^2}. \end{aligned}$$

Once again, plotting these functions in Figure 6.17, it is easy to see that, for our value of c, these are both positive. Lemma E.1 therefore gives that $\rho_2^{C,2} > \rho_2^{C,1,2}$.

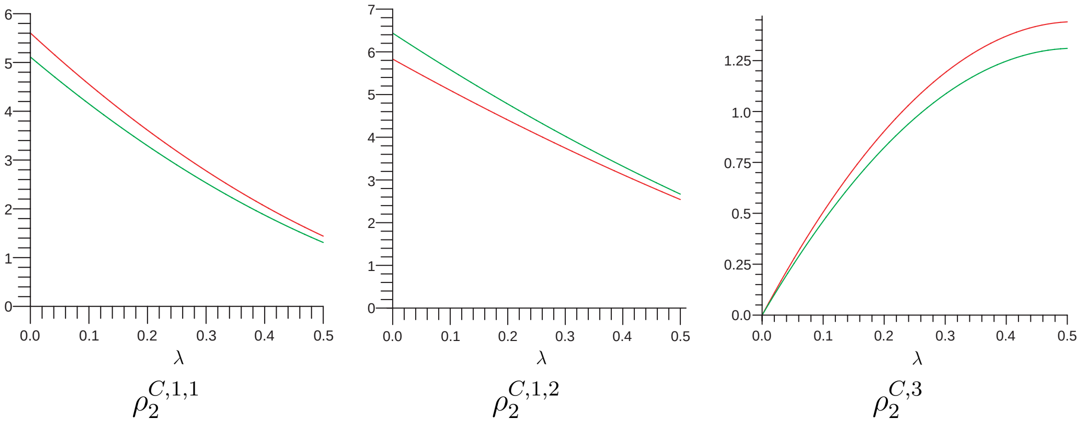
The final subcase is for $\rho_2^{C,3}$ we have

$$\begin{aligned} e_{3,1} &= -6c^2\lambda + 3\lambda + 1, \\ e_{3,2} &= -3c^2\lambda + c^2 + 2c\sqrt{1-c^2}\lambda, \\ e_{3,2} &= -c^2\lambda - \tfrac{1}{2} + \lambda. \end{aligned} \tag{6.87}$$

and hence

$$h_1 = 2(-1+6c^2-8c^4)\lambda^2 + 2(-4+7c^2)\lambda, \text{ and}$$
$$h_2 = 2(c^2+4c^4+2(2c^2-1)c\sqrt{1-c^2})\lambda^2 + 2(c^2+2c^4-2c\sqrt{1-c^2})\lambda.$$

Plotting these functions with our value of c (see Figure 6.17) shows that they are both non-negative and, by Lemma E.1, $\rho_2^{C,2} > \rho_2^{C,3}$. In fact for this case we need only determine the sign of a linear function but, for consistency, we plot the quadratic.



$\rho_2^{C,1,1}$ $\rho_2^{C,1,2}$ $\rho_2^{C,3}$

Figure 6.17: Plots of h_1 (red) and h_2 (green) from Lemma E.1 for $\rho_2^{(1)} = \rho_2^{C,2}$, $\rho_2^{(2)}$ the indicated critical point, and $c = 0.98576$. It is clear that all functions are non-negative.

We now move on to compare $\rho_2^{C,2}$ when $f(r,\lambda) < 0$ with $\rho_2^{C,4,i}$, $\rho_2^{C,5,i}$, $i = 1,2$ and $\rho_2^{C,6,2}$ for which we use Lemma E.2. We set $\rho_2^{(2)} = \rho_2^{C,2}$ and $f_3 = f(r,\lambda)$, whilst $\rho_2^{(2)}$ and f_2 are the pair density and angular condition to be tested. We have already shown that, for $A_3(r) > 0$, $-\alpha_3 A_3(r) > 0$ and $\delta_3 > 0$.

The coefficients α_i, β_i, etc are given explicitly in (6.79), (6.81), (6.83), and the g_{ij} can be read off from (6.80), (6.82) and (6.84). In particular it is trivial to see that $-\alpha_2 A_3(r) > 0$ in all cases.

It is clear that all of these coefficients are at most linear in λ and hence coefficients in the upper bound on $\rho_2^{(2)}$ in the proof of Lemma E.2 also are at most linear in λ. In particular we have the explicit forms for $2e_{1,3} - (\delta_2 + g_{2,3})$ given by $-3(1-c^2) - 3\lambda c^2 + \lambda$, $3(c^2-1)$, and $-2 - \lambda + 3\lambda c^2$ for $\rho_2^{C,4,i}$, $\rho_2^{C,5,i}$, $i = 1,2$, and $\rho_2^{C,6,2}$ respectively, and all three are negative.

It remains to consider the coefficients h_1 and h_2 from Lemma E.2, which are given by

$\rho_2^{C,4,1}$:

$$\begin{aligned}h_1 &= (1-5c^2+6c^4)\lambda^2 + (5-9c^2-6c^4)\lambda - 4 + 9c^2, \text{ and}\\ h_2 &= (-c^2+3c^4+2(1-3c^2)c\sqrt{1-c^2})\lambda^2\\ &\quad + (-c^2-9c^4+2(3c^2-1)c\sqrt{1-c^2})\lambda + 2c^2 + 3c^4 - 2c\sqrt{1-c^2}.\end{aligned}$$

$\rho_2^{C,4,2}$:

$$\begin{aligned}h_1 &= 4(2-5c^2)\lambda - 4 + 10c^2, \text{ and}\\ h_2 &= 4(-3c^4-2c\sqrt{1-c^2})\lambda + 6c^4 + 4c\sqrt{1-c^2}.\end{aligned}$$

$\rho_2^{C,5,1}$:

$$\begin{aligned}h_1 &= 2(1-5c^2+6c^4)\lambda^2 + 2(3-4c^2-6c^4)\lambda - 8 + 18c^2, \text{ and}\\ h_2 &= 2(-c^2+3c^4+2(1-3c^2)c\sqrt{1-c^2})\lambda^2\\ &\quad + 2(-c^2-6c^4+2(3c^2-1)c\sqrt{1-c^2})\lambda + 4c^2 + 6c^4 - 4c\sqrt{1-c^2}.\end{aligned}$$

$\rho_2^{C,5,2}$:

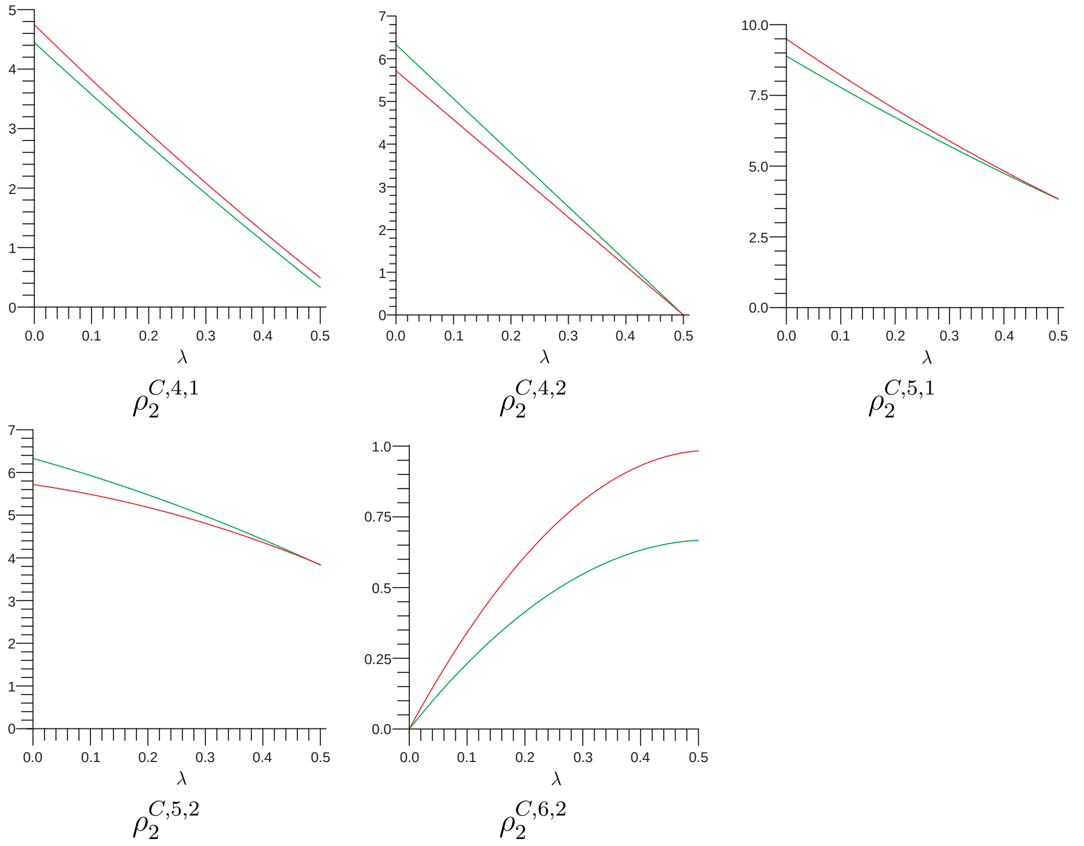
$$\begin{aligned}h_1 &= 2(-1+5c^2-6c^4)\lambda^2 - 2c^2\lambda - 4 + 10c^2, \text{ and}\\ h_2 &= 2(c^2-3c^4-2(1-3c^2)c\sqrt{1-c^2})\lambda^2\\ &\quad + 2(2c^2-3c^4-12c\sqrt{1-c^2})\lambda + 6c^4 + 4c\sqrt{1-c^2}.\end{aligned}$$

$\rho_2^{C,6,2}$:

$$\begin{aligned}h_1 &= -2(3c^2-1)(2c^2-1)\lambda^2 + 4(2c^2-1)\lambda, \text{ and}\\ h_2 &= -2(3c^2-1)(c^2-2c\sqrt{1-c^2})\lambda^2 + 4(c^2-2c\sqrt{1-c^2})\lambda.\end{aligned}$$

The cases of $\rho_2^{C,4,2}$ and $\rho_2^{C,6,2}$ are clearly equivalent to finding the sign of functions linear in λ and it is therefore trivial to show that these h_i are all non-negative on $\lambda \in [0,1/2]$ for our value of c. The remaining cases are quadratics in λ, which follows from the fact that all coefficients are at most linear in λ and the form of the h_i. These are plotted in Figure 6.18 and are clearly all non-negative.

It follows from Lemma E.2 that, when $\rho_2^{C,6,1}$ does not exist, $\rho_2^{C,2} > \rho_2^{C,i,j}$ where $(i,j) = (4,1), (4,2), (5,1), (5,2)$ and $(6,2)$. This concludes the analysis of the case when $\rho_2^{C,max} = \rho_2^{C,2}$ and shows that this is indeed the global maximum of the Carbon pair density.



$\rho_2^{C,4,1}$ $\rho_2^{C,4,2}$ $\rho_2^{C,5,1}$

$\rho_2^{C,5,2}$ $\rho_2^{C,6,2}$

Figure 6.18: Plots of h_1 (red) and h_2 (green) from Lemma E.2 for $\rho_2^{(1})=\rho_2^{C,2}$, $\rho_2^{(2)}$ the indicated critical point, f_2 its existence condition, $f_3=f(r,\lambda)$ from Theorem 6.7.1, and $c=0.98576$. It is clear that all functions are non-negative.

We must now consider the case from Theorem 6.7.1 where $\rho_2^{C,max}=A_0(r)+\rho_2^{C,6,1}$, for which we must have $f(r,\lambda)\in[0,1]$. Using Lemma E.3 we compare this to the four critical points which are always present. However, the fact that $\rho_2^{C,6,1}\geq\rho_2^{C,2}$ whenever it exists is trivial since in this case we have $-2\alpha A_3(r)>0$ and $g_i=e_i$.

For the cases of $\rho_2^{C,1,1}$ and $\rho_2^{C,1,2}$ we use Lemma E.3. We have already shown that $-\alpha A_3(r)>0$ whenever $A_3(r)>0$ and hence it remains to consider the $\tilde{f}_{ij}$.

For $\rho_2^{C,1,1}$ we have

$$\begin{aligned}
f_{11}&=(2+2c^2\lambda-\lambda)^2,\\
f_{22}&=c^2(4-3c^2+4c\sqrt{1-c^2})\lambda^2+4c^2(c^2+8c\sqrt{1-c^2})\lambda+4c^2,\\
f_{12}&=2(2+2c^2\lambda-\lambda)((2+\lambda)c^2+2\lambda c\sqrt{1-c^2}),\\
f_{33}&=(1-14c^2+17c^4)\lambda^2+4(3-3c^2-8c^4)\lambda+8c^2-4,
\end{aligned}$$

$$\begin{aligned}
f_{13} &= 2(1 - 17c^2 + 30c^4)\lambda^2 + 4(6 - 5c^2 - 12c^4)\lambda - 32 + 48c^2, \\
f_{23} &= 2(-c^2 + 15c^4 - 2(c^2+1)c\sqrt{1-c^2})\lambda^2 \\
&\quad + 4(-8c^2 - 3c^4 + 2(2c^2+1)c\sqrt{1-c^2})\lambda + 16(c^2 - c\sqrt{1-c^2}).
\end{aligned}$$

It is clear that all of f_{11}, f_{22} and f_{12} are positive, and since it is a quadratic in λ, it is easy to show that, for $c > 1/\sqrt{2}$, f_{33} is also positive.

Thus, using the notation of Lemma E.3, it follows that

$$\tilde{f}_{11} = f_{11} + \tfrac{1}{2}\chi_{13} f_{13}, \quad \tilde{f}_{22} = f_{22} + \tfrac{1}{2}\chi_{23} f_{23}, \quad \tilde{f}_{12} = f_{12},$$

$$\tilde{f}_{33} = f_{33} + \tfrac{1}{2}(\chi_{13} f_{13} + \chi_{23} f_{23}), \quad \tilde{f}_{13} = (1 - \chi_{13}) f_{13}, \text{ and } \tilde{f}_{23} = (1 - \chi_{23}) f_{23}.$$

Figure 6.19 show that, for our value of c, each of the $\tilde{f}_{ij}$ is non-negative on $\lambda \in [0, 1/2]$ and by Lemma E.3 we have $\rho_2^{C,6,1} > \rho_2^{C,1,1}$ It is worth noting that, for our value of c, only f_{23} is ever negative.

For $\rho_2^{C,1,2}$ we note that f_{11}, f_{22} and f_{12} depend only on the coefficients of $\rho_2^{C,6,1}$ and hence are the same as for $\rho_2^{C,1,1}$. Furthermore, by (6.85) and (6.86) we see that $e_{(1,1),3} = e_{(1,2),3}$ and hence f_{33} is also unchanged. The new terms are given by

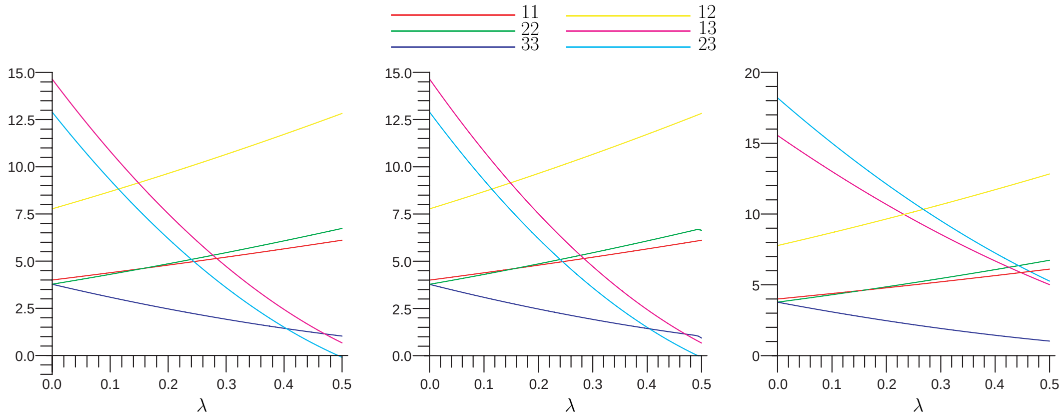
$$\begin{aligned}
f_{13} &= 2(1 - 9c^2 + 14c^4)\lambda^2 + 4(2 - 5c^2 - 4c^4)\lambda + 16c^2, \\
f_{23} &= 2(-c^2 + 7c^4 + 2(7c^2 - 1)c\sqrt{1-c^2})\lambda^2 \\
&\quad + 4(-4c^2 - 3c^4 - 2(2c^2+6)c\sqrt{1-c^2})\lambda + 16(c^2 + c\sqrt{1-c^2}).
\end{aligned}$$

Inserting our value of c results in all six coefficients being positive for all λ (see Figure 6.19) and thus we have $\tilde{f}_{ij} = f_{ij}$ for all i and j and Lemma E.3 gives that $\rho_2^{C,6,1} > \rho_2^{C,1,2}$.

For the case of $\rho_2^{C,3}$, we must use Lemma E.4 with $f = 8(1 - \lambda c^2) A_3(r)(\rho_2^{C,6,1} - \rho_2^{C,3})$, which has the same sign as $\rho_2^{C,6,1} - \rho_2^{C,3}$. This gives

$$\begin{aligned}
e_{11} &= (2\lambda c^2 + 2 - \lambda)^2, \\
e_{22} &= (4 + 4\lambda - 3\lambda^2)c^2 + 4\lambda c^3\sqrt{1-c^2}(\lambda + 2) + 4\lambda c^2, \\
e_{12} &= 4\lambda(2+\lambda)c^4 + 8\lambda^2 c^3\sqrt{1-c^2} + 2(4 - \lambda^2)c^2 + 4\lambda(2-\lambda)c\sqrt{1-c^2}, \\
e_{33} &= 4(1 - \lambda(1 + c^2)) + \lambda^2(c^4 + 2c^2 + 1).
\end{aligned}$$

It is clear that, for $c \in [0, 1)$, all of these are strictly positive, satisfying the



f_{ij} for $\rho_2^{C,1,1}$ $\tilde{f}_{ij}$ for $\rho_2^{C,1,1}$ $f_{ij} = \tilde{f}_{ij}$ for $\rho_2^{C,1,2}$

Figure 6.19: Plots of f_{ij} and $\tilde{f}_{ij}$ from Lemma E.3 for $\rho_2^{(1)} = \rho_2^{C,6,1}$, $\rho_2^{(2)}$ the indicated critical point, and $c = 0.98576$.

restrictions in Lemma E.4. In addition it is simple to check that the identity $4e_{11}e_{22} - e_{12}^2 = 0$ holds. The remaining coefficients are given by

$$\begin{aligned} e_{13} &= 2(1 + 7c^2 - 18c^4)\lambda^2 + 4(13c^2 - 6)\lambda - 8, \\ e_{23} &= 2(-c^2(1 + 9c^2) + (14c^2 - 2)c\sqrt{1 - c^2})\lambda^2 \\ &\quad + 4(c^2(4 + 3c^2) - 14c\sqrt{1 - c^2})\lambda - 8c^2, \end{aligned}$$

from which it is trivial to compute $r_1 = -\frac{1}{2}e_{12}e_{13} + e_{11}e_{23}$ and $r_2 = -\frac{1}{4}e_{13}^2 + e_{11}e_{33}$ and the two test functions $q_1 := -r_1e_{13} + r_2e_{12}$ and $q_2 := -e_{23}^2 + 4e_{22}e_{33}$.

We find that

$$\begin{aligned} q_1 &= 32c\lambda(2 + 2c^2\lambda - \lambda)(1 - \lambda c^2)Q_1, \\ q_2 &= 64c^2\lambda(1 - \lambda c^2)Q_2, \end{aligned}$$

where Q_1 and Q_2 are explicit polynomials in λ of degree 3 and 2 respectively, with coefficients depending on c. It is clear that the prefactors in q_1 and q_2 are positive and hence we need only determine the signs of r_1 (which must be non-positive), Q_1 and Q_2, which must both be non-negative.

The functions r_1, Q_1 and Q_2 for our value of c are shown in Figure 6.20 and it is clear that $r_1 \leq 0$, $Q_1 \geq 0$ and $Q_2 \geq 0$. Hence Lemma E.4 allows us to conclude that $\rho_2^{C,6,1} \geq \rho_2^{C,3}$ and equality holds only when $\lambda = 0$, and for this case it is trivial to directly check that $\rho_2^{C,6,1} > \rho_2^{C,3}$.

The final case to consider is when both critical points have an existence con-

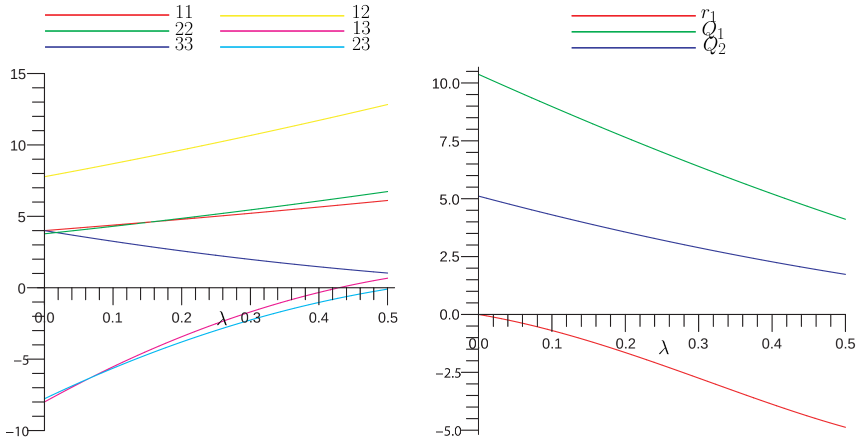


Figure 6.20: Plots of e_{ij} from Lemma E.4 for $f = 8(1-\lambda^2)A_3(r)(\rho_2^{C,6,1} - \rho_2^{C,3})$, along with the resulting r_1, Q_1 and Q_2, all evaluated at $c = 0.98576$.

dition. Recall that if, for each $\lambda \in [0, 1/2]$, at least one of Lemmas E.5 and E.6 apply then $\rho_2^{(1)} > \rho_2^{(2)}$ everywhere. As already stated, we have that $-\alpha_1 A_3(r) > 0$ and $\alpha_2 A_3(r) > 0$ whenever $A_3(r) > 0$ in all cases. We also have that $\delta_1 > 0$ and, by the assumption that $\rho_2^{C,max} = \rho_2^{C,6,1}$, that $f_1 \geq 0$. Further, to apply Lemma E.6, we have that $f_2 \leq 1$ (since both critical points exist), but we must check that in each case we have $-(\alpha_2 + \delta_2) > 0$. It then remains to show that e_{11}, e_{12}, e_{22}, h_{11}, h_{12} and h_{22} are all non-negative.

All the necessary coefficients are given in (6.79)-(6.84) and are at most linear in λ. It follows that e_{ij} are at most cubic and h_{ij} at most quartic in λ.

For the case with $\rho_2^1 = \rho_2^{C,6,1}$ and $\rho_2^{C,6,2}$ we see that there is a quicker argument than Lemmas E.5 and E.6. Since $\alpha_1 = \alpha_2$ we construct $2\alpha_1 A_3(r)(\rho_2^{C,6,1} - \rho_2^{C,6,2})$ which is, as in the proof of Lemma E.5, of the form

$$\begin{aligned} 2\alpha_1 A_3(r)(\rho_2^{(1)} - \rho_2^{(2)}) &= e_{11}A_1^2(r) + e_{22}A_2^2(r) + e_{33}A_3^2(r) \\ &\quad + e_{12}A_1(r)A_2(r) + e_{13}A_1(r)A_3(r) + e_{23}A_2(r)A_3(r), \end{aligned}$$

where we find

$$\begin{aligned} e_{11} &= \lambda(8(-1+4c^2-4c^4)\lambda + 16(2c^2-1)), \\ e_{22} &= \lambda(8c^2(-c^2+2c\sqrt{1-c^2})\lambda + 16c^4), \\ e_{33} &= 0, \\ e_{12} &= \lambda(16(c^2-2c^4+(2c^2-1)c\sqrt{1-c^2})\lambda + 32c^4), \\ e_{13} &= \lambda(4(-1+5c^2-6c^4)\lambda + 8(3c^2-3)), \\ e_{23} &= \lambda(4(c^2-3c^4+2(3c^2-1)c\sqrt{1-c^2})\lambda + 8(c^4-2c\sqrt{1-c^2}). \end{aligned}$$

It is therefore clear that we need only determine the sign of a function linear in λ and see that, for our value of c, these are all clearly positive, as shown in Figure 6.21.

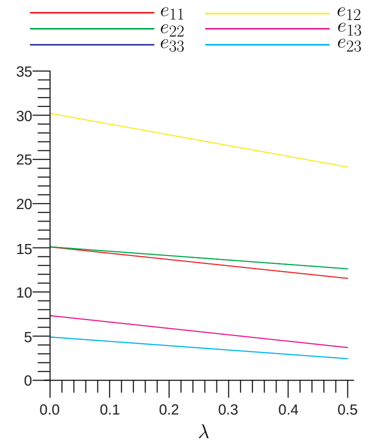


Figure 6.21: Plots of e_{ij} from Lemma E.5 for $\rho_2^{(1)} = \rho_2^{C,6,1}$ and $\rho_2^{(2)} = \rho_2^{C,6,2}$. All functions are evaluated at $c = 0.98576$.

For the cases of $\rho_2^{C,4,1}$ and $\rho_2^{C,4,2}$, we use Lemma E.6 and have $-(\alpha_2 + \delta_2) = 3(1 - c^2) + \lambda(3c^2 - 1)$, which is clearly positive for $c \in [1/\sqrt{3}, 1]$ and hence our value of c.

We first check that in each case e_{13}, e_{23} and e_{33} are non-negative, which can be seen from Figure 6.22. We also show e_{11}, e_{12} and e_{22} to demonstrate that these are not always non-negative, and hence these cases are non-trivial. Furthermore, also from Figure 6.22, we see that h_{11}, h_{12} and h_{13} are non-negative and so Lemma E.6 shows that $\rho_2^{C,6,2} > \rho_2^{C,4,i}$, $i = 1, 2$ whenever both exist.

The case of $\rho_2^{C,5,1}$ is less straightforward and requires the use of both Lemmas E.5 and E.6. We see that $-(\alpha_2 + \delta_2) = (3 - 2\lambda)(1 - c^2) \geq 0$ and so it remains to check the conditions on the e_{ij} and h_{ij}. It is clear that that the e_{ij} are the same in both lemmas and, as shown in Figure 6.23, these are positive. Once again the other e_{ij} are not always non-negative and thus this case is non-trivial.

Figure 6.23 shows that the resulting h_{ij} from Lemma E.5 are all positive for $\lambda \in [0.1, 0.35]$ whilst those from Lemma E.6 are all positive for $\lambda \in [0, 0.1]$ and $\lambda \in [0.35.1/2]$. From this it follows that, for each λ, at least one of Lemmas E.6 and E.5 apply and thus, when they both exist, $\rho_2^{C,6,1} > \rho_2^{C,5,1}$.

For the final case of $\rho_2^{C,5,2}$ we again see (Figure 6.24) that e_{13}, e_{23} and e_{33} are positive and in this case the other e_{ij} are negative so this case is also non-trivial.

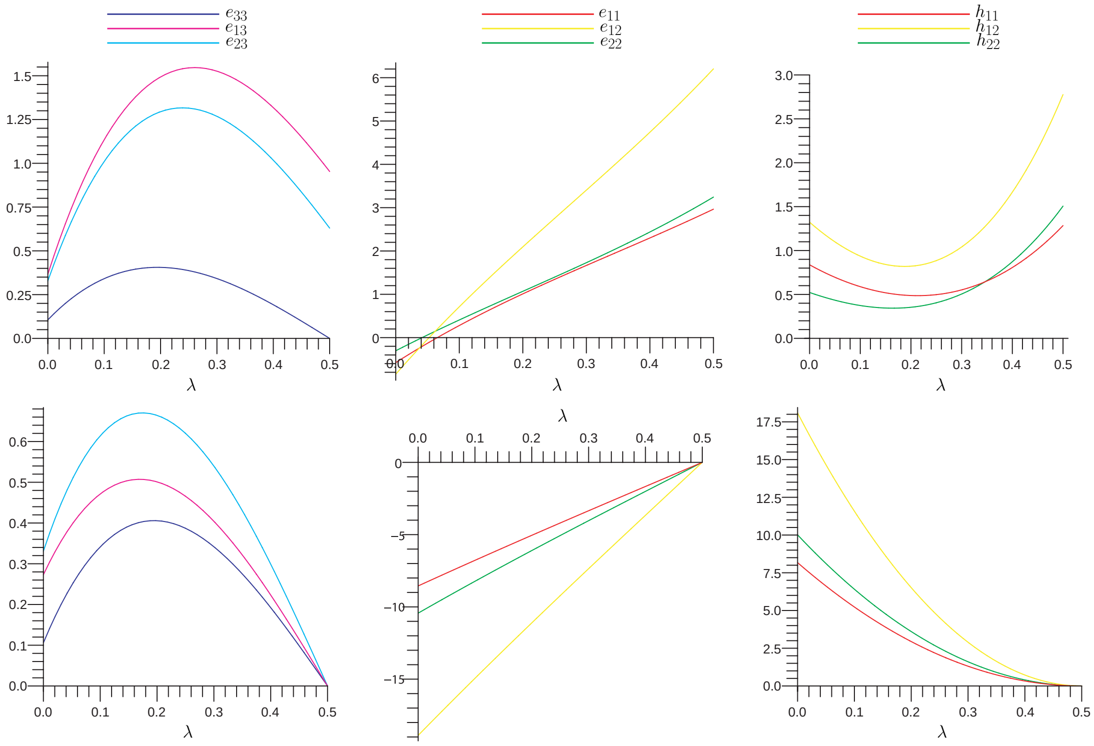


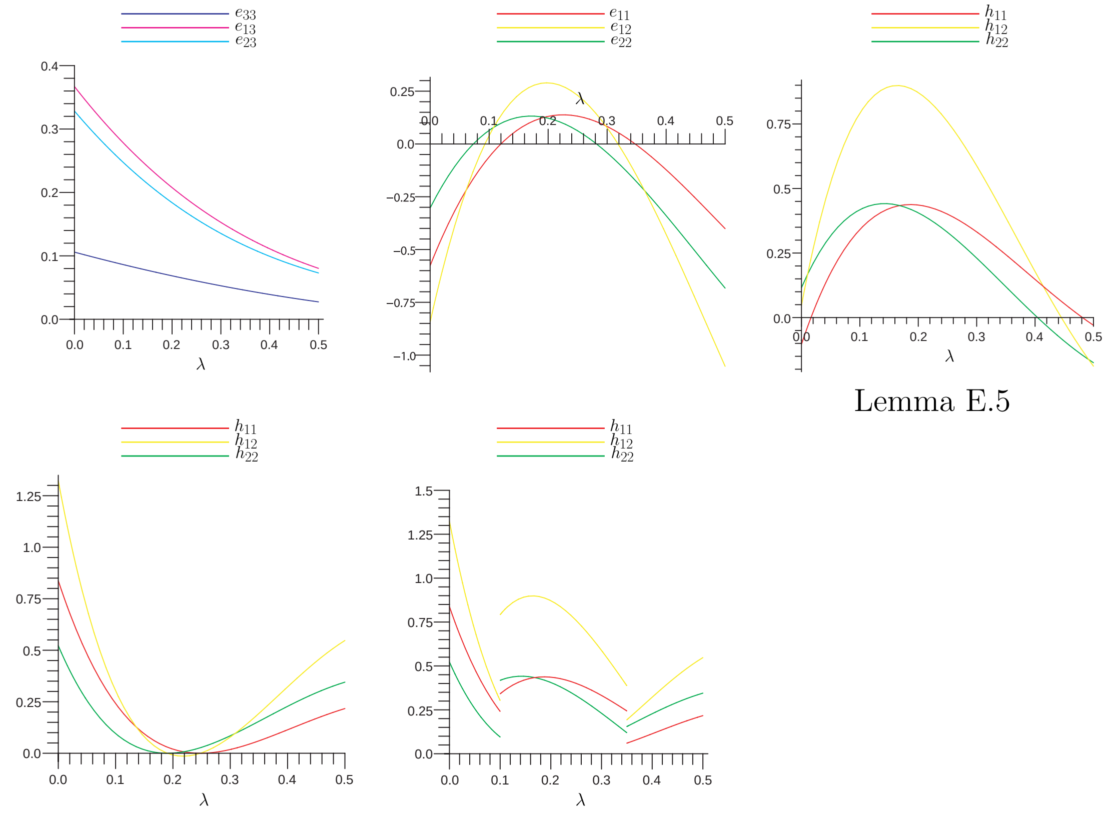
Figure 6.22: Plots of e_{ij} from Lemma E.6 for $\rho_2^{(1)} = \rho_2^{C,6,1}$ and $\rho_2^{(2)} = \rho_2^{C,4,1}$ for the top line and $\rho_2^{C,4,2}$ for the bottom line. Also shown are the with the resulting h_{11}, h_{12} and h_{22}. All functions are evaluated at $c = 0.98576$.

Applying Lemma E.6 we have the same $-(\alpha_2+\delta_2) \geq 0$ as for $\rho_2^{C,5,2}$ and Figure 6.24 shows that the h_{ij} are once again positive. Thus, when both exist, $\rho_2^{C,6,1} > \rho_2^{C,5,2}$, which completes the proof of Theorem 6.7.1.

6.7.4 Bond Angle

It remains to determine the angle of the maximizing pair density. We begin by noting that $\rho_2^{C,2}$ corresponds to $\cos\theta = 0$, $\sin\phi = 0$ and $\epsilon_{c\theta} = \pm 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = \pm 1$ and hence has an angle of 180°. Now, $\rho_2^{C,6,1}$ corresponds to $\sin^2\phi = f_6$ with $\cos\theta = 0$ and $\epsilon_{c\theta} = \pm 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = 1$. Hence we may set $\theta_1 = \theta_2 = \frac{\pi}{2}$ and $\phi_1 = \pi - \phi_2$, which gives an angle of $\Phi = \pi - 2\phi_1$, with $\phi_1 = \arcsin((\max\{0, f(r,\lambda)\})^{1/2})$, where $f(r,\lambda)$ is given in Theorem 6.7.1.

As in the previous cases we plot the bond angle range for Carbon. In this case the monotonicity is less trivial. Differentiating the expression for $f(r,\lambda)$ in



Lemma E.5

Lemma E.6

Combined

Figure 6.23: Plots of e_{ij} for Lemmas E.5 and E.6 for $\rho_2^{(1)} = \rho_2^{C,6,1}$ and $\rho_2^{(2)} = \rho_2^{C,5,1}$, along with the resulting h_{11}, h_{12} and h_{22} and the combined h_{ij}'s which are always positive. All functions are evaluated at $c = 0.98576$.

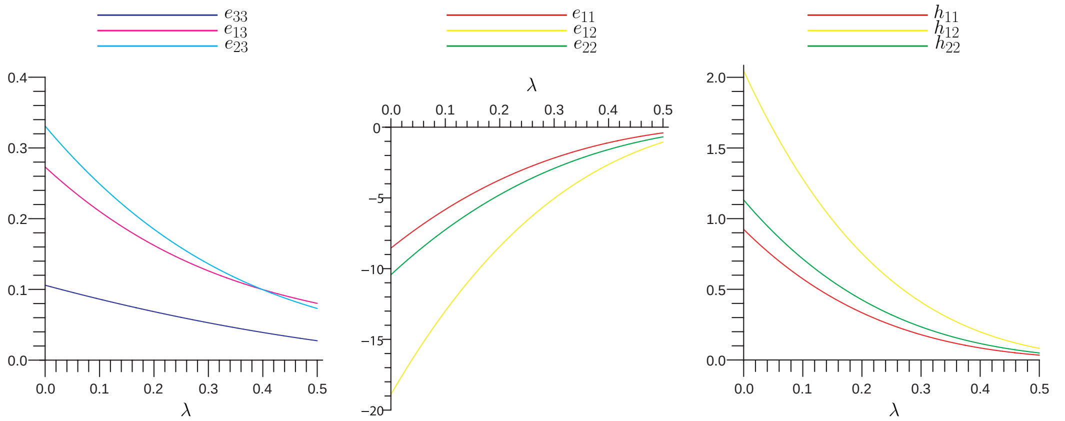


Figure 6.24: Plots of e_{ij} from Lemma E.6 for $\rho_2^{(1)} = \rho_2^{C,6,1}$ and $\rho_2^{(2)} = \rho_2^{C,5,2}$. Also shown are the with the resulting h_{11}, h_{12} and h_{22}. All functions are evaluated at $c = 0.98576$.

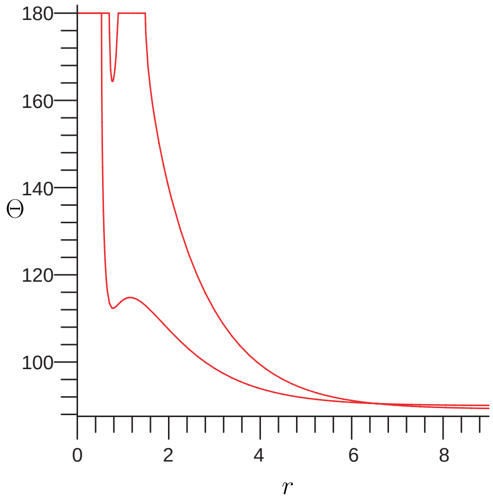


Figure 6.25: Bond angle range for Carbon. The area between the two curves gives the range in which the angle may vary with no energy penalty.

Theorem 6.7.1 and multiplying the result by a positive term, we have

$$4(c^2\lambda-1)^2 A_3(r)\frac{d}{d\lambda}f(r,\lambda) = (1-4c^2)A_1(r) + (-c^2-2c^4-2c\sqrt{1-c^2})A_2(r) + (1-c^2)A_3(r).$$

Hence the derivative is independent of λ and thus the bond angle is always monotone (but unlike in previous cases, not uniformly increasing or decreasing for each value of r).

For small values of r we expect $A_1(r)$ and $A_2(r)$ to dominate $A_3(r)$ and hence the derivative will be negative. A similar argument to the previous cases shows that the bond angle will be increasing with λ. For large enough r this will reverse so that the bond angle is decreasing with λ. The resulting angle ranges are shown in Figure 6.25.

As in the previous cases we may use this monotonicity and the fact that $A_2(r)$ decays exponentially slower than $A_1(r)$ to give the range in the limit $r \to \infty$. We see that the limiting range of $f(r,\lambda)$ is $-\delta/\alpha = -(2-3c^2\lambda+\lambda)/(4(c^2\lambda-1)$. For $\lambda = 0$ this gives a value of $1/2$ and for $\lambda = 1/2$ it gives a value of $(5-3c^2)/(8-4c^2)$, leading to a range $\Phi \in [\pi - 2\arcsin(((5-3c^2)/(8-4c^2))^{1/2}), 90°] \sim [89.2°, 90°]$.

Inserting $r = 1.254$, the position of the outer maximum of the single particle density from Table 5.2, we find a bond angle range of $[114.6°, 180°]$. From Table 6.1 we see that the experimental bond angle for CH_2 is 136° and is contained in the predicted bond angle range.

Figures 6.26-6.29 show the densities and pair densities when one electron is

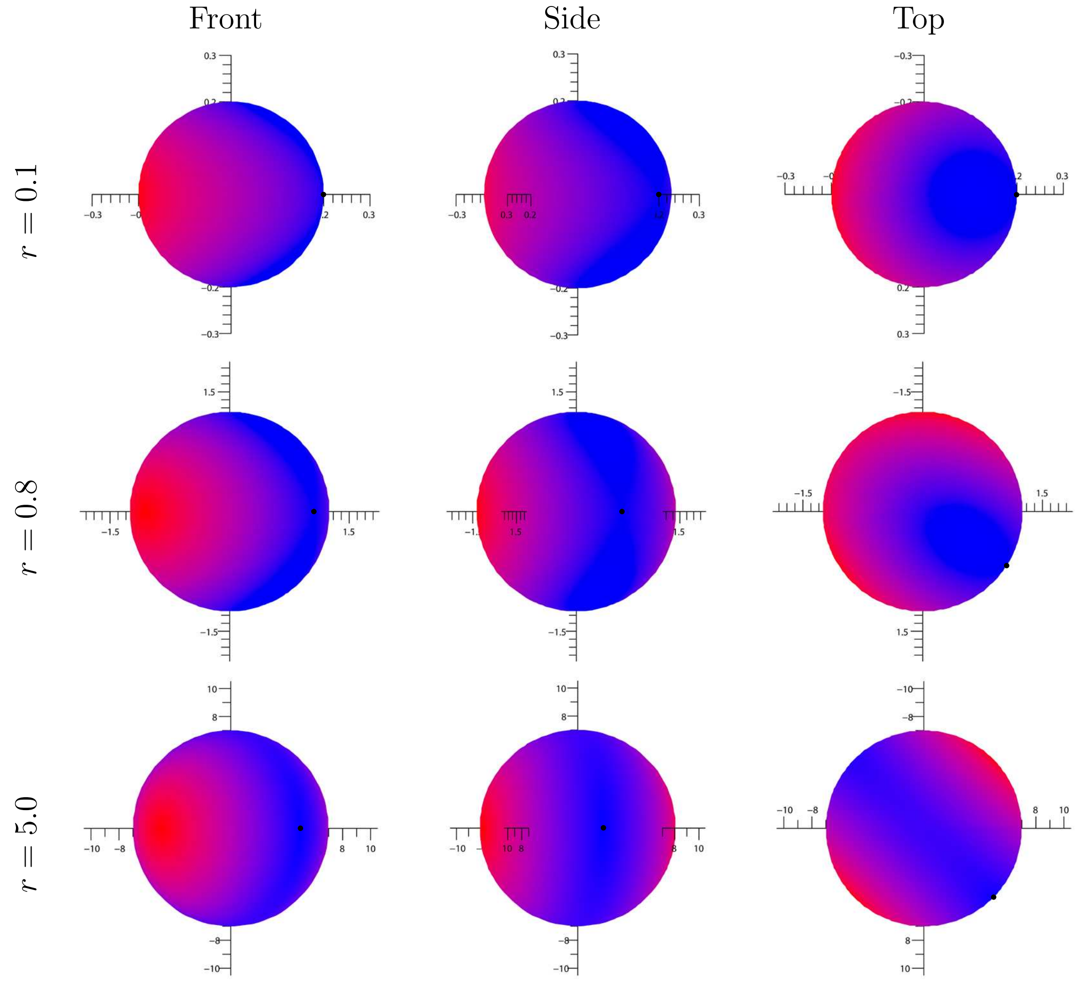


Figure 6.26: Carbon pair densities for various distances with $\lambda = 0$. Front denotes the view along the positive e_1 axis, Top the view along the positive e_3 axis and Side the view with the e_1 axis out of the plane and the e_3 axis vertical. The black dot denotes the maximizing position of one of the electrons.

placed in the maximizing position. The plots are very similar to those of Fluorine in Figures 6.13-6.16, which is due to the similarity in the wavefunctions when $c \sim 1$.

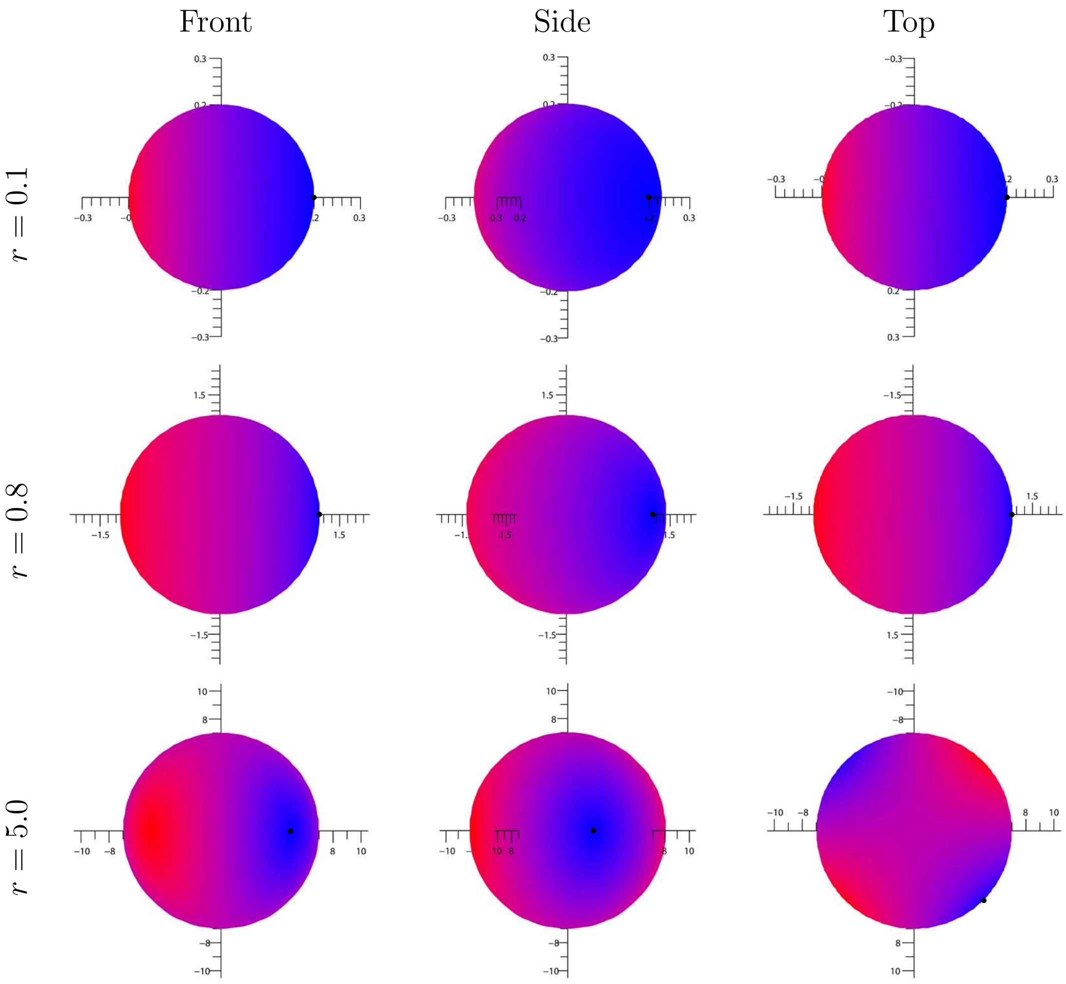


Figure 6.27: Carbon pair densities for various distances with $\lambda = 1/2$, orientations and description as in Figure 6.26.

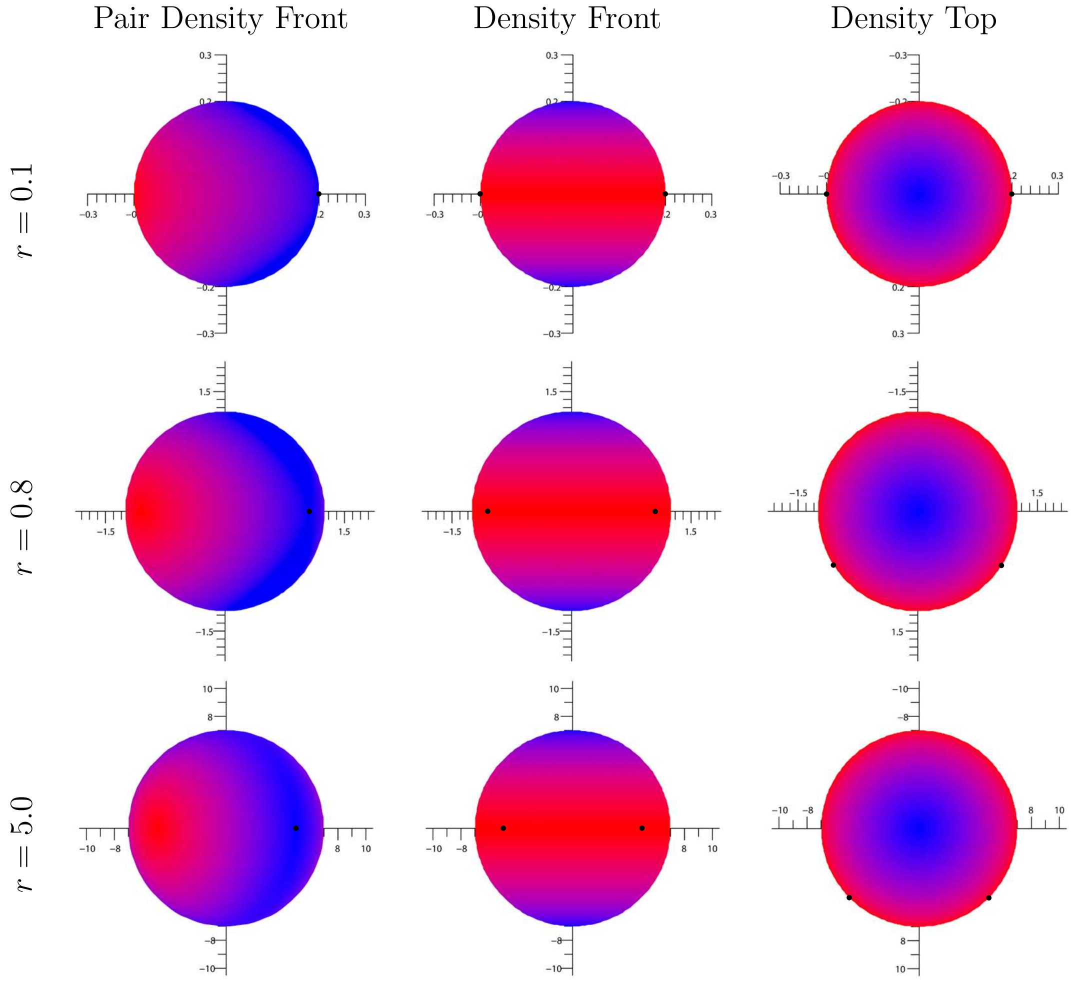
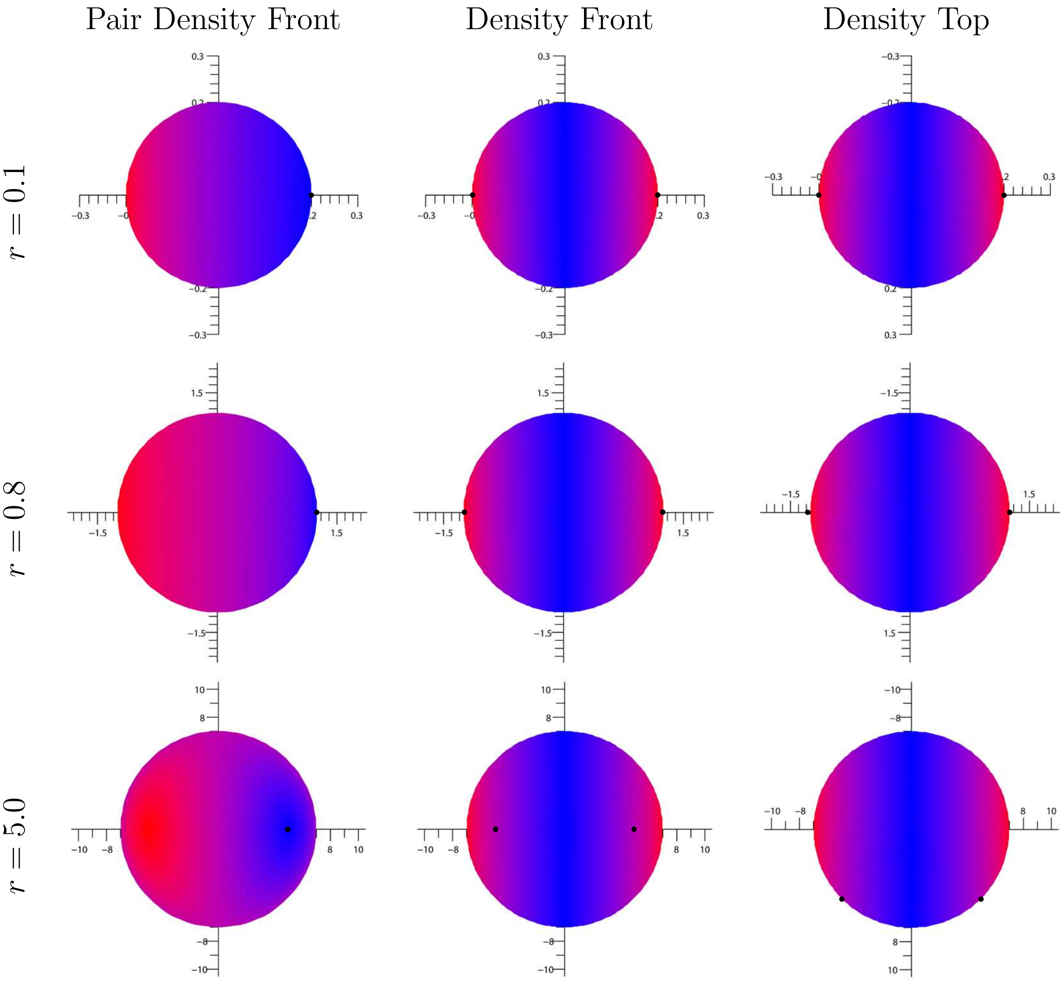


Figure 6.28: Carbon pair density and single particle density with $\lambda = 0$, orientations as in Figure 6.26. The black dots denote the positions of the electrons in the maximum pair density.



Pair Density Front

Density Front

Density Top

$r = 0.1$

$r = 0.8$

$r = 5.0$

Figure 6.29: Carbon pair density and single particle density with $\lambda = 1/2$, details as in Figure 6.28.

6.8 Boron

We wish to find a general form for the Boron ground state wavefunction and by Lemma 6.2.1 we may choose the case $S_3\Psi = \Psi$. The form used to calculate the energy was

$$\Psi = c|1s\overline{1s}2s\overline{2s}p_3\rangle - \sqrt{1-c^2}\big(|1s\overline{1s}p_1\overline{p_1}p_3\rangle + |1s\overline{1s}p_2\overline{p_2}p_3\rangle\big),$$

which is an eigenfunction of L_3 with eigenvalue zero. As shown by Lemma 6.6.1, we may take $\Psi = c\Psi_1 - \sqrt{1-c^2}\Psi_2$ with $c \in \mathbb{R}$. By an almost identical argument as for Oxygen in Section 6.3.1 we know that we may write a general form of the first Slater determinant as

$$\Psi_1 := \alpha_1|1s\overline{1s}2s\overline{2s}p_1\rangle + \alpha_2|1s\overline{1s}2s\overline{2s}p_2\rangle + \alpha_3|1s\overline{1s}2s\overline{2s}p_3\rangle$$

where $\alpha_i \in \mathbb{C}$, $\sum |\alpha_i|^2 = 1$, which has a pair density of the form

$$\begin{aligned}\rho_2^{\Psi_1}(x,y) = {} & \rho_2^{|1s\overline{1s}2s\overline{2s}\rangle}(x,y) \\ & + \tfrac{1}{2}\big[\rho^{|1s\overline{1s}2s\overline{2s}\rangle}(y)k^2e^{-Z|x|}\big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)x\big) + x \leftrightarrow y\big] \\ & - \big(1s(x)1s(y) + 2s(x)2s(y)\big)k^2e^{-Z|x|/2}e^{-Z|y|/2} \\ & \qquad\qquad \times \big(x\cdot(a_1\otimes a_1 + b_1\otimes b_1)y\big),\end{aligned} \tag{6.88}$$

where $k((a_1+ib_1)\cdot x)e^{-Z|x|/2} := \alpha_1p_1(x) + \alpha_2p_2(x) + \alpha_3p_3(x)$. We also note that by Lemma 6.6.1 we may choose c to be real.

We may also write the general form of the second part of the wavefunction as

$$\begin{aligned}\Psi_2 := {} & \tfrac{\beta_1}{\sqrt{2}}\big(|1s\overline{1s}p_2\overline{p_2}p_1\rangle + |1s\overline{1s}p_3\overline{p_3}p_1\rangle\big) + \tfrac{\beta_2}{\sqrt{2}}\big(|1s\overline{1s}p_1\overline{p_1}p_2\rangle + |1s\overline{1s}p_3\overline{p_3}p_2\rangle\big) \\ & + \tfrac{\beta_3}{\sqrt{2}}\big(|1s\overline{1s}p_1\overline{p_1}p_3\rangle + |1s\overline{1s}p_2\overline{p_2}p_3\rangle\big)\end{aligned}$$

where $\beta_i \in \mathbb{C}$, $\sum |\beta_i|^2 = 1$. We now wish to use the restriction that $\Psi(\alpha,\beta)$ is a minimizer of the energy to derive restrictions on the β_i in terms of the α_i. Firstly we see that $\langle\Psi_2|V_{ee}|\Psi_2\rangle$ is independent of the choice of β_i, which follows, as in the Carbon case, from the symmetry of the p-orbitals.

We now choose a suitable simultaneous rotation of all coordinates such that Ψ_1 obtains the canonical form

$$\Psi_1 = \sqrt{1-\lambda}|1s\overline{1s}2s\overline{2s}p_1\rangle + i\sqrt{\lambda}|1s\overline{1s}2s\overline{2s}p_2\rangle$$

which is clearly possible from the form of the pair density of Ψ_1 and arguments analogous to that in the Oxygen case. We wish to determine the energy for this choice of Ψ_1 and a general choice of β_i and note that the energy only differs in the term $\langle\Psi_1|V_{ee}|\Psi_2\rangle$ which, removing terms differing by 3 orbitals, is given by

$$\begin{aligned}
&\langle\Psi_1|V_{ee}|\Psi_2\rangle \\
&= \tfrac{\sqrt{1-\lambda}}{\sqrt{2}}\big[\beta_1^*\langle 1s\overline{1s}2s\overline{2s}p_1|V_{ee}|\big(|1s\overline{1s}p_2\overline{p_2}p_1\rangle + |1s\overline{1s}p_3\overline{p_3}p_1\rangle\big)\rangle \\
&\qquad\quad + \beta_2^*\langle 1s\overline{1s}2s\overline{2s}p_1|V_{ee}|1s\overline{1s}p_1\overline{p_1}p_2\rangle + \beta_3^*\langle 1s\overline{1s}2s\overline{2s}p_1|V_{ee}|1s\overline{1s}p_1\overline{p_1}p_3\rangle\big] \\
&\quad + \tfrac{i\sqrt{\lambda}}{\sqrt{2}}\big[\beta_2^*\langle 1s\overline{1s}2s\overline{2s}p_2|V_{ee}|\big(|1s\overline{1s}p_1\overline{p_1}p_2\rangle + |1s\overline{1s}p_3\overline{p_3}p_2\rangle\big)\rangle \\
&\qquad\quad + \beta_1^*\langle 1s\overline{1s}2s\overline{2s}p_2|V_{ee}|1s\overline{1s}p_2\overline{p_2}p_1\rangle + \beta_3^*\langle 1s\overline{1s}2s\overline{2s}p_2|V_{ee}|1s\overline{1s}p_2\overline{p_2}p_3\rangle\big] \\
&= \sqrt{2}\big(\sqrt{1-\lambda}\beta_1^* + i\sqrt{\lambda}\beta_2^*\big)(23|32).
\end{aligned}$$

In order for the energy to be minimized we therefore have the two restrictions

$$\sqrt{1-\lambda}\beta_1^* + i\sqrt{\lambda}\beta_2^* = 1, \quad |\beta_1|^2 + |\beta_2|^2 + |\beta_3|^2 = 1,$$

which, as in the Carbon case, imply that

$$\beta_1 = \sqrt{1-\lambda}, \quad \beta_2 = i\sqrt{\lambda}, \quad \beta_3 = 0.$$

The next stage is to determine the canonical pair density, which will clearly be a one-parameter family. We have

$$\rho_2^{\Psi} = c^2\rho_2^{\Psi_1}(x,y) + (1-c^2)\rho_2^{\Psi_2}(x,y) - 2c\sqrt{1-c^2}\langle\Psi_1|V_{x,y}|\Psi_2\rangle.$$

The pair density $\rho_2^{\Psi_1}(x,y)$ follows trivially from (6.88) and the result from Oxygen, and is

$$\begin{aligned}
\rho_2^{\Psi_1}(x,y) = \rho_2^{|1s\overline{1s}2s\overline{2s}\rangle}&(|x|,|y|) \\
&+ \tfrac{1}{2}\big[\rho^{|1s\overline{1s}2s\overline{2s}\rangle}(|y|)k^2e^{-Z|x|}\big((1-\lambda)x_1^2 + \lambda x_2^2\big) + x \leftrightarrow y\big] \\
&- \big(1s(|x|)1s(|y|) + 2s(|x|)2s(|y|)\big)k^2e^{-Z(|x|+|y|)/2} \\
&\qquad\qquad\qquad \times \big((1-\lambda)x_1y_1 + \lambda x_2y_2\big).
\end{aligned} \tag{6.89}$$

The canonical pair density for Ψ_2 is less straightforward as there is no obvious radially-symmetric pair density from which is can be derived by adding or subtracting a single orbital. The previous methods would suggest using $|1s\overline{1s}p_1\overline{p_1}\rangle + |1s\overline{1s}p_2\overline{p_2}\rangle + |1s\overline{1s}p_3\overline{p_3}\rangle$, which unfortunately leads to problems resulting from the fact that $|1s\overline{1s}p_i\overline{p_i}p_i\rangle = 0$, destroying the symmetry of the result. It therefore seems more sensible to directly compute the pair density:

$$\begin{aligned}\rho_2^{\Psi_2} = \tfrac{1}{2}(1-\lambda)\Big[&\rho_2^{|1s\overline{1s}p_2\overline{p_2}p_1\rangle} + \rho_2^{|1s\overline{1s}p_3\overline{p_3}p_1\rangle} + 2\langle 1s\overline{1s}p_2\overline{p_2}p_1|V_{x,y}|1s\overline{1s}p_3\overline{p_3}p_1\rangle\Big]\\ &+ \tfrac{1}{2}\lambda\Big[\rho_2^{|1s\overline{1s}p_1\overline{p_1}p_2\rangle} + \rho_2^{|1s\overline{1s}p_3\overline{p_3}p_2\rangle} + 2\langle 1s\overline{1s}p_1\overline{p_1}p_2|V_{x,y}|1s\overline{1s}p_3\overline{p_3}p_2\rangle\Big].\end{aligned}$$

Computing the two general terms required we see that

$$\begin{aligned}\rho_2^{|1s\overline{1s}p_i\overline{p_i}p_k\rangle} = \tfrac{1}{2}\Big[&|1s(x)|^2|1s(y)|^2 + 4|1s(x)|^2|p_i(y)|^2 + 2|1s(x)|^2|p_j(y)|^2\\ &+ |p_i(x)|^2|p_i(y)|^2 + 2|p_i(x)|^2|p_k(y)|^2 + x\leftrightarrow y\Big]\\ -\Big[&2\cdot 1s(x)1s(y)p_i(x)p_i(y) + 1s(x)1s(y)p_k(x)p_k(y)\\ &+ p_i(x)p_i(y)p_k(x)p_k(y)\Big]\end{aligned}$$

and $\langle 1s\overline{1s}p_i\overline{p_i}p_k|V_{x,y}|1s\overline{1s}p_j\overline{p_j}p_k\rangle = p_i(x)p_i(y)p_j(x)p_j(y)$.

We now combine the above results and note that all orbitals are real to obtain

$$\begin{aligned}&\rho_2^{|1s\overline{1s}p_i\overline{p_i}p_k\rangle} + \rho_2^{|1s\overline{1s}p_j\overline{p_j}p_k\rangle} + 2\langle 1s\overline{1s}p_i\overline{p_i}p_k|V_{x,y}|1s\overline{1s}p_j\overline{p_j}p_k\rangle =\\ &\tfrac{1}{2}\Big[2|1s(x)|^2|1s(y)|^2 + 4|1s(x)|^2\big(|p_i(y)|^2 + |p_j(y)|^2 + |p_k(y)|^2\big)\\ &\qquad + |p_i(x)|^2|p_i(y)|^2 + |p_j(x)|^2|p_j(y)|^2\\ &\qquad + 2\big(|p_i(x)|^2 + |p_j(x)|^2\big)|p_k(y)|^2 + x\leftrightarrow y\Big]\\ &-\Big[2\cdot 1s(x)1s(y)\big(p_i(x)p_i(y) + p_j(x)p_j(y) + p_k(x)p_k(y)\big)\\ &\qquad + \big(p_i(x)p_i(y) + p_j(x)p_j(y)\big)p_k(x)p_k(y) - 2p_i(x)p_i(y)p_j(x)p_j(y)\Big],\end{aligned}$$

and add and subtract the same terms to make the symmetry more obvious, obtaining

$$\begin{aligned}&\rho_2^{|1s\overline{1s}p_i\overline{p_i}p_k\rangle} + \rho_2^{|1s\overline{1s}p_j\overline{p_j}p_k\rangle} + 2\langle 1s\overline{1s}p_i\overline{p_i}p_k|V_{x,y}|1s\overline{1s}p_j\overline{p_j}p_k\rangle\\ &= \tfrac{1}{2}\big[2|1s(x)|^2|1s(y)|^2 + 4|1s(x)|^2\big(|p_1(y)|^2 + |p_2(y)|^2 + |p_3(y)|^2\big)\\ &\qquad + \big(|p_1(x)|^2|p_1(y)|^2 + |p_2(x)|^2|p_2(y)|^2 + |p_3(x)|^2|p_3(y)|^2\big)\\ &\qquad + 2\big(|p_1(x)|^2 + |p_2(x)|^2 + |p_3(x)|^2\big)|p_k(y)|^2 - 3|p_k(x)|^2|p_k(y)|^2 + x\leftrightarrow y\big]\\ &\quad - \big[2\cdot 1s(x)1s(y)\big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big)\\ &\qquad + \big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big)p_k(x)p_k(y) - |p_k(x)|^2|p_k(y)|^2\\ &\qquad - 2p_i(x)p_i(y)p_j(x)p_j(y)\big]\end{aligned}$$

We now use the symmetry of a number of terms under $x \leftrightarrow y$ and simplify to

$$\begin{aligned}
&\rho_2^{|1s\overline{1s}p_i\overline{p_i}p_k\rangle} + \rho_2^{|1s\overline{1s}p_j\overline{p_j}p_k\rangle} + 2\langle 1s\overline{1s}p_i\overline{p_i}p_k|V_{x,y}|1s\overline{1s}p_j\overline{p_j}p_k\rangle \\
&\quad = 2|1s(x)|^2|1s(y)|^2 + |p_1(x)|^2|p_1(y)|^2 + |p_2(x)|^2|p_2(y)|^2 + |p_3(x)|^2|p_3(y)|^2 \\
&\qquad - 2\cdot 1s(x)1s(y)\big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big) \\
&\qquad - \big[p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big]p_k(x)p_k(y) \\
&\qquad - 2|p_k(x)|^2|p_k(y)|^2 + 2p_i(x)p_i(y)p_j(x)p_j(y) \\
&\qquad + \Big[\big(2|1s(x)|^2 + |p_k(x)|^2\big)\big(|p_1(y)|^2 + |p_2(y)|^2 + |p_3(y)|^2\big) + x \leftrightarrow y\Big],
\end{aligned}$$

which shows that the canonical form of $\rho_2^{\Psi_2}$ is

$$\begin{aligned}
\rho_2^{\Psi_2} = {} & |1s(x)|^2|1s(y)|^2 + \tfrac{1}{2}\big(|p_1(x)|^2|p_1(y)|^2 + |p_2(x)|^2|p_2(y)|^2 + |p_3(x)|^2|p_3(y)|^2\big) \\
& - 1s(x)1s(y)\big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big) \\
& - (1-\lambda)|p_1(x)|^2|p_1(y)|^2 - \lambda|p_2(x)|^2|p_2(y)|^2 \\
& + (1-\lambda)p_2(x)p_2(y)p_3(x)p_3(y) + \lambda p_1(x)p_1(y)p_3(x)p_3(y) \\
& - \tfrac{1}{2}\big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big) \\
& \qquad\qquad \times \big((1-\lambda)p_1(x)p_1(y) + \lambda p_2(x)p_2(y)\big) \\
& + \tfrac{1}{2}\big]\big(2|1s(x)|^2 + (1-\lambda)|p_1(x)|^2 + \lambda|p_2(x)|^2\big) \\
& \qquad \times \big(|p_1(y)|^2 + |p_2(y)|^2 + |p_3(y)|^2\big) + x \leftrightarrow y\big].
\end{aligned}$$

In the same spirit as (6.89), we write $p_i(x) = kx_ie^{-Z|x|/2}$ and hence

$$\begin{aligned}
\rho_2^{\Psi_2} = {} & |1s(|x|)|^2|1s(|y|)|^2 + \tfrac{1}{2}k^4e^{-Z|x|}e^{-Z|y|}(x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2) \\
& - 1s(|x|)1s(|y|)k^2e^{-Z|x|/2}e^{-Z|y|/2}(x_1y_1 + x_2y_2 + x_3y_3) \\
& - \tfrac{1}{2}k^4e^{-Z|x|}e^{-Z|y|}(x_1y_1 + x_2y_2 + x_3y_3)\big((1-\lambda)x_1y_1 + \lambda x_2y_2\big) \\
& - k^4e^{-|x|}e^{-|y|}\big((1-\lambda)x_1^2y_1^2 + \lambda x_2^2y_2^2\big) \\
& + k^4e^{-|x|}e^{-|y|}\big((1-\lambda)x_2x_3y_2y_3 + \lambda x_1x_3y_1y_3\big) \\
& + \big(|1s(|x|)|^2k^2e^{-|y|}|y|^2 + |1s(|y|)|^2k^2e^{-|x|}|x|^2\big) \\
& + \tfrac{1}{2}k^4e^{-|x|}e^{-|y|}\big[(1-\lambda)(x_1^2|y|^2 + y_1^2|x|^2) + \lambda(x_2^2|y|^2 + y_2^2|x|^2)\big]\big].
\end{aligned}$$

It remains to compute the cross term $\langle\Psi_1|V_{x,y}|\Psi_2\rangle$ which, noting that terms containing Slater determinants that differ by more than 2 orbitals give zero, is given by

$$\begin{aligned}
&\langle\Psi_1|V_{x,y}|\Psi_2\rangle \\
&\quad= (1-\lambda)\big(\langle 1s\overline{1s}2s\overline{2s}p_1|V_{x,y}|1s\overline{1s}p_2\overline{p_2}p_1\rangle + \langle 1s\overline{1s}2s\overline{2s}p_1|V_{x,y}|1s\overline{1s}p_3\overline{p_3}p_1\rangle\big) \\
&\qquad+ \lambda\big(\langle 1s\overline{1s}2s\overline{2s}p_2|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2\rangle + \langle 1s\overline{1s}2s\overline{2s}p_2|V_{x,y}|1s\overline{1s}p_3\overline{p_3}p_2\rangle\big) \\
&\qquad+ i\sqrt{\lambda(1-\lambda)}\big(\langle 1s\overline{1s}2s\overline{2s}p_1|V_{x,y}|1s\overline{1s}p_1\overline{p_1}p_2\rangle \\
&\qquad\qquad\qquad\qquad+ \langle 1s\overline{1s}2s\overline{2s}p_2|V_{x,y}|1s\overline{1s}p_2\overline{p_2}p_1\rangle\big).
\end{aligned}$$

It is trivial to check that the two terms with imaginary coefficient vanish. The two remaining terms give

$$\begin{aligned}
\langle\Psi_1|V_{x,y}|\Psi_2\rangle &= (1-\lambda)2s(x)2s(y)\big(p_2(x)p_2(y) + p_3(x)p_3(y)\big) \\
&\quad+ \lambda 2s(x)2s(y)\big(p_1(x)p_1(y) + p_3(x)p_3(y)\big) \\
&= 2s(x)2s(y)k^2e^{-Z(|x|+|y|)/2}\big(p_1(x)p_1(y) + p_2(x)p_2(y) + p_3(x)p_3(y)\big) \\
&\quad- 2s(x)2s(y)k^2e^{-Z(|x|+|y|)/2}\big((1-\lambda)p_1(x)p_1(y) + \lambda p_2(x)p_2(y)\big).
\end{aligned}$$

Hence the total canonical pair density for the Boron ground state is given by

$$\begin{aligned}
&\rho_2^B = \\
&c^2\bigg[\rho_2^{|1s\overline{1s}2s\overline{2s}\rangle}(|x|,|y|) + \Big[\big(|1s(|y|)|^2 + |2s(|y|)|^2\big)k^2e^{-|x|}\big((1-\lambda)x_1^2 + \lambda x_2^2\big) + x \leftrightarrow y\Big] \\
&\qquad- \big(1s(|x|)1s(|y|) + 2s(|x|)2(|y|)\big)k^2e^{-|x|/2}e^{-|y|/2}\big((1-\lambda)x_1y_1 + \lambda x_2y_2\big)\bigg] \\
&\quad+ (1-c^2)\bigg[|1s(|x|)|^2|1s(|y|)|^2 + \tfrac{1}{2}k^4e^{-|x|}e^{-|y|}(x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2) \\
&\qquad\qquad- 1s(|x|)1s(|y|)k^2e^{-\frac{|x|}{2}}e^{-\frac{|y|}{2}}(x_1y_1 + x_2y_2 + x_3y_3) \\
&\qquad\qquad- \tfrac{1}{2}k^4e^{-|x|}e^{-|y|}(x_1y_1 + x_2y_2 + x_3y_3)\big((1-\lambda)x_1y_1 + \lambda x_2y_2\big) \\
&\qquad\qquad- k^4e^{-|x|}e^{-|y|}\big((1-\lambda)x_1^2y_1^2 + \lambda x_2^2y_2^2\big) \\
&\qquad\qquad+ k^4e^{-|x|}e^{-|y|}\big((1-\lambda)x_2x_3y_2y_3 + \lambda x_1x_3y_1y_3\big) \\
&\qquad\qquad+ \Big(|1s(|x|)|^2k^2e^{-|y|}|y|^2 + \tfrac{1}{2}k^4e^{-|x|}e^{-|y|}\big((1-\lambda)x_1^2|y|^2 + \lambda x_2^2|y|^2\big) \\
&\qquad\qquad\qquad\qquad+ x \leftrightarrow y\Big)\bigg] \\
&\quad- \sqrt{2}c\sqrt{1-c^2}2s(|x|)2s(|y|)k^2e^{-\frac{|x|}{2}}e^{-\frac{|y|}{2}} \\
&\qquad\qquad\qquad\qquad\times\big((x_1y_1 + x_2y_2 + x_3y_3) - (1-\lambda)x_1y_1 - \lambda x_2y_2\big)
\end{aligned}$$

Unlike all previous cases, this is not of a suitable form to apply Corollary 6.3.4.

This is due principally to the terms of the form $x_i^2 y_i^2$ for which are not present in the general symmetric function in (6.37). There seems to be no analogous proof for the symmetry of the critical points in this case.

As discussed in Section 6.3.11 it is difficult to perform numerics on these pair densities, due in part to the number of variables involved and also the range of numerical values encountered.

6.9 Dependence on the Core Orbitals

As mentioned in Sections 6.1.3 and 6.1.4, methods such as VSEPR and Walsh Diagrams ignore the core orbitals when determining bond angles. We wish to investigate whether our method allows the same approach, i.e. whether we get the same predicted bond angles if we ignore the core $1s$ electrons when calculating the bond angle.

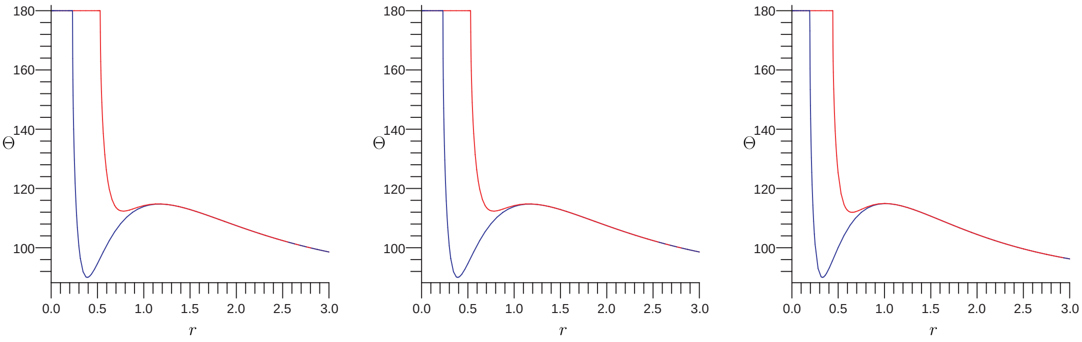
Intuitively this would indeed be the case if the $1s$-orbitals are suitably localised around the nucleus (i.e. have a high enough decay rate) to prevent significant interaction with the valence shell electrons at bond-like distances.

It is clear that the energy calculations require us to include the $1s$-orbitals in order to accurately determine the atomic ground state. However, the proofs of the global maxima of the above pair densities require only that the $A_i(r)$ are non-negative functions. We may therefore set $|\psi_{1s}|^2 \equiv 0$ and re-plot the bond angle diagrams.

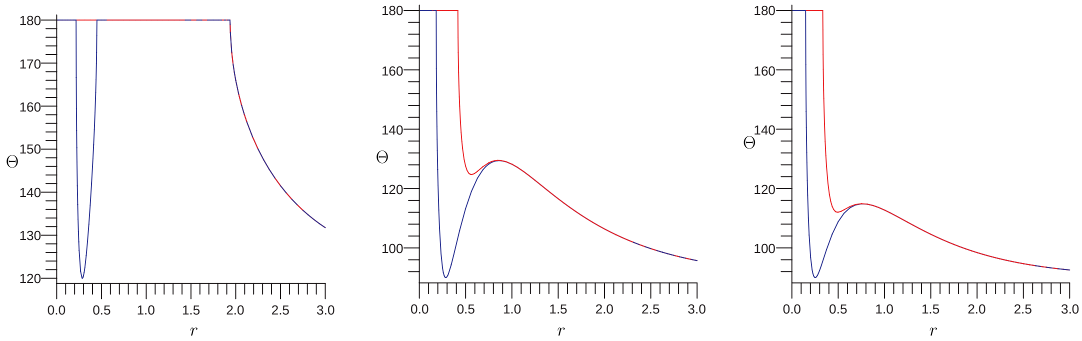
As can be seen from Figure 6.30, the removal of the $1s$-orbital dependence has a very small effect on the predicted bond angle at bonding distances. It is also clear that the effects at short range are very large, and those at long range are negligible. This agrees with the intuition that the $1s$-orbitals are dominant only close to the nucleus.

6.10 Conclusions and Open Problems

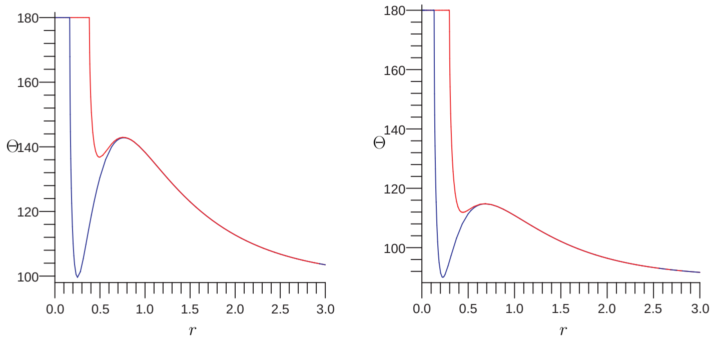
In this chapter we have seen that a simple (although in no way mathematically motivated) way of calculating the bond angle in an AH_2 molecule (Proposition 6.1.1), combined with the simple, but hopefully accurate, canonical wavefunctions derived in the previous chapters, leads to very good qualitative and relatively accurate quantitative results. Of particular interest is the discovery of a parameter which can be freely varied without affecting the energy of the central atom, but which leads to a range of predicted bond angles. The only case we have been



Carbon $\lambda = 0$ Carbon $\lambda = 1/2$ Nitrogen



Oxygen $\lambda = 0$ Oxygen $\lambda = 1/2$ Fluorine $\lambda = 0$



Fluorine $\lambda = 1/2$ Neon

Figure 6.30: Bond angle plots without the core $1s$-orbitals, the red curves include the $1s$-orbitals, whilst the blue curves use the same bond angle formulae but with $|\psi_{1s}|^2 \equiv 0$.

unable to treat using this method is that of Boron, which would be an interesting area for further investigation.

A further area of interest would be to consider low-lying excited states of the central atom. This should only be important when the spectral gaps are small, otherwise the energy gained by bonding will be insufficient to overcome the energy penalty for excitation of the central atom.

One other motivation for doing this comes from Pauling's rules in Section 6.1.2, which suggest that an atom can only form two bonds if it has two unpaired electrons. This is not the case in the ground states of Beryllium and Boron and suggests that the first excited states (which do indeed have two unpaired electrons in each case) should be considered.

Another extension would be to use higher order densities to explain notions such as double and triple bonds. For example, we would conjecture that a double bond may only form between two atoms if the four point density of the (valence electrons) of the full dimer system has a maximum that is suitably localised between the atoms. The precise mathematical definition of 'suitably localised between' is unclear but the investigation of a number of systems would hopefully shed light on the formation of higher order bonds.

These calculations require a relatively simple but energetically accurate ground state for a molecular system. The presence of multi-centre integrals (which, in the case of Slater type orbitals, are more complicated than single centre ones, see e.g. [Sug27, ZG29]) and the exponentially increasing size of the CI matrix make this more difficult than the atomic case.

However, the relative success of the simple approximations in the above cases suggests that it may be possible to reduce the dimension of CI calculations for molecules and still obtain accurate results. For example, it is conceivable that in order to perform an accurate molecular calculation on AH_2 it is only necessary to consider wavefunctions corresponding to the ground state configurations of each atom.

From Tables 2.1 and 5.1 it is clear that this method would drastically reduce the dimension of the full CI matrix. It is hoped that the calculations would be simple enough to enable understanding of why a particular choice from the range of bond angles with the same central atom energy is favourable to different molecules. It is also expected that in some cases the true bond angle lies outside this range, with the energy penalty to the central atom being lower than the energy gain from the outer atoms.

A very interesting direction would be to use the densities and pair densities to investigate the correlation between electrons. This would require a definition of correlation, probably similar to that used in multi-variate statistics. From DFT, one would hope to see that the correlation function is (to a large extent) independent of the number of electrons in the system. This could be seen as the first step in a mathematical justification of the correlation functions used in DFT, although more accurate densities and pair densities than those derived here may be necessary to see any interesting effects.

Chapter 7

Rate of Convergence of CI for Helium

7.1 Introduction

The aim of this chapter is to reproduce some of the results of [Hil85] with as much mathematical rigour as possible. The paper deals with rates of convergence of approximations, in particular, the result we are interested in concerns the rate of convergence of a Configuration Interaction calculation on the ground state of Helium. Specifically, if the bases chosen span the subspace with total angular momentum less than or equal to L, it gives the asymptotic rate of convergence as $E_L - E_{L-1} = \mathcal{O}(L+\frac{1}{2})^{-4}$ as $L \to \infty$. This gives that the overall energy error converges as $E - E_L = \mathcal{O}(L+\frac{1}{2})^{-3}$.

Furthermore, Hill derives the constant for the leading order term, finding it to be (Hill's (5.11a))

$$C_1 = 2\pi^2 \int_0^\infty |\psi(r,r,0)|^2 r^5 dr,$$

where ψ is the exact wavefunction whose coordinates are r_1, r_2 and r_{12}, the two nuclear-electron distances and the interelectron distance respectively.

The main interest in this result is that it may be used to extrapolate the results of low dimensional CI calculations to give a more accurate value for the energy with little computational cost. (See for example the recent work [BM06].) For this application it is important that we are able to derive the leading order constant.

In general we will follow Hill's method but hope to make a number of improvements both in terms of mathematical rigour and understandability. In particular Hill's Section III on the convergence rate of a general problem is very hard to follow and there are a number of typographical mistakes throughout the paper.

The main changes in our method are that we use an alternative, simpler, formulation of the energy error. This leads to replacing Hill's integrals containing ℓ-projections of $r_{12}^{-1}\psi$ (denoted ϕ_ℓ) by integrals involving derivatives with respect to r_1 and r_2 of the ℓ-projections of the wavefunction.

A further difference in our analysis is that we construct an explicit form of $f_\ell^k(r_<, r_>)$ whereas Hill uses a recursion formula to obtain an asymptotic result. The analysis of the remainder terms $R_{j,\ell}(r_<, r_>)$ is similar to that of Hill but we produce a more detailed version of his Appendix C, the necessary conditions seem optimal in this formulation of the problem.

One noteworthy point is that, in order to obtain Hill's asymptotic result, the wavefunction is required to be three times differentiable in r_{12}. The validity of this assumption will be discussed later.

7.2 Form of the Energy Error

Let ψ be the actual ground state solution to the Schrödinger equation for the Helium atom, $H\psi = E\psi$ where

$$H := -\frac{1}{2}\Delta_1 - \frac{1}{2}\Delta_2 - \frac{2}{|\mathbf{r_1}|} - \frac{2}{|\mathbf{r_2}|} + \frac{1}{|\mathbf{r_1} - \mathbf{r_2}|},$$

where $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^3$ are the positions of the two electrons and we later use the notation $|\mathbf{r}_1 - \mathbf{r}_2| := r_{12}$.

The ground state of the Helium atom has angular momentum zero (see for example the experimental data of [NIS06]) and is therefore invariant under simultaneous rotation of $\mathbf{r_1}$ and $\mathbf{r_2}$. Hence the ground state may be described by the lengths of the two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, denoted respectively r_1 and r_2, and θ, the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$. Equivalently, θ may be written in terms of r_1, r_2 and r_{12}, giving

$$\psi = \psi(r_<, r_>, r_{12})$$

where

$$r_< := \min\{r_1, r_2\}, \quad r_> := \max\{r_1, r_2\}$$

and

$$r_{12} := |\mathbf{r}_1 - \mathbf{r}_2| = (r_<^2 + r_>^2 - 2r_<r_> \cos\theta)^{1/2}.$$

We wish to investigate the energy of the approximate ground state wavefunction given by projection of ψ onto the space spanned by functions with angular momentum less than or equal to L. It is well known [AS72] that the Legendre polynomials

$$P_\ell := \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2-1)^\ell$$

are orthogonal with respect to the weighted inner product on L^2 given by

$$(f,g) := \int_0^\pi f(\theta)\overline{g(\theta)}\sin\theta d\theta.$$

It follows from the normalisation formula [AS72]

$$\int_0^\pi P_m(\cos\theta)P_n(\cos\theta)\sin\theta d\theta = \delta_{m,n}\frac{2}{2n+1}$$

that the corresponding orthonormal wavefunctions are given by

$$\Phi_\ell(\theta)(x) := (l+\tfrac{1}{2})^{1/2}P_\ell(\cos\theta), \tag{7.1}$$

which form a complete basis under this inner product. These $\Phi_\ell(\theta)$ are eigenfunctions of the angular momentum operator with eigenvalues $-\ell(\ell+1)$ and we must therefore take the projection onto the span of the first L of these, i.e.

$$P_L := \sum_{\ell=1}^{L}\big(\,\cdot\,,\Phi_\ell(\theta)\big)\Phi_\ell(\theta).$$

We denote the expansion coefficients by $\psi_\ell(r_<,r_>)$:

$$\psi_\ell(r_<,r_>) = \int_0^\pi \psi(r_<,r_>,r_{12})\Phi_\ell(\theta)\sin\theta d\theta. \tag{7.2}$$

and hence have

$$P_L\psi(r_<,r_>,r_{12}) = \sum_{\ell=0}^{L}\psi_\ell(r_<,r_>)\Phi_\ell(\theta).$$

By the completeness of the $\Phi_\ell(\theta)$ we denote the part of the wavefunction whose angular part does not lie in the span of the first L Legendre polynomials by

$$\psi_L^\perp := \psi - P_L\psi,$$

with an analogous expansion:

$$\psi_L^\perp(r_<,r_>,r_{12}) = \sum_{\ell=L+1}^{\infty}\psi_\ell(r_<,r_>)\Phi_\ell(\theta).$$

We now wish to investigate the error in the energy in terms of $\psi_L^\perp$. The following result [Fri04] gives a suitable form:

Given a symmetric operator H_0 defined on a dense domain $D \subseteq X$ of some Hilbert space X, we say $H_0 > 0$ if $\langle \psi, H_0 \psi \rangle > 0$ for all $\psi \in D \backslash \{0\}$. If $H_0 > 0$, the associated 'energy norm' and 'dual energy norm' are denoted by $\|\psi\|_e := (\langle \psi, H_0 \psi \rangle)^{1/2}$ $(\psi \in D)$, $\|\psi\|_{e'} := \sup_{\chi \in D \backslash \{0\}} (|\langle \chi, \psi \rangle| / \|\chi\|_e)$ $(\psi \in X)$. If $S \subseteq X$ is a subspace, the orthogonal complement $\{\psi \in X \,|\, \langle \chi, \psi \rangle = 0 \, \forall \chi \in S\}$ is denoted $S^\perp$.

Theorem 7.2.1 (Eigenvalue error asymptotics). *Consider a linear operator of form*

$$H = H_0 + H_1$$

where H, H_0, H_1 are symmetric operators defined on a joint dense domain $D \subseteq X$ of a Hilbert space X. Assume $H_0 > 0$ and let $\| \cdot \|_e$, $\| \cdot \|_{e'}$ denote the associated energy norm and dual energy norm.

(H1) (boundedness of H_1 with respect to energy norm) $\|H_1 \psi\| \leq C_0 \|\psi\|_e \; \forall \psi \in D$

(H2) (Existence of ground state with spectral gap) There exist $\psi \in D \backslash \{0\}$ (ground state), $E \in \mathbb{R}$ (ground state energy) and $\Delta > 0$ (spectral gap) such that $H\psi = E\psi$, $H\Big|_{\{Span \psi\}^\perp} \geq (E + \Delta) I$.

Let S be an arbitrary subspace of X with $S \subseteq D$, let P_S be the orthogonal projection operator onto the closure of S, let $P_S^\perp = I - P_S$, and let

$$\tilde{E} := \inf_{\phi \in S \backslash \{0\}} \frac{\langle \phi, H \phi \rangle}{\|\phi\|^2}$$

denote the 'Ritz-Galerkin ground state eigenvalue'. (If S is finite dimensional, the infimum is attained and equals the bottom eigenvalue of $P_S H P_S$ on S.) Assume $\tilde{E} \leq C_1$. There exists a constant $C > 0$ independent of the subspace S (and depending only on C_0, C_1 and Δ) such that the eigenvalue error $\tilde{E} - E$ satisfies

$$\Big| (\tilde{E} - E) - \langle H \rangle_{P_S^\perp \psi} \Big| \leq C \Big(\|P_S^\perp \psi\|^2 + \|P_S^\perp \psi\|_e^4 + \|[P_S, H_0] P_S^\perp \psi\|_{e'}^2 \Big). \qquad (7.3)$$

A simple lemma now reduces the number of terms to be calculated in our case:

Lemma 7.2.2. *Let* $H := -\Delta + V(x)$, $X = L^2(\mathbb{R}^d)$, $D = H^2(\mathbb{R}^d)$, $V : \mathbb{R}^d \to \mathbb{R}$, *be a general Schrödinger operator with*

$$V(x_1, ..., x_N) = -\sum_{\alpha=1}^{M}\sum_{i=1}^{N} \frac{Z_\alpha}{|x_i - R_\alpha|} + \sum_{1\le i<j\le N} \frac{1}{|x_i - x_j|}.$$

Then $H_0 = -\Delta$ *is symmetric and positive, the energy norm is given by* $\|\psi\|_e = \|\nabla\psi\|_{L^2}$ *and*

$$\|V(x)\psi\| \le C_1\|\psi\|_e \tag{7.4}$$

for some $C_1 > 0$. *Furthermore, for the Helium Ground State with*

$$S_L := \mathit{Span}\{\Phi_\ell(\theta),\ \ell = 0, \dots, L\},$$

we have $[P_{S_L}, H_0] = 0$ *and defining*

$$E_L := \inf_{\phi\in S_L\setminus\{0\}} \frac{\langle\phi, H\phi\rangle}{\|\phi\|^2}$$

and $\psi_L^\perp := (I - P_{S_L})\psi$ *gives*

$$E_L - E \le \|\psi_L^\perp\|_e^2 + C_1\|\psi_L^\perp\|\,\|\psi_L^\perp\|_e + C_2\Big(\|\psi_L^\perp\|^2 + \|\psi_L^\perp\|_e^4\Big),$$

for some $C_2 > 0$.

Proof The symmetry, positivity and norm equivalence follow trivially from integration by parts. The bound on $\|V(x)\psi\|$ follows from Hardy's inequality in $\mathbb{R}^3$, $\int_{\mathbb{R}^3} \frac{|\phi(x)|^2}{|x|^2}dx \le 4\int_{\mathbb{R}^3} |\nabla\phi(x)|^2dx$.

Noting that the Helium ground state is symmetric, we use that, for symmetric functions $f(r_<, r_>, \theta)$, H_0 may be rewritten [KB77] as $H_0 = -\frac{1}{2}(\Delta_1' + \Delta_2')$ where

$$\Delta_i' := \frac{1}{r_i^2}\frac{\partial}{\partial r_i} r_i^2 \frac{\partial}{\partial r_i} + \frac{1}{r_i^2 \sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}. \tag{7.5}$$

Hence, since

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Phi_\ell(\theta) = -\ell(\ell+1)\Phi_\ell(\theta) \tag{7.6}$$

we have that $[P_{S_L}, H_0] = 0$.

This, along with (7.3), gives

$$0 \le \tilde{E} - E \le \langle H\rangle_{\psi_L^\perp} + C_2\Big(\|\psi_L^\perp\|^2 + \|\psi_L^\perp\|_e^4\Big).$$

Decomposing the first term on the right hand side as

$$\langle H\rangle_{\psi_L^\perp} = \|\psi_L^\perp\|_e^2 + \langle V\rangle_{\psi_L^\perp},$$

and using the Cauchy-Schwarz inequality along with (7.4) gives the result.

Note in particular that the constant of one for the $\|\psi_L^\perp\|_e^2$ term is optimal. □

We therefore see that, in order to determine the rate of decay of the energy error, we need to determine the rate of decay of the two terms $\|\psi_L^\perp\|$ and $\|\psi_L^\perp\|_e$. As mentioned in the introduction, this is in contrast to Hill's method, the energy norm replacing $\|\phi_L^\perp\|$ [Hil85].

7.3 Form of the Error Estimate Norms

Consider the part of the wavefunction not captured by the first L angular momentum eigenfunctions,

$$\psi_L^\perp = \sum_{\ell=L+1}^{\infty} \psi_\ell(r_1, r_2)\Phi_\ell(\theta). \tag{7.7}$$

Lemma 7.3.1. *The following identities hold:*

$$\|\psi_L^\perp\|^2 = 16\pi^2 \sum_{\ell=L+1}^{\infty} \int_0^\infty \int_0^{r_>} |\psi_\ell(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_>, \tag{7.8}$$

$$\begin{aligned}\|\psi_L^\perp\|_e^2 = 8\pi^2 &\sum_{\ell=L+1}^{\infty} \int_0^\infty \int_0^{r_>} \big(|\tfrac{\partial}{\partial r_<}\psi_\ell(r_<, r_>)|^2 + |\tfrac{\partial}{\partial r_>}\psi_\ell(r_<, r_>)|^2\big) r_<^2 r_>^2 dr_< dr_> \\ &+ 8\pi^2 \sum_{\ell=L+1}^{\infty} \int_0^\infty \int_0^{r_>} \ell(\ell+1)(r_<^2 + r_>^2)|\psi_\ell(r_<, r_>)|^2 dr_< dr_>\end{aligned} \tag{7.9}$$

Proof Noting that $\int_0^\pi |\psi_L^\perp|^2 d\theta = \sum_{\ell=L+1}^\infty |\psi_\ell|^2$, using spherical polar coordinates and integrating over the three angles that are independent of θ (essentially ϕ_1, ϕ_2 and θ_1, leaving $\theta_1 - \theta_2 =: \theta$) gives a factor of $8\pi^2$. We are hence left with only the radial integrals, and using the Dominated Convergence Theorem gives

$$\|\psi_L^\perp\|^2 = 8\pi^2 \sum_{\ell=L+1}^{\infty} \int_0^\infty \int_0^\infty |\psi_\ell(r_1, r_2)|^2 r_1^2 r_2^2 dr_1 dr_2.$$

Using the general identity

$$\int_0^\infty\int_0^\infty f(r_1,r_2)dr_1dr_2 = \int_0^\infty\int_0^{r_>}\big(f(r_<,r_>)+f(r_>,r_<)\big)dr_<dr_> \tag{7.10}$$

and the fact that $|\psi_\ell(r_1,r_2)|^2r_1^2r_2^2$ is symmetric in r_1 and r_2, gives the first result.

For the second case we use (7.5) and (7.6) which give

$$H_0\psi_L^\perp = -\frac{1}{2}\sum_{\ell=L+1}^{\infty}\sum_{i=1}^{2}\Big(\frac{1}{r_i^2}\frac{\partial}{\partial r_i}r_i^2\frac{\partial}{\partial r_i} - \frac{\ell(\ell+1)}{r_i^2}\Big)\psi_\ell(r_1,r_2)\Phi_\ell(\theta)$$

and, as with the previous case, forming $\langle\psi_L^\perp, H_0\psi_L^\perp\rangle$, integrating out over the three independent angular coordinates, and using the fact that $\Phi_\ell(\theta)$ is an orthonormal basis gives

$$\begin{aligned}\|\psi_L^\perp\|_e^2 = -4\pi^2\int_0^\infty\int_0^\infty\sum_{\ell=L+1}^{\infty}\sum_{i=1}^{2}\psi_\ell(r_1,r_2)\Big(\frac{1}{r_i^2}\frac{\partial}{\partial r_i}r_i^2\frac{\partial}{\partial r_i} - \frac{\ell(\ell+1)}{r_i^2}\Big)\\ \times\,\psi_\ell(r_1,r_2)^*r_1^2r_2^2dr_1dr_2.\end{aligned}$$

Using the Dominated Convergence Theorem to swap the sum and integrals (the modulus of the sum up to a finite N is bounded by $\sum_{\ell=L+1}^{N}(|\psi_\ell(r_<,r_>)|^2+|H_0\psi_\ell(r_<,r_>)|^2)\leq|\psi_L^\perp|^2+|H_0\psi_L^\perp|^2$, which is integrable) and integrating the first term by parts gives

$$\|\psi_L^\perp\|_e^2 = 4\pi^2\sum_{\ell=L+1}^{\infty}\int_0^\infty\int_0^\infty\sum_{i=1}^{2}\Big(|\tfrac{\partial}{\partial r_i}\psi_\ell(r_1,r_2)|^2+\frac{\ell(\ell+1)}{r_i^2}|\psi_\ell(r_1,r_2)|^2\Big)r_1^2r_2^2dr_1dr_2.$$

Noting that the integrand is once again symmetric in r_1 and r_2 and using (7.10) gives the result. □

Equation (7.8) mirrors (5.87) of [Hil85], although the proof is more rigorous, whilst the second result is new.

7.4 Form of the Wavefunction

7.4.1 Expansion of the Wavefunction

For the start of this section we follow the proof of Theorem 2 of [Hil85], which derives a form for $b_\ell(r_1,r_2)$ (which is later set to $\psi_\ell(r_1,r_2)$), essentially by applying Taylor's theorem to the definition of $\psi_\ell(r_<,r_>)$ in (7.2) (equivalent to Hill's (4.34)).

We will instead apply Taylor's theorem to the wavefunction $\psi(r_<, r_>, r_{12})$ and then project onto $\Phi_\ell(\theta)$.

Lemma 7.4.1. *Suppose that* $\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})$ *exist for* $r_> - r_< \leq r_{12} \leq r_< + r_>$, $0 \leq j \leq J-1$ *and* $\frac{\partial^{J-1}}{\partial r_{12}^{J-1}}\psi(r_<, r_>, r_{12})$ *is absolutely continuous for* $r_> - r_< \leq r_{12} \leq r_< + r_>$. *Then* $\psi(r_<, r_>, r_{12})$ *can be expanded as*

$$\begin{aligned}\psi(r_<, r_>, r_{12}) = &\sum_{j=0}^{J-1} \frac{(r_{12} - (r_> - r_<))^j}{j!} \left[\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})\right]_{r_{12}=r_>-r_<} \\ &+ \int_{r_>-r_<}^{r_{12}} \frac{(r_{12}-t)^{J-1}}{(J-1)!} \frac{\partial^J}{\partial r_{12}^J}\psi(r_<, r_>, t)dt.\end{aligned}$$

Proof This is simply the application of Taylor's theorem about $\theta = 0$ (which is equivalent to $r_{12} = r_> - r_<$) with the Cauchy form of the remainder. □

It is now clear that we may project each of the terms in the (finite) expansion of $\psi(r_<, r_>, r_{12})$ to obtain an expression for the partial waves $\psi_\ell(r_<, r_>)$. The first thing of note is that, for $\ell > 0$, the term $\left[\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})\right]_{r_{12}=r_>-r_<}$ is simply a multiplicative factor as it is independent of θ. For the same reason, for $\ell > 0$, the term corresponding to $j = 0$ contributes zero. To simplify notation we denote

$$\xi_\ell^j(r_<, r_>) := \int_0^\pi \frac{(r_{12} - (r_> - r_<))^j}{j!}\Phi_\ell(\theta)\sin\theta d\theta,$$

the coefficients in the expansion by

$$A_j(r_<, r_>) := \left[\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})\right]_{r_{12}=r_>-r_<}, j = 0, \ldots, J-1,$$

the last term in the expansion by

$$R_J(r_<, r_>, r_{12}) := \int_{r_>-r_<}^{r_{12}} \frac{(r_{12}-t)^J}{J!} \frac{\partial^{J+1}}{\partial r_{12}^{J+1}}\psi(r_<, r_>, t)dt,$$

and its projection onto $\Phi_\ell(\theta)$ as

$$R_{J,\ell}(r_<, r_>) := \int_0^\pi R_J(r_<, r_>, r_{12})\Phi_\ell(\theta)\sin\theta d\theta.$$

Hence we see that, for $\ell > 0$, and $\psi(r_<, r_>, r_{12})$ J-times differentiable with respect to r_{12}, the partial waves can be written as

$$\psi_\ell(r_<, r_>) = \sum_{j=1}^{J-1} A_j(r_<, r_>)\xi_\ell^j(r_<, r_>) + R_{J-1,\ell}(r_<, r_>). \tag{7.11}$$

The norms given in (7.8) and (7.9) depend on $|\psi_\ell(r_<, r_>)|^2$ and $|\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)|^2$, where $r_\gamma \in r_<, r_>$. It is immediately clear from (7.11) that, for $\ell > 0$, these correspond to

$$\begin{aligned}|\psi_\ell(r_<, r_>)|^2 = &\sum_{i,j=1}^{J-1} A_i(r_<, r_>)\xi_\ell^i(r_<, r_>)A_j(r_<, r_>)^*\xi_\ell^j(r_<, r_>)^* \\ &+ 2\mathrm{Re}\Big(R_{J-1,\ell}(r_<, r_>)^* \sum_{k=1}^{J-1} A_k(r_<, r_>)\xi_\ell^k(r_<, r_>)\Big) \\ &+ |R_{J-1,\ell}(r_<, r_>)|^2\end{aligned} \tag{7.12}$$

and

$$\begin{aligned}|\tfrac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)|^2 = &\sum_{i,j=1}^{J-1} \tfrac{\partial}{\partial r_\gamma}\big(A_i(r_<, r_>)\xi_\ell^i(r_<, r_>)\big)\tfrac{\partial}{\partial r_\gamma}\big(A_j(r_<, r_>)\xi_\ell^j(r_<, r_>)\big)^* \\ &+ 2\mathrm{Re}\Big(\tfrac{\partial}{\partial r_\gamma}R_{J-1,\ell}(r_<, r_>)^* \sum_{k=1}^{2J-1} \tfrac{\partial}{\partial r_\gamma}\big(A_k(r_<, r_>)\xi_\ell^k(r_<, r_>)\big)\Big) \\ &+ |\tfrac{\partial}{\partial r_\gamma}R_{J-1,\ell}(r_<, r_>)|^2.\end{aligned} \tag{7.13}$$

We now require $\iint |\psi_\ell(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_>$, $\iint |\psi_\ell(r_<, r_>)|^2 r_>^2 dr_< dr_>$, $\iint |\psi_\ell(r_<, r_>)|^2 r_<^2 dr_< dr_>$ and $\iint |\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_>$. In order to simplify notation, we introduce the norm on symmetric functions

$$\|f\|_{a,b} := \left(\int_0^\infty \int_0^{r_>} |f(r_<, r_>)|^2 r_<^a r_>^b \, dr_< dr_>\right)^{1/2},$$

which comes from the inner product $(f, g)_{a,b} := \iint f g^* r_<^a r_>^b dr_< dr_>$. Inserting the expansions (7.12) and (7.13) into $\|\psi_\ell(r_<, r_>)\|_{a,b}$ and $\|\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)\|_{a,b}$, using the linearity of the integrals, and applying Cauchy-Schwarz, we obtain the bounds

$$\begin{aligned}\|\psi_\ell(r_<, r_>)\|_{a,b}^2 \leq &\sum_{i,j=1}^{J-1} \|A_i(r_<, r_>)\xi_\ell^i(r_<, r_>)\|_{a,b} \, \|A_j(r_<, r_>)\xi_\ell^j(r_<, r_>)\|_{a,b} \\ &+ 2\sum_{k=1}^{J-1} \|A_k(r_<, r_>)\xi_\ell^k(r_<, r_>)\|_{a,b} \, \|R_{J-1,\ell}(r_<, r_>)\|_{a,b} + \|R_{J-1,\ell}(r_<, r_>)\|_{a,b}^2\end{aligned} \tag{7.14}$$

and, using the product rule for differentiation,

$$\begin{aligned}
&\|\tfrac{\partial}{\partial r_\gamma}\psi_\ell(r_<,r_>)\|_{a,b}^2 \qquad\qquad (7.15)\\
&\le \sum_{i,j=1}^{J-1}\Big[\|\xi_\ell^i(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}A_i(r_<,r_>)\|_{a,b}\,\|\xi_\ell^j(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}A_j(r_<,r_>)\|_{a,b}\\
&\qquad + 2\|\xi_\ell^i(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}A_i(r_<,r_>)\|_{a,b}\,\|A_j(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<,r_>)\|_{a,b}\\
&\qquad + \|A_i(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^i(r_<,r_>)\|_{a,b}\,\|A_j(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<,r_>)\|_{a,b}\Big]\\
&\quad + 2\|\tfrac{\partial}{\partial r_\gamma}R_{J-1,\ell}(r_<,r_>)\|_{a,b}\sum_{k=1}^{J-1}\|A_k(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^k(r_<,r_>)\|_{a,b}\\
&\quad + 2\|\tfrac{\partial}{\partial r_\gamma}R_{J-1,\ell}(r_<,r_>)\|_{a,b}\sum_{k=1}^{2J-1}\|\xi_\ell^k(r_<,r_>)\tfrac{\partial}{\partial r_\gamma}A_k(r_<,r_>)\|_{a,b}\\
&\quad + \|\tfrac{\partial}{\partial r_\gamma}R_{J-1,\ell}(r_<,r_>)\|_{a,b}^2.
\end{aligned}$$

We now have a complete list of the required integrals and see that we need only consider those involving a single $\xi_\ell^j(r_<,r_>)$ or $\frac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<,r_>)$. The next stage is to obtain a more explicit form for the projections $\xi_\ell^j(r_<,r_>)$ and their derivatives. However, before we do this, we state the main result of this chapter.

7.5 Statement of the Result

Theorem 7.5.1. *Let $\psi(r_<,r_>,r_{12})$ be such that $\frac{\partial^j}{\partial r_{12}^j}\psi(r_<,r_>,r_{12})$ exist for $r_> - r_< \le r_{12} \le r_< + r_>$, $0 \le j \le 2J$. Define the functions*

$$\begin{aligned}
\xi_\ell^j(r_<,r_>) :&= \int_0^\pi \frac{(r_{12}-(r_>-r_<))^j}{j!}\Phi_\ell(\theta)\sin\theta\, d\theta,\\
A_j(r_<,r_>) :&= \left[\frac{\partial^j}{\partial r_{12}^j}\psi(r_<,r_>,r_{12})\right]_{r_{12}=r_>-r_<},\ j = 0,\dots,2J-1 \qquad (7.16)
\end{aligned}$$

and

$$R_j(r_<,r_>,r_{12}) := \int_{r_>-r_<}^{r_{12}} \frac{(r_{12}-t)^j}{j!}\frac{\partial^{j+1}}{\partial r_{12}^{j+1}}\psi(r_<,r_>,t)\, dt.$$

The large ℓ behaviour of

$$\psi_\ell(r_<,r_>) := \int_0^\pi \psi(r_<,r_>,r_{12})\Phi_\ell(\theta) sin\theta\, d\theta$$

is given by

$$\psi_\ell(r_<, r_>) = \sum_{j=1}^{2J-1} A_j(r_<, r_>)\xi_\ell^j(r_<, r_>) + R_{2J-1,\ell}(r_<, r_>) \tag{7.17}$$

where

$$R_{2J-1,\ell}(r_<, r_>) := \int_0^\pi R_{2J-1}(r_<, r_>, r_{12})\Phi_\ell(\theta)\sin\theta\, d\theta.$$

Furthermore suppose that $A(r_<, r_>)$ *satisfies* $A(\,\cdot\,, r_>) \in C^N(\mathbb{R}^+)$*, and there exists* $\tilde{A} : \mathbb{R}^+ \to \mathbb{R}$ *such that*

$$\tilde{A}(r_>) \geq \left| \left[\tfrac{\partial^n}{\partial r_<^n} |A(r_<, r_>)|^2 \right]_{r_<=s} \right|, \ \forall\, s \in \mathbb{R}^+,\ n = 0, \ldots, N.$$

Define a norm on the set of symmetric functions $f : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ *by*

$$\|f\|_{a,b} := \left(\int_0^\infty \int_0^{r_>} |f(r_<, r_>)|^2 r_<^a r_>^b \, dr_< dr_> \right)^{1/2}$$

then, for any $M < N$*,*

$$\begin{aligned} &\|A(r_<, r_>)\xi_\ell^j(r_<, r_>)\|_{a,b}^2 \\ &= \mathcal{O}(\ell^{-2j-4}) \Bigg[\sum_{m=0}^{M} \int_0^\infty r_>^{2j+1+a+b} \frac{(r_< - r_>)^m}{m!} \left[\tfrac{\partial^m}{\partial r_<^m} |A(r_<, r_>)|^2 \right]_{r_<=r_>} dr_> \\ &\qquad + \sum_{n=M+1}^{N} \mathcal{O}(\ell^{-n}) \int_0^\infty r_>^{2j+1+a+b+n} \tilde{A}(r_>) dr_> \Bigg] \end{aligned} \tag{7.18}$$

Define

$$I_j(r_<, r_>) := \int_{r_>-r_<}^{r_<+r_>} \left| \frac{\partial^j}{\partial r_{12}^j} \psi(r_<, r_>, t) \right|^2 dt$$

and suppose that $I_{2J} < \infty$*. Then*

$$\lim_{\ell\to\infty} \ell^{2J} R_{2J-1,\ell}(r_<, r_>) = 0 \tag{7.19}$$

If, in addition,

$$\int_0^\infty \int_0^{r_>} r_<^{4J-1} I_{2J}(r_<, r_>) r_<^2 r_>^2 \, dr_< dr_> < \infty$$

then

$$\lim_{\ell\to\infty} \ell^{4J} \int_0^\infty \int_0^{r_>} |R_{2J-1,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 \, dr_< dr_> = 0. \tag{7.20}$$

Theorem 7.5.2. *Let* $\psi(r_<, r_>, r_{12})$ *be such that* $\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})$ *exist for* $r_> - r_< \leq r_{12} \leq r_< + r_>$, $0 \leq j \leq 2J+1$ *and let* $r_\gamma \in \{r_<, r_>\}$. *With the same notation as in Theorem 7.5.1, the large-*ℓ *behaviour of* $\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)$ *is given by*

$$\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>) = \sum_{j=1}^{2J} \frac{\partial}{\partial r_\gamma}(A_j(r_<, r_>)\xi_\ell^j(r_<, r_>)) + \frac{\partial}{\partial r_\gamma}R_{2J,\ell}(r_<, r_>). \tag{7.21}$$

Furthermore, for $A(r_<, r_>)$ *and* $\tilde{A}(r_>)$ *as in Theorem 7.5.1, for any* $M < N$,

$$\begin{aligned}
&\|A(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<, r_>)\|_{a,b}^2 \\
&= \mathcal{O}(\ell^{-2j-2})\left[\sum_{m=0}^{M}\int_0^\infty r_>^{2j-1+a+b}(r_< - r_>)^m \left[\tfrac{\partial^m}{\partial r_<^m}|A(r_<, r_>)|^2\right]_{r_<=r_>} dr_> \right.\\
&\qquad\qquad \left. + \sum_{n=M+1}^{N}\mathcal{O}(\ell^{-n})\int_0^\infty r_>^{2j+1+a+b}\tilde{A}(r_>)dr_>\right].
\end{aligned} \tag{7.22}$$

If $I_{2J+1} < \infty$ *then*

$$\lim_{\ell\to\infty}\ell^{2J}\frac{\partial}{\partial r_\gamma}R_{2J,\ell}(r_<, r_>) = 0, \tag{7.23}$$

and if, in addition,

$$\int_0^\infty\int_0^{r_>} r_<^{4J-1}I_{2J+1}(r_<, r_>)r_<^2 r_>^2\, dr_< dr_> < \infty,$$

then

$$\lim_{\ell\to\infty}\ell^{4J}\int_0^\infty\int_0^{r_>} |\tfrac{\partial}{\partial r_\gamma}R_{2J,\ell}(r_<, r_>)|^2 r_<^2 r_>^2\, dr_< dr_> = 0. \tag{7.24}$$

These two results are analogous to Theorems 2 and 3 of [Hil85], although formulated in a slightly different way.

7.6 The Decay Rate of $\xi_\ell^j(r_<, r_>)$

For this section we follow the proof of Hill's Theorem 2 (specifically page 1180) by noting that

$$\xi_\ell^j(r_<, r_>) = \frac{1}{j!}\sum_{k=0}^{j}\int_0^\pi \binom{j}{k}(-1)^k(r_> - r_<)^{j-k}r_{12}^k\Phi_\ell(\theta)\sin\theta d\theta.$$

Since the Legendre polynomial expansion of r_{12}^{2k} terminates at $\ell = k$, taking $\ell > j/2$ gives

$$\xi_\ell^j(r_<, r_>) = \frac{1}{j!} \sum_{k=0}^{\lfloor (j+1)/2 \rfloor} \binom{j}{2k-1} (-1)^{j-2k+1} (r_> - r_<)^{j-2k+1} \int_0^\pi r_{12}^{2k-1} \Phi_\ell(\theta) \sin\theta d\theta$$

where the notation $\lfloor n \rfloor$ means the greatest integer $m \leq n$. We would now like to show that all terms in this sum decay at the same rate, i.e. that the rate is independent of k.

The first stage of this is to write the Legendre polynomial expansions of odd powers of r_{12} as

$$r_{12}^{2k-1} =: \sum_{\ell=0}^{\infty} (\ell + \tfrac{1}{2})^{-1/2} f_\ell^k(r_<, r_>) \Phi_\ell(\theta), \tag{7.25}$$

from which it immediately follows (using the orthonormality of $\Phi_\ell(\theta)$) that

$$\int_0^\pi r_{12}^{2k-1} \Phi_\ell(\theta) \sin\theta d\theta = (\ell + \tfrac{1}{2})^{-1/2} f_\ell^k(r_<, r_>).$$

We would therefore like to find an explicit form for the $f_\ell^k(r_<, r_>)$ which will allow us to estimate the necessary decay rates. Hill bypasses this step by deriving recursion relations for the $f_\ell^k(r_<, r_>)$ but we feel that an exact formulation will give more insight and also make the calculation of the derivatives easier.

7.6.1 An Explicit Form for $f_\ell^k(r_<, r_>)$

We begin by considering the form of $f_\ell^k(r_<, r_>)$, the (up to a multiplicative constant) projection of r_{12}^{2k-1} onto the ℓth angular momentum eigenstate $\Phi_\ell(\theta)$:

Lemma 7.6.1.

$$\begin{aligned} f_\ell^k(r_<, r_>) &= r_<^\ell r_>^{-\ell-1} \sum_{j=0}^{k} \frac{a_j^k r_<^{2(k-j)} r_>^{2j}}{\prod_{m=1}^{k-j} (2\ell + 2m + 1) \prod_{n=1}^{j} (2\ell - 2n + 1)} \\ &=: r_<^\ell r_>^{-\ell-1} \sum_{j=0}^{k} \frac{a_j^k r_<^{2(k-j)} r_>^{2j}}{P_j^k(\ell)} \end{aligned} \tag{7.26}$$

where $P_j^k(\ell)$ is a polynomial in ℓ of degree k and $a_j^k \in \mathbb{Z}$ are given by

$$a_0^k = \prod_{i=1}^{k} (2k - 2i + 1), \quad a_j^k = (-1)^j \binom{k}{j} a_0^k, \ j = 1, \ldots, k,$$

with the convention throughout that an empty product takes the value 1.

Proof We prove Lemma 7.6.1 in a number of stages. Firstly we show that

Lemma 7.6.2.

$$f_\ell^k(r_<, r_>) = r_<^\ell r_>^{-\ell-1} \sum_{j=0}^{k} c_j^{k,\ell} r_<^{2(k-j)} r_>^{2j} \tag{7.27}$$

where $c_j^{k,\ell}$ *is given by the recursion relation*

$$c_j^{k+1,\ell} = c_j^{k,\ell} + c_{j-1}^{k,\ell} - \frac{2\ell}{2\ell-1} c_{j-1}^{k,\ell-1} - \frac{2(\ell+1)}{2\ell+3} c_j^{k,\ell+1}$$

and the initial condition $c_0^{0,\ell} = 1$.

We then derive a suitable expression for $c_j^{k,\ell}$ which proves Lemma 7.6.1.
Proof of Lemma 7.6.2 We recall expansion (7.25), and note that

$$r_{12}^{2k+1} = (r_<^2 + r_>^2 - 2r_< r_> \cos\theta) r_{12}^{2k-1}. \tag{7.28}$$

Inserting (7.25) into (7.28) gives

$$r_{12}^{2k+1} = (r_<^2 + r_>^2 - 2r_< r_> \cos\theta) \sum_{\ell=0}^{\infty} (\ell + \tfrac{1}{2})^{-1/2} f_\ell^k(r_<, r_>) \Phi_\ell(\theta). \tag{7.29}$$

In order to determine the action of $\cos\theta$ on $\Phi_\ell(\theta)$, we recall the definition (7.1) of $\Phi_\ell(\theta)$ and the recursion relation for the Legendre polynomials [AS72]

$$(\ell+1) P_{\ell+1}(\cos\theta) = (2\ell+1)\cos\theta P_\ell(\cos\theta) - \ell P_{\ell-1}(\cos\theta). \tag{7.30}$$

It follows from (7.30) that

$$\cos\theta (\ell + \tfrac{1}{2})^{1/2} P_\ell = \frac{\ell+1}{2\ell+1} (\ell + \tfrac{1}{2})^{1/2} P_{\ell+1} + \frac{\ell}{2\ell+1} (\ell + \tfrac{1}{2})^{1/2} P_{\ell-1}$$

and hence by (7.1)

$$\begin{aligned} \cos\theta\, \Phi_\ell(\theta) &= \frac{1}{\sqrt{2}(2\ell+1)^{1/2}} \Big(\frac{\ell+1}{(\ell+\frac{3}{2})^{1/2}} \Phi_{\ell+1}(\theta) + \frac{\ell}{(\ell-\frac{1}{2})^{1/2}} \Phi_{\ell-1}(\theta) \Big) \\ &= \frac{\ell+1}{[(2\ell+1)(2\ell+3)]^{1/2}} \Phi_{\ell+1}(\theta) + \frac{\ell}{[(2\ell+1)(2\ell-1)]^{1/2}} \Phi_{\ell-1}(\theta). \end{aligned} \tag{7.31}$$

Inserting (7.31) into (7.29) and comparing coefficients of $\Phi_\ell(\theta)$ with (7.25) gives

$$\begin{aligned}(\ell+\tfrac{1}{2})^{-1/2} f_\ell^{k+1}(r_<, r_>) =& (\ell+\tfrac{1}{2})^{-1/2}(r_<^2 + r_>^2) f_\ell^k(r_<, r_>) \\ & - 2r_< r_> \frac{\ell(\ell-\frac{1}{2})^{-1/2}}{[(2\ell+1)(2\ell-1)]^{1/2}} f_{\ell-1}^k(r_<, r_>) \\ & - 2r_< r_> \frac{(\ell+1)(\ell+\frac{3}{2})^{-1/2}}{[(2\ell+1)(2\ell+3)]^{1/2}} f_{\ell+1}^k(r_<, r_>),\end{aligned}$$

which implies the recursion relation

$$\begin{aligned}f_\ell^{k+1}(r_<, r_>) = (r_<^2 + r_>^2) f_\ell^k(r_<, r_>) - r_< r_> \frac{2\ell}{2\ell-1} f_{\ell-1}^k(r_<, r_>) \\ - r_< r_> \frac{2(\ell+1)}{2\ell+3} f_{\ell+1}^k(r_<, r_>). \qquad (7.32)\end{aligned}$$

It is clear from the standard form

$$r_{12}^{-1} = \sum_{l=0}^{\infty} (\ell+\tfrac{1}{2})^{-1/2} \frac{r_<^\ell}{r_>^{\ell+1}} \Phi_\ell(\theta)$$

that the expansion (7.27) holds for $n = 0$. Assuming the expansion (7.27) for $f_\ell^k(r_<, r_>)$ and inserting it into (7.32) we have

$$\begin{aligned}f_\ell^{k+1}(r_<, r_>) = r_<^\ell r_>^{-\ell-1} \Bigg[\sum_{j=0}^{k} c_j^{k,\ell} r_<^{2(k-j+1)} r_>^{2j} + \sum_{j=0}^{k} c_j^{k,\ell} r_<^{2(k-j)} r_>^{2(j+1)} \\ - \sum_{j=0}^{k} \frac{2\ell}{2\ell-1} c_j^{k,\ell-1} r_<^{2(k-j)} r_>^{2(j+1)} - \sum_{j=0}^{k} \frac{2(\ell+1)}{2\ell+3} c_j^{k,\ell+1} r_<^{2(k-j+1)} r_>^{2j} \Bigg]\end{aligned}$$

and equating powers of $r_<$ and $r_>$ with (7.27) gives the result

$$c_j^{k+1,\ell} = c_j^{k,\ell} + c_{j-1}^{k,\ell} - \frac{2\ell}{2\ell-1} c_{j-1}^{k,\ell-1} - \frac{2(\ell+1)}{2\ell+3} c_j^{k,\ell+1}. \qquad (7.33)$$

□

We now wish to show that $c_j^{k,\ell}$ is of the required form:

Lemma 7.6.3.

$$c_j^{k,\ell} = \frac{a_j^k}{\prod_{m=1}^{k-j} (2\ell+2m+1) \prod_{n=1}^{j} (2\ell-2n+1)} =: \frac{a_j^k}{P_j^k(\ell)} \qquad (7.34)$$

where

$$a_0^k = \prod_{i=1}^{k}(2k-2i+1), \tag{7.35}$$

and

$$a_j^k = (-1)^j \binom{k}{j} a_0^k,\ j = 1, \dots, k, \tag{7.36}$$

with the convention throughout that an empty product takes the value 1.

Proof of Lemma 7.6.3 We start with the case $j = 0$, $k = 0$ which is clearly of the correct form as $a_0^0 = 1$. Assuming (7.34) and substituting into (7.33) gives

$$\frac{a_j^{k+1}}{P_j^{k+1}(\ell)} = \frac{a_j^k}{P_j^k(\ell)} + \frac{a_{j-1}^k}{P_{j-1}^k(\ell)} - \frac{2\ell}{2\ell-1}\frac{a_{j-1}^k}{P_{j-1}^k(\ell-1)} - \frac{2(\ell+1)}{2\ell+3}\frac{a_j^k}{P_j^k(\ell+1)}.$$

Which, multiplying by $P_j^{k+1}(\ell)$, gives

$$\begin{aligned} a_j^{k+1} &= a_j^k \left[(2\ell+2(k+1-j)+1) - \frac{2(\ell+1)}{2\ell+3}(2\ell+3)\frac{(2\ell-2j+1)}{(2\ell+1)}\right] \\ &\quad + a_{j-1}^k \left[(2\ell-2j+1) - \frac{2\ell}{2\ell-1}\frac{(2\ell+2(k-(j-1)+1)}{(2\ell+1)}(2\ell-1)\right] \\ &= \frac{4k\ell+2\ell+2k+2j+1}{2\ell+1}a_j^k + \frac{-4k\ell-2\ell-2j+1}{2\ell+1}a_{j-1}^k. \end{aligned} \tag{7.37}$$

For $j = 0$, using (7.35) for k and that $a_{-1}^k = 0$, this gives

$$\begin{aligned} a_0^{k+1} &= \frac{4k\ell+2\ell+2k+1}{2\ell+1}a_j^k = (2k+1)a_0^k = (2k+1)\prod_{i=1}^{k}(2k-2i+1) \\ &= \prod_{i=1}^{k+1}(2(k+1)-2i+1), \end{aligned}$$

showing that (7.35) holds for $k+1$.

Noting now that (7.36) implies, for $1 \le j \le k$,

$$a_{j-1}^k = -\frac{j}{k-j+1}a_j^k \tag{7.38}$$

and hence, for $1 \le j \le k$, inserting (7.38) into (7.37) gives

$$a_j^{k+1} = \frac{(k+1)(2\ell+4k\ell+2k+1)}{(2\ell+1)(k-j+1)}a_j^k = \frac{(k+1)(1+2k)}{(k-j+1)}a_j^k. \tag{7.39}$$

It remains to show (7.36) holds for a_j^{k+1}, which is equivalent to

$$a_{j-1}^{k+1} = -\frac{j}{k-j+2} a_j^{k+1}, \ 1 \le j \le k+1. \tag{7.40}$$

Using (7.39) for a_{j-1}^{k+1} and applying (7.38) we get

$$a_{j-1}^{k+1} = \frac{(k+1)(1+2k)}{(k-j+2)} a_{j-1}^k = -\frac{(k+1)(1+2k)}{(k-j+2)} \frac{j}{(k-j+1)} a_j^k,$$

which holds for $1 \le j \le k$. Using (7.39) gives (7.40), showing that (7.36) holds for $k+1$, $j = 1, \dots, k$.

The final case is $j = k+1$. By (7.37) and the fact that $a_{k+1}^k = 0$ we have

$$a_{k+1}^{k+1} = a_k^k \left(\frac{-4k\ell - 2k - 2(k+1) + 1}{2\ell + 1} \right) = -(2k+1) a_k^k.$$

Using (7.36) for $j = k$ and (7.35) for $k+1$ this gives

$$a_{k+1}^{k+1} = (-1)^{k+1}(2k+1) a_0^k = (-1)^{k+1} a_0^{k+1},$$

and since $\binom{k+1}{k+1} = 1$ this gives the result. ☐

Combining Lemmas 7.6.2 and 7.6.3 proves Lemma 7.6.1. ☐

Lemma 7.6.1 appears to be entirely new and may well be useful in other applications involving partial wave expansions.

7.6.2 Decay of $\xi_\ell^j(r_<, r_>)$

We now wish to use the explicit form of $f_\ell^k(r_<, r_>)$ given in Lemma 7.6.1 along with the expansion

$$\xi_\ell^j(r_<, r_>) = (\ell + \tfrac{1}{2})^{-1/2} \frac{1}{j!} \sum_{k=0}^{\lfloor (j+1)/2 \rfloor} \binom{j}{2k-1} (-1)^{j-2k+1} (r_> - r_<)^{j-2k+1} f_\ell^k(r_<, r_>) \tag{7.41}$$

derived at the start of this section to derive the decay rate of the $\xi_\ell^j(r_<, r_>)$ in terms of the decay rates of the $f_\ell^k(r_<, r_>)$. Following the ideas of Hill, we write

$$(r_> - r_<)^{j-2k+1} = r_>^{j-2k+1} \left(1 - \tfrac{r_<}{r_>}\right)^{j-2k+1}$$

and hence, up to a constant, (7.26) shows that the term in the summand of (7.41) may be written as

$$r_>^{j-2k}\Big(1-\tfrac{r_<}{r_>}\Big)^{j-2k+1}\Big(\tfrac{r_<}{r_>}\Big)^{\ell}\sum_{i=0}^{k}\frac{a_i^k r_<^{2(k-i)}r_>^{2i}}{P_i^k(\ell)}.$$

We now investigate a general function of the form $h_{p,n}(x)=(1-x)^p x^n$, $x\in[0,1]$, $n,p\in\mathbb{R}\backslash\{0\}$, where x will correspond to $r_</r_>$, and in particular we would like to find a bound for $h_{p,n}(x)$ in terms of p and n. It is clear that

$$\frac{d}{dx}h_{p,n}(x)=n(1-x)^p x^{n-1}-p(1-x)^{p-1}x^n,$$

from which it follows that $h_{p,n}(x)$ is maximized at $x=\frac{n}{n+p}$. Noting that this value is indeed in the allowed range $x\in[0,1]$, we find that the maximum value of $h_{p,n}$ can be bounded from above by $(\frac{p}{n})^p$.

In our case, we see that $p=j-2k+1$ and that we may take $n=\ell/m$ with $m>0$. This gives that

$$\begin{aligned}(r_>-r_<)^{j-2k+1}\Big(\tfrac{r_<}{r_>}\Big)^{\ell/m}&=r_>^{j-2k+1}\Big(1-\tfrac{r_<}{r_>}\Big)^{j-2k+1}\Big(\tfrac{r_<}{r_>}\Big)^{\ell/m}\\&\le r_>^{j-2k+1}\big(m(j-2k+1)\big)^{j-2k+1}\ell^{-j+2k-1},\end{aligned}\tag{7.42}$$

which shows that the decay rate in ℓ is independent of the choice of m. Inserting this bound into (7.41), and using that $1/(\ell+1/2)<1/\ell$, we see that

$$\begin{aligned}|\xi_\ell^j(r_<,r_>)|^2\le{}&\ell^{-2j-3}\Big(\frac{1}{j!}\Big)^2\sum_{k=0}^{\lfloor(j+1)/2\rfloor}\sum_{k'=0}^{\lfloor(j+1)/2\rfloor}\ell^{2(k+k')}\binom{j}{2k-1}\binom{j}{2k'-1}\\&\times r_>^{2j-2(k+k')+2}\big(m(j-2k+1)\big)^{j-2k+1}\big(m(j-2k'+1)\big)^{j-2k'+1}\\&\times\Big|\Big(\tfrac{r_<}{r_>}\Big)^{-2\ell/m}f_\ell^k(r_<,r_>)f_\ell^{k'}(r_<,r_>)\Big|.\end{aligned}\tag{7.43}$$

Similarly, we need an expression for $|\frac{\partial}{\partial r_\alpha}\xi_\ell^j(r_<,r_>)|^2$ and differentiating (7.41) we see that

$$\begin{aligned}&\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<,r_>)\\&=\pm(\ell+\tfrac12)^{-1/2}\frac{1}{j!}\sum_{k=0}^{\lfloor(j+1)/2\rfloor}\binom{j}{2k-1}(-1)^{j-2k+1}\Big[(r_>-r_<)^{j-2k+1}\tfrac{\partial}{\partial r_\gamma}f_\ell^k(r_<,r_>)\\&\qquad+(j-2k+1)(r_>-r_<)^{j-2k}f_\ell^k(r_<,r_>)\Big].\end{aligned}$$

Using (7.42) and the analogous version for lower powers of $(r_> - r_<)$ gives

$$
\begin{aligned}
&|\tfrac{\partial}{\partial r_\gamma}\xi^j_\ell(r_<, r_>)|^2 \leq \\
&\quad \ell^{-2j-1}\Big(\frac{1}{j!}\Big)^2 \sum_{k=0}^{\lfloor (j+1)/2\rfloor} \sum_{k'=0}^{\lfloor (j+1)/2\rfloor} \ell^{2(k+k')} \binom{j}{2k-1}\binom{j}{2k'-1} r_>^{2j-2(k+k')} \\
&\quad \times \Big[(j-2k+1)(j-2k'+1)\big(m(j-2k)\big)^{j-2k}\big(m(j-2k')\big)^{j-2k'} \\
&\qquad\qquad \times \big|\big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} f^k_\ell(r_<, r_>) f^{k'}_\ell(r_<, r_>)\big| \\
&\qquad + 2\ell^{-1} r_>(j-2k+1)\big(m(j-2k)\big)^{j-2k}\big(m(j-2k'+1)\big)^{j-2k'+1} \\
&\qquad\qquad \times \big|\big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} f^k_\ell(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f^{k'}_\ell(r_<, r_>)\big| \\
&\qquad + \ell^{-2} r_>^2 \big(m(j-2k+1)\big)^{j-2k+1}\big(m(j-2k'+1)\big)^{j-2k'+1} \\
&\qquad\qquad \times \big|\big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} \tfrac{\partial}{\partial r_\gamma} f^k_\ell(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f^{k'}_\ell(r_<, r_>)\big|\Big]. \qquad (7.44)
\end{aligned}
$$

It remains to determine the decay rate of the $f^k_\ell(r_<, r_>)$ and their derivatives. We have an explicit form for $f^k_\ell(r_<, r_>)$, given in (7.26) and the following lemma gives an explicit form for $\frac{\partial}{\partial r_\gamma} f^k_\ell(r_<, r_>)$:

Lemma 7.6.4. *The derivatives of $f^k_l(r_<, r_>)$ are given by*

$$
\frac{\partial}{\partial r_<} f^k_\ell(r_<, r_>) = r_<^{\ell-1} r_>^{-\ell-1} \sum_{j=0}^{k} \frac{a^k_j r_<^{2(k-j)} r_>^{2j} (\ell + 2(k-j))}{\prod_{m=1}^{k-j}(2\ell+2m+1) \prod_{n=1}^{j}(2\ell-2n+1)} \qquad (7.45)
$$

$$
\frac{\partial}{\partial r_>} f^k_\ell(r_<, r_>) = r_<^{\ell} r_>^{-\ell-2} \sum_{j=0}^{k} \frac{a^k_j r_<^{2(k-j)} r_>^{2j} (-\ell - 1 + 2j)}{\prod_{m=1}^{k-j}(2\ell+2m+1) \prod_{n=1}^{j}(2\ell-2n+1)} \qquad (7.46)
$$

Proof This follows trivially from differentiating (7.26) term by term. ☐

The remaining step is to determine the decay rates of integrals of products of two of $f^k_l(r_<, r_>)$ and $\frac{\partial}{\partial r_\gamma} f^k_l(r_<, r_>)$ with a general function of $r_<$ and $r_>$, which will later to be chosen to be products of the $A_j(r_<, r_>)$ or their derivatives. Again by Cauchy-Schwarz it is necessary only to consider integrals of the products

$$
r_<^n r_>^m g(r_<, r_>)(r_</r_>)^{-2\ell/m} f^k_\ell(r_<, r_>) f^{k'}_\ell(r_<, r_>)
$$

and

$$
r_<^n r_>^m g(r_<, r_>)(r_</r_>)^{-2\ell/m} \tfrac{\partial}{\partial r_\gamma} f^k_\ell(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f^{k'}_\ell(r_<, r_>).
$$

7.7 Results on the Integrals

Following the discussion in the previous section we need to derive the rate of decay of integrals of the form

$$\int_0^\infty \int_0^{r_>} r_<^\alpha r_>^\beta g(r_<, r_>) \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< dr_> \tag{7.47}$$

and

$$\int_0^\infty \int_0^{r_>} r_<^\alpha r_>^\beta g(r_<, r_>) \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |\tfrac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f_\ell^{k'}(r_<, r_>)| dr_< dr_>, \tag{7.48}$$

where $\alpha, \beta \in \mathbb{N} \cup \{0\}$ and $r_\gamma \in \{r_<, r_>\}$.

The evaluation of these integrals is obviously highly dependent on the form of $g(r_<, r_>)$ and, whilst we have an explicit form for $f_\ell^k(r_<, r_>)$, we do not have an explicit form for $g(r_<, r_>)$. We begin by noting the following result on $f_\ell^k(r_<, r_>)$ and $\frac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>)$

Lemma 7.7.1. *For* $m > 1$ *and* $\ell \gg k$*, the functions* $(r_</r_>)^{-\ell/m} f_\ell^k(r_<, r_>)$ *and* $(r_</r_>)^{-\ell/m} \frac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>)$ *for* $r_\gamma \in \{r_<, r_>\}$ *are strongly peaked on the line* $r_< = r_>$*.*

Proof From (7.26), (7.45) and (7.46) we see that a general term in $r_<$ and $r_>$ is of the form $r_<^{\tilde{\ell}+a}/r_>^{\tilde{\ell}+b}$ where $a, b \in \{-2, \dots, 2k\}$ and $\tilde{\ell} = (m-1)\ell/m > 0$. Rewriting this and using that $r_< = r_> - x$ where $x \in [0, r_>]$ we have that

$$\frac{r_<^{\tilde{\ell}+a}}{r_>^{\tilde{\ell}+b}} = r_>^{a-b} \Big(\frac{r_<}{r_>}\Big)^{\tilde{\ell}+a} = r_>^{a-b} \Big(1 - \frac{x}{r_>}\Big)^{\tilde{\ell}+b}.$$

For $\ell \gg k$ we have $\tilde{\ell} + b \gg 1$ and this expression is strongly peaked when $x = 0$, giving the result. □

Although stated in [Hil85] (page 1187), this result was not proved.

For suitably smooth $g(r_<, r_>)$, this result implies that it is sensible to to Taylor expand $g(r_<, r_>)$ around $r_< = r_>$ giving

$$g(r_<, r_>) = g(r_>, r_>) + (r_< - r_>) g'(r_>, r_>) + (r_< - r_>)^2 g''(r_>, r_>) + \dots,$$

where $g'(x, r_>) := \frac{\partial}{\partial r_<} g(r_<, r_>)|_{r_<=x}$ and so on. This result is made rigorous by the next lemma:

Lemma 7.7.2. *Let $g(\,\cdot\,, r_>) \in C^N(\mathbb{R})$ be such that $|g^{(n)}(s, r_>)| \leq \tilde{g}(r_>)$ for all $s \in \mathbb{R}_+$ and for $n = 0, \dots, N$. Then, for $M < N$, $p > 0$, and $f(r_<, r_>)$ an arbitrary function,*

$$\begin{aligned}\int_0^{r_>} & g(r_<, r_>)|f(r_<, r_>)|dr_< \\ &= \sum_{m=0}^{M} g^{(m)}(r_>, r_>) \int_0^{r_>} \frac{(r_< - r_>)^m}{m!}|f(r_<, r_>)|dr_< \\ &\qquad + \sum_{n=M+1}^{N} \mathcal{O}(\ell^{-n})\tilde{g}(r_>) \int_0^{r_>} r_<^n \Big(\frac{r_<}{r_>}\Big)^{-\ell/p}|f(r_<, r_>)|dr_<\end{aligned} \tag{7.49}$$

Proof We begin by noting that we may Taylor expand $g(r_<, r_>)$ in $r_<$ about $r_>$ and, since $g^{(N)}$ is continuous, we may use the Lagrange form of the remainder, i.e.

$$g(r_<, r_>) = \sum_{n=0}^{N-1} \frac{(r_< - r_>)^n}{n!} g^{(n)}(r_>, r_>) + \frac{(r_< - r_>)^N}{N!} g^{(N)}(s, r_>)$$

for some point $s \in (r_<, r_>)$. It is clear that the first M terms of this expansion contribute the terms in the sum in (7.49). We now note that, in a similar way to the derivation of (7.42),

$$(r_< - r_>)^n \left(\frac{r_<}{r_>}\right)^{\ell/p} \leq c r_>^n \ell^{-n},$$

for any $p > 0$ and for some $c > 0$ (which depends on p). Hence, since $|g^{(n)}(r, r_>)| \leq \tilde{g}(r_>)$ for all r, the result follows. □

This result replaces the non-rigorous argument of [Hil85].

Applying this result to the integrals (7.47) and (7.48), we find the general integrals are now of the form

$$\int_0^\infty \int_0^{r_>} r_<^\alpha r_>^\beta g(r_>) \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< dr_> \tag{7.50}$$

and

$$\int_0^\infty \int_0^{r_>} r_<^\alpha r_>^\beta g(r_>) \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |\tfrac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f_\ell^{k'}(r_<, r_>)| dr_< dr_>. \tag{7.51}$$

The advantage of this form of the integrals is that, since the unknown function g is no longer dependent on $r_<$, we may explicitly evaluate the $r_<$ integral. We now move on to discuss the decay rates of integrals of the type (7.50) and (7.51).

7.7.1 Decay rates of the integrals

From Hill's result we expect the norms (7.8) and (7.9) to decay at least as $\mathcal{O}(l^{-4})$. In fact, since the norm in (7.8) is the L^2 norm and that in (7.9) is a higher Sobolev norm, we expect the slowest decay to come from the energy norm. Recall that we have factors of $\ell^{-2j-3+2(k+k')}$, $\ell^{-2j-1+2(k+k')}$ and $\ell^{-2j-3+2(k+k')}$ for terms involving $f_\ell^k(r_<,r_>)f_\ell^{k'}(r_<,r_>)$ from (7.43), $f_\ell^k(r_<,r_>)f_\ell^{k'}(r_<,r_>)$ from (7.44) and $\frac{\partial}{\partial r_\alpha}f_\ell^k(r_<,r_>)\frac{\partial}{\partial r_\alpha}f_\ell^{k'}(r_<,r_>)$ from (7.44) respectively.

The main aim of this section is to remove the dependence on k and k' from the decay rates of $|\xi_\ell^j(r_<,r_>)|^2$ and $|\frac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<,r_>)|$. We begin by finding the k-dependence of integrals of the forms (7.50) and (7.51).

Lemma 7.7.3. *For $k,k'\geq 1$, $\alpha\in\mathbb{N}\cup\{0\}$, $r_\gamma\in\{r_<,r_>\}$, and $m>1$, let $\tilde{\ell}:=(m-1)\ell/m$ then*

$$\int_0^{r_>} r_<^\alpha\left(\tfrac{r_<}{r_>}\right)^{-2\ell/m}|f_\ell^k(r_<,r_>)f_\ell^{k'}(r_<,r_>)|dr_< = \frac{r_>^{2(k+k')+\alpha-1}}{P^k(\ell)P^{k'}(\ell)Q^{k,k'}(\tilde{\ell};\alpha)}\sum_{j=0}^{k}\sum_{j'=0}^{k'}a_j^k a_{j'}^{k'}R_{j,j'}^{k,k'}(\ell,\tilde{\ell};\alpha), \tag{7.52}$$

$$\begin{aligned}\int_0^{r_>} r_<^\alpha\left(\tfrac{r_<}{r_>}\right)^{-2\ell/m}|\tfrac{\partial}{\partial r_<}f_\ell^k(r_<,r_>)\tfrac{\partial}{\partial r_<}f_\ell^{k'}(r_<,r_>)|dr_< = \\ \frac{r_>^{2(k+k')+\alpha-3}}{P^k(\ell)P^{k'}(\ell)Q^{k,k'}(\tilde{\ell}-1;\alpha)} \\ \times\sum_{j=0}^{k}\sum_{j'=0}^{k'}a_j^k a_{j'}^{k'}R_{j,j'}^{k,k'}(\ell,\tilde{\ell}-1;\alpha)(\ell+2(k-j))(\ell+2(k'-j')),\end{aligned} \tag{7.53}$$

and

$$\int_0^{r_>} r_<^\alpha\left(\tfrac{r_<}{r_>}\right)^{-2\ell/m}|\tfrac{\partial}{\partial r_>}f_\ell^k(r_<,r_>)\tfrac{\partial}{\partial r_>}f_\ell^{k'}(r_<,r_>)|dr_< = \frac{r_>^{2(k+k')+\alpha-3}}{P^k(\ell)P^{k'}(\ell)Q^{k,k'}(\tilde{\ell};\alpha)}\sum_{j=0}^{k}\sum_{j'=0}^{k'}a_j^k a_{j'}^{k'}R_{j,j'}^{k,k'}(\ell,\tilde{\ell};\alpha)(\ell-2j)(\ell-2j') \tag{7.54}$$

where $P^k(\ell)$ is a polynomial in ℓ of degree $2k$, given explicitly in (7.55), $Q^{k,k'}(\tilde{\ell};\alpha)$ is a polynomial in $\tilde{\ell}$ of degree $k+k'+1$ with coefficients parameterized by α and is given explicitly in (7.57). The remaining term $R_{j,j'}^{k,k'}(\ell,\tilde{\ell};\alpha)$ is the product of a *polynomial in ℓ with a polynomial in $\tilde{\ell}$ (both of degree $k+k'$), with coefficients parameterized by α and is given explicitly in (7.58).*

Proof We prove the result for $f_\ell^k(r_<, r_>)$, the proofs for $\frac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>)$ are analogous. By (7.26) we have that

$$\left(\tfrac{r_<}{r_>}\right)^{-2\ell/m} f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>) = \frac{r_<^{2\tilde{\ell}}}{r_>^{2\tilde{\ell}+2}} \sum_{j=0}^{k} \sum_{j'=0}^{k'} \frac{a_j^k a_{j'}^{k'} r_<^{2(k+k'-j-j')} r_>^{2(j+j')}}{P_j^k(\ell) P_{j'}^{k'}(\ell)}$$

and hence

$$\begin{aligned}
&\int_0^{r_>} r_<^\alpha \left(\tfrac{r_<}{r_>}\right)^{-2\ell/m} f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>) dr_< \\
&= \int_0^{r_>} r_<^{2\tilde{\ell}+\alpha} r_>^{-2\tilde{\ell}-2} \sum_{j=0}^{k} \sum_{j'=0}^{k'} \frac{a_j^k a_{j'}^{k'} r_<^{2(k+k'-j-j')} r_>^{2(j+j')}}{P_j^k(\ell) P_{j'}^{k'}(\ell)} dr_< \\
&= r_>^{2(k+k')+\alpha-1} \sum_{j=0}^{k} \sum_{j'=0}^{k'} \frac{a_j^k a_{j'}^{k'}}{P_j^k(\ell) P_{j'}^{k'}(\ell) (2\tilde{\ell} + 2(k+k'-j-j') + \alpha + 1)}.
\end{aligned}$$

We now define the polynomial

$$P^k(\ell) := \prod_{m=1}^{k} (2\ell + 2m + 1) \prod_{n=1}^{k} (2\ell - 2n + 1), \tag{7.55}$$

which is the product of all distinct terms in the denominators of (7.26) over all j. We also note that the terms given by the integration are

$$\{(2\tilde{\ell} + 2(k+k'-j-j') + \alpha + 1)\}_{j+j'=0}^{k+k'} = \{(2\tilde{\ell} + 2i + \alpha - 1)\}_{i=1}^{k+k'+1}, \tag{7.56}$$

and we therefore define

$$Q^{k,k'}(\tilde{\ell}; \alpha) = \prod_{i=1}^{k+k'+1} (2\tilde{\ell} + 2i + \alpha - 1). \tag{7.57}$$

It follows that

$$\begin{aligned}
&\int_0^{r_>} r_<^\alpha \left(\tfrac{r_<}{r_>}\right)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< = \\
&\qquad \frac{r_>^{2(k+k')+\alpha-1}}{P^k(\ell) P^{k'}(\ell) Q^{k,k'}(\tilde{\ell}; \alpha)} \Big| \sum_{j=0}^{k} \sum_{j'=0}^{k'} a_j^k a_{j'}^{k'} R_{j,j'}^{k,k'}(\ell, \tilde{\ell}; \alpha) \Big|
\end{aligned}$$

where

$$\begin{aligned} R^{k,k'}_{j,j'}(\ell,\tilde{\ell};\alpha) = &\prod_{m=k-j+1}^{k}(2\ell+2m+1)\prod_{n=j+1}^{k}(2\ell-2n+1) \\ &\times \prod_{m'=k'-j'+1}^{k'}(2\ell+2m'+1)\prod_{n'=j'+1}^{k'}(2\ell-2n'+1) \\ &\times \prod_{\substack{i=1 \\ i\neq(k+k'-j-j'+1)}}^{k+k'+1}(2\tilde{\ell}+2i+\alpha-1). \end{aligned} \tag{7.58}$$

For $\frac{\partial}{\partial r_>}f^k_\ell(r_<,r_>)$ the power of $r_<$ is unchanged, leading to the same coefficients from the integration but a reduced power of $r_>$. The extra two terms in ℓ come from the initial differentiation as given in (7.46). For $\frac{\partial}{\partial r_<}f^k_\ell(r_<,r_>)$ the initial power of $r_<$ is reduced by two, and we once again gain two factors from the initial differentiation as shown in (7.45). □

The results of Lemma 7.7.3 show (noting that $\mathcal{O}(\tilde{\ell}) = \mathcal{O}(\ell)$) that the integral (7.52) is of order $\ell^{-(k+k')-1}$ and the integrals (7.53) and (7.54) are of order $\ell^{-(k+k')+1}$. However, this is not strong enough to counter the $\ell^{2(k+k')}$ in the summands of (7.43) and (7.44). We would therefore like to show that the summation over j and j' in (7.52)-(7.54) causes an increase in order.

The first step is to note that $R^{k,k'}_{j,j'}(\ell,\tilde{\ell};\alpha)$ and the $(\ell+\beta)$ terms in these sums contribute powers of j and j'. The terms in (7.56), i.e those in the final product of $R^{k,k'}_{j,j'}(\ell,\tilde{\ell};\alpha)$ in (7.58), may be written in the form $(2\tilde{\ell}+\gamma_i-(j+j'))$ where the γ_i depend on k, k' and α but are independent of j, j' and ℓ. There are $k+k'$ of these terms and this product may be expanded as

$$\prod_{\substack{i=1 \\ i\neq(k+k'-j-j'+1)}}^{k+k'+1}(2\tilde{\ell}+2i+\alpha-1) = \sum_{n=0}^{k+k'} c_n\tilde{\ell}^n(j+j')^{k+k'-n} = \sum_{n=0}^{k+k'}\tilde{\ell}^n\sum_{m=0}^{k+k'-n} d_{n,m}j^m j'^{k+k'-n-m}$$

where c_n and $d_{n,m}$ depend on k,k' and α. Similarly, we see that the extra two terms in (7.53) and (7.54) lead to sums of the form

$$\sum_{n=0}^{k+k'} c_n\tilde{\ell}^n(j+j')^{k+k'-n}j^x j'^y\ell^{2-x-y} = \sum_{n=0}^{k+k'}\tilde{\ell}^n\ell^{2-x-y}\sum_{m=0}^{k+k'-n} d_{n,m}j^{m+x}j'^{k+k'-n+y}$$

where $x,y\in\{0,1\}$.

From these expansions and the explicit form of $R^{k,k'}_{j,j'}(\ell,\tilde{\ell};\alpha)$ given in (7.58) we see that, for fixed n and m, the double sum over j and j' in (7.52) is, up to a constant, of the form

$$\sum_{j=0}^{k}\sum_{j'=0}^{k'} a_j^k a_{j'}^{k'} \tilde{\ell}^n j^m j'^{k+k'-n-m} \prod_{\mu=k-j+1}^{k}(2\ell+2\mu+1)\prod_{\nu=j+1}^{k}(2\ell-2\nu+1)$$
$$\times \prod_{\mu'=k'-j'+1}^{k'}(2\ell+2\mu'+1)\prod_{\nu'=j'+1}^{k'}(2\ell-2\nu'+1). \tag{7.59}$$

This double sum separates into the product of two sums:

$$\tilde{\ell}^n \sum_{j=0}^{k} a_j^k j^m \prod_{\mu=k-j+1}^{k}(2\ell+2\mu+1)\prod_{\nu=j+1}^{k}(2\ell-2\nu+1) \tag{7.60}$$
$$\times \sum_{j'=0}^{k'} a_{j'}^{k'} j'^{k+k'-n-m} \prod_{\mu'=k'-j'+1}^{k'}(2\ell+2\mu'+1)\prod_{\nu'=j'+1}^{k'}(2\ell-2\nu'+1).$$

The sums in (7.53) and (7.54) are of the same form but the sum over j contains either an extra power of ℓ or and extra power of j and analogously for j'. We now wish to show:

Lemma 7.7.4. *For $a_j^k = (-1)^j\binom{k}{j}a_0^k$, and $p \in \mathbb{N}\cup\{0\}$, the sum*

$$S_{k,p} := \sum_{j=0}^{k} a_j^k j^p \prod_{m=k-j+1}^{k}(2\ell+2m+1)\prod_{n=j+1}^{k}(2\ell-2n+1)$$

is $\mathcal{O}(\ell^p)$.

Proof We first note that if $p \geq k$ then the result is trivial as the highest order of ℓ possible is k. Hence we need only consider $0 \leq p < k$ and begin by dividing through by 2^k and rearranging to give

$$cS_{k,p} = \sum_{j=0}^{k} a_j^k j^p \prod_{m=k-j+1}^{k}\big((\ell+\tfrac{1}{2})+m\big)\prod_{n=j+1}^{k}\big((\ell+\tfrac{1}{2})-n\big),$$

where c is some constant. For ease of notation we let $\alpha := (\ell+\frac{1}{2})$, giving

$$cS_{k,p} = \sum_{j=0}^{k}(-1)^j\binom{k}{j} j^p \prod_{m=k-j+1}^{k}(\alpha+m)\prod_{n=j+1}^{k}(\alpha-n).$$

Recall the Gamma functions [AS72],

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt,$$

which are the generalized factorial function and satisfy

$$\Gamma(x) = (x-1)\Gamma(x-1), \; x \in \mathbb{R}. \tag{7.61}$$

This gives that

$$\prod_{m=k-j+1}^{k} (\alpha + m) = \frac{\Gamma(\alpha + k + 1)}{\Gamma(\alpha + k + 1 - j)},$$

$$\prod_{n=j+1}^{k} (\alpha - n) = (-1)^{k-j} \frac{\Gamma(k + 1 - \alpha)}{\Gamma(j + 1 - \alpha)},$$

and the standard form of the binomial coefficient is

$$\binom{k}{j} = \frac{\Gamma(k+1)}{\Gamma(j+1)\Gamma(k-j+1)}. \tag{7.62}$$

It follows that

$$cS_{k,p} = (-1)^k \sum_{j=0}^{k} j^p \frac{\Gamma(k+1)}{\Gamma(j+1)\Gamma(k-j)} \frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-j)} \frac{\Gamma(k+1-\alpha)}{\Gamma(j+1-\alpha)}.$$

We will prove the cases $0 \leq p < k$ by induction on p and begin with the case $p = 0$. Using the identity

$$\frac{\Gamma(z)}{\Gamma(w)} = \Gamma(z-w+1) \frac{\Gamma(z)}{\Gamma(z-w+1)\Gamma(w)} = (z-w)! \binom{z-1}{z-w} \tag{7.63}$$

and the Vandermonde identity [AS72] for the sum of binomial coefficients

$$\sum_{j=0}^{k} \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k} \tag{7.64}$$

gives

$$\frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-j)} = j! \binom{\alpha+k}{j}$$

and

$$\frac{\Gamma(k+1-\alpha)}{\Gamma(j+1-\alpha)} = (k-j)! \binom{k-\alpha}{k-j}.$$

Hence

$$\begin{aligned} cS_{k,0} &= (-1)^k \Gamma(k+1) \sum_{j=0}^{k} \binom{\alpha+k}{j} \binom{k-\alpha}{k-j} \\ &= (-1)^k \Gamma(k+1) \binom{2k}{k} = (-1)^k \Gamma(k+1) \frac{\Gamma(2k+1)}{\Gamma(k+1)\Gamma(k+1)} \\ &= (-1)^k \frac{\Gamma(2k+1)}{\Gamma(k+1)} = \mathcal{O}(1), \end{aligned}$$

which shows the result holds for $p = 0$. For $0 < p < k$ we consider the sum

$$\begin{aligned} T_{k,p} := (-1)^k \sum_{j=0}^{k} j(j-1)\ldots(j-(p-1)) &\frac{\Gamma(k+1)}{\Gamma(j+1)\Gamma(k-j+1)} \\ \times &\frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-j)} \frac{\Gamma(k+1-\alpha)}{\Gamma(j+1-\alpha)} \end{aligned}$$

which is clearly zero when $j \in \{0, \ldots, (p-1)\}$ and thus

$$\begin{aligned} &T_{k,p} \\ &= (-1)^k \sum_{j=p}^{k} \frac{\Gamma(j+1)}{\Gamma(j-p+1)} \frac{\Gamma(k+1)}{\Gamma(j+1)\Gamma(k-j+1)} \frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-j)} \frac{\Gamma(k+1-\alpha)}{\Gamma(j+1-\alpha)} \\ &= (-1)^k \sum_{j=0}^{k-p} \frac{\Gamma(k+1)}{\Gamma(j+1)\Gamma(k-j+1-p)} \frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-j-p)} \frac{\Gamma(k+1-\alpha)}{\Gamma(j+p+1-\alpha)} \\ &= (-1)^k \sum_{j=0}^{k-p} \frac{\Gamma(k+1)\Gamma(\alpha+k+1)}{\Gamma(\alpha+k-p+1))} \frac{\Gamma(\alpha+k-p+1)}{\Gamma(j+1)\Gamma(\alpha+k-p-j+1)} \\ &\qquad \times \frac{\Gamma(k-\alpha+1)}{\Gamma(k-j-p+1)\Gamma(p+j-\alpha+1)} \end{aligned}$$

We now apply (7.62) to the last two terms to obtain

$$T_{k,p} = (-1)^k \Gamma(k+1) \frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k+1-p)} \sum_{j=0}^{k-p} \binom{\alpha+k-p}{j} \binom{k-\alpha}{k-p-j},$$

and using (7.64) gives

$$T_{k,p} = (-1)^k \Gamma(k+1) \frac{\Gamma(\alpha+k+1)}{\Gamma(\alpha+k-p+1)} \binom{2k-p}{k-p}.$$

It is clear from (7.61) that $\Gamma(\alpha+k+1)/\Gamma(\alpha+k-p+1)$ is $\mathcal{O}(\alpha^p)$ which is equivalent to $\mathcal{O}(\ell^p)$. Finally noting that $j(j-1)\dots(j-(p-1)) = j^p + \sum_{i=0}^{p-1} c_i j^i$ and using the inductive hypothesis gives the result. ☐

We are now in a position to determine the decay rates of the integrals (7.52)-(7.54):

Lemma 7.7.5. *Let* $m > 1$ *then*

$$\int_0^{r_>} r_<^\alpha \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< = r_>^{2(k+k')+\alpha-1} \mathcal{O}(\ell^{-2(k+k')-1})$$

$$\int_0^{r_>} r_<^\alpha \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |\tfrac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f_\ell^{k'}(r_<, r_>)| dr_< = r_>^{2(k+k')+\alpha-3} \mathcal{O}(\ell^{-2(k+k')+1})$$

where $r_\gamma \in \{r_<, r_>\}$.

Proof Lemma 7.7.4 and (7.60) shows that the sums in (7.59) are overall of order $\mathcal{O}(\ell^{k+k})$ and the analogous sums for the derivative cases given in (7.53) and (7.54) are of the form $\mathcal{O}(\ell^{k+k'+2})$. Combining this order with that of the polynomials outside the sum in (7.52)-(7.54), which are of order $\ell^{-3(k+k')-1}$, gives the result for the order. The result for the power of $r_>$ follow directly from the results of Lemma 7.7.3. ☐

Recall from Section 7.6.2 that there exist bounds on $|\xi_\ell^j(r_<, r_>)|^2$ and $|\frac{\partial}{\partial r_\gamma} \xi_\ell^j(r_<, r_>)|^2$ (given explicitly by (7.43) and (7.44)) such that

$$\begin{aligned}
&\iint r_<^\alpha r_<^\beta g(r_>) |\xi_\ell^j(r_<, r_>)|^2 dr_< dr_> \\
&\leq \ell^{-2j-3} \sum_{k=0}^{\lfloor (j+1)/2 \rfloor} \sum_{k'=0}^{\lfloor (j+1)/2 \rfloor} \ell^{2(k+k')} c_{j,k,k',m}^{(1)} \\
&\qquad \times \iint r_<^\alpha r_>^{2j-2(k+k')+\beta+2} g(r_>) \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< dr_>
\end{aligned}$$

and, using Cauchy-Schwarz to bound the middle term of (7.44),

$$\begin{aligned}
&\iint r_<^\alpha r_>^\beta g(r_>) |\xi_\ell^j(r_<, r_>)|^2 dr_< dr_> \leq \ell^{-2j-1} \sum_{k=0}^{\lfloor (j+1)/2 \rfloor} \sum_{k'=0}^{\lfloor (j+1)/2 \rfloor} \ell^{2(k+k')} \\
&\quad \times \Big[c_{j,k,k',m}^{(2)} \iint r_<^\alpha r_>^{2j-2(k+k')+\beta} \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} |f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)| dr_< dr_> \\
&\qquad + \ell^{-2} c_{j,k,k',m}^{(3)} \iint r_<^\alpha r_>^{2j-2(k+k')+2+\beta} \big(\tfrac{r_<}{r_>}\big)^{-2\ell/m} \\
&\qquad\qquad \times |\tfrac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} f_\ell^{k'}(r_<, r_>)| dr_< dr_> \Big],
\end{aligned}$$

for some constants $c^{(i)}_{j,k,k',m}$, $i = 1, 2, 3$.

This holds for any $m > 1$ and the bounds in Lemma 7.7.5 are independent of the choice of m and so, inserting these decay rates into the two above expressions, and noting that $\mathcal{O}(\ell) = \mathcal{O}(\tilde{\ell})$, we find

$$\iint r_<^{\alpha} r_<^{\beta} g(r_>) |\xi_\ell^j(r_<, r_>)|^2 dr_< dr_> = \mathcal{O}(\ell^{-2j-4}) \int r_>^{2j+1+\alpha+\beta} g(r_>) dr_>, \quad (7.65)$$

$$\iint r_<^{\alpha} r_>^{\beta} g(r_>) |\tfrac{\partial}{\partial r_\gamma} \xi_\ell^j(r_<, r_>)|^2 dr_< dr_> = \mathcal{O}(\ell^{-2j-2}) \int r_>^{2j-1+\alpha+\beta} g(r_>) dr_> \quad (7.66)$$

These bounds lead to the following result about the decay rates of the norms of $\psi_\ell(r_<, r_>)$ and $\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)$ given in (7.14) and (7.15).

Corollary 7.7.6. *Let $g(\,\cdot\,, r_>) \in C^N(\mathbb{R})$ be such that $|g^{(n)}(s, r_>)| \leq \tilde{g}(r_>)$ for some $\tilde{g}(r_>)$, all $s \in \mathbb{R}$ and $n = 0, \ldots, N$. Recalling the notation*

$$\|f\|_{a,b} := \Big(\iint |f|^2 r_<^a r_>^b dr_< dr_> \Big)^{1/2}$$

we have, for any $M < N$,

$$\begin{aligned} &\|g(r_<, r_>) \xi_\ell^j(r_<, r_>)\|_{a,b}^2 \\ &= \mathcal{O}(\ell^{-2j-4}) \Bigg[\sum_{m=0}^{M} \int_0^\infty r_>^{2j+1+a+b} \frac{(r_< - r_>)^m}{m!} \Big[\tfrac{\partial^m}{\partial r_<^m} |g(r_<, r_>)|^2 \Big]_{r_< = r_>} dr_> \\ &\qquad + \sum_{n=M+1}^{N} \mathcal{O}(\ell^{-n}) \int_0^\infty r_>^{2j+1+a+b+n} \tilde{g}(r_>) dr_> \Bigg] \end{aligned} \quad (7.67)$$

and

$$\begin{aligned} &\|g(r_<, r_>) \tfrac{\partial}{\partial r_\gamma} \xi_\ell^j(r_<, r_>)\|_{a,b}^2 \\ &= \mathcal{O}(\ell^{-2j-2}) \Bigg[\sum_{m=0}^{M} \int_0^\infty r_>^{2j-1+a+b} \frac{(r_< - r_>)^m}{m!} \Big[\tfrac{\partial^m}{\partial r_<^m} |g(r_<, r_>)|^2 \Big]_{r_< = r_>} dr_> \\ &\qquad + \sum_{n=M+1}^{N} \mathcal{O}(\ell^{-n}) \int_0^\infty r_>^{2j-1+a+b+n} \tilde{g}(r_>) dr_> \Bigg] \end{aligned} \quad (7.68)$$

for $r_\gamma \in \{r_<, r_>\}$.

Proof We first insert the expansions for $|\xi_\ell^j(r_<, r_>)|^2$ and $|\frac{\partial}{\partial r_\gamma} \xi_\ell^j(r_<, r_>)|^2$ given in (7.43) and (7.44), with $m = 2$ into the $\|g(r_<, r_>)\xi_\ell^j(r_<, r_>)\|_{a,b}^2$ and $\|g(r_<, r_>)\frac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<, r_>)\|_{a,b}^2$ respectively.

Next we apply Lemma 7.7.2 with $p = 2$, $g(r_<, r_>) = |g(r_<, r_>)|^2$ and

$f(r_<, r_>) = r_<^a r_>^b f_\ell^k(r_<, r_>) f_\ell^{k'}(r_<, r_>)$ or $r_<^a r_>^b \frac{\partial}{\partial r_\gamma} f_\ell^k(r_<, r_>) \frac{\partial}{\partial r_\gamma} f_\ell^{k'}(r_<, r_>)$ to each of the $r_<$ integrals, resulting in the sums over m and n in the above.

The final stage is to apply Lemma 7.7.5 to each of the $r_<$ integrals, where, by the choice $m = p = 2$ for the preceding lemmas, the coefficient m in Lemma 7.7.5 is either 2 or 4/3, both of which are greater than 1.

Hence all powers of ℓ in terms of k and k' cancel. The calculation of the powers of ℓ and $r_>$ is trivial by following the application of the above lemmas. □

Although the decay rate results are analogous to those of [Hil85], the method of proof is entirely different. We use the explicit form for $f_\ell^k(r_<, r_>)$ and its derivatives, whilst Hill proceeds inductively, the proof of which contains a number of typographical errors.

We will later choose $g(r_<, r_>)$ to be $A_j(r_<, r_>)$ or $\frac{\partial}{\partial r_\gamma} A_j(r_<, r_>)$. It remains to estimate the decay rate of the norm of the remainder and its derivatives in (7.14) and (7.15).

7.8 Bound on the Remainder

We now wish to determine the rate of decay of the remainder terms $R_{2J-1,\ell}(r_<, r_>)$ and $\frac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)$ in Theorems 7.5.1 and 7.5.2. In particular, from (7.14) and (7.15), we see that we are interested in the integrals

$$\iint |R_{2J-1,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_> \quad \text{and} \quad \iint |\tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_>.$$

7.8.1 Remainder Decay Rates with Assumptions

For much of this section we follow the method used by Hill, starting with the definition of the differential operators $\mathcal{L}_j$ ([Hil85, page 1181]), which are chosen so that $\mathcal{L}_{2n}$ is equivalent to n applications of the angular Laplacian. We have the recursive definitions

$$\mathcal{L}_1 := \tfrac{\partial}{\partial \theta}, \quad \mathcal{L}_{2n} := \left(\tfrac{\partial}{\partial \theta} + \tfrac{\cos\theta}{\sin\theta}\right) \mathcal{L}_{2n-1}, \quad \mathcal{L}_{2n+1} := \tfrac{\partial}{\partial \theta} \mathcal{L}_{2n}, \tag{7.69}$$

and it is clear that $\mathcal{L}_{2n}\Phi_\ell(\theta) = [-\ell(\ell+1)]^n \Phi_\ell(\theta)$. Consider now an integral of the form

$$I := \int_0^\pi f(\theta)\Phi_\ell(\theta)\sin\theta d\theta,$$

(where $f(\theta)$ will be $R_{2J-1}(r_<, r_>, r_{12})$ or $\frac{\partial}{\partial r_\gamma} R_{2J}(r_<, r_>, r_{12})$), for which we have

$$[-\ell(\ell+1)]^J I = \int_0^\pi f(\theta)\big(\mathcal{L}_{2J}\Phi_\ell(\theta)\big)\sin\theta d\theta. \tag{7.70}$$

We now wish to rewrite this integral so that the operators $\mathcal{L}_j$ apply to $f(\theta)$ rather than to $\Phi_\ell(\theta)$:

Lemma 7.8.1. *For a general* $2J$*-times differentiable function* $f(\theta)$ *and operators* $\mathcal{L}_j$ *defined in (7.69) we have*

$$\begin{aligned}\int_0^\pi f(\theta)\big(\mathcal{L}_{2J}\Phi_\ell(\theta)\big)\sin\theta d\theta &= \Big[\sin\theta \sum_{j=0}^{2J-1}(-1)^j\big(\mathcal{L}_j f(\theta)\big)\big(\mathcal{L}_{2J-1-j}\Phi_\ell(\theta)\big)\Big]_0^\pi \\ &\quad + \int_0^\pi \big(\mathcal{L}_{2J} f(\theta)\big)\Phi_\ell(\theta)\sin\theta d\theta.\end{aligned}$$

Proof We prove the result by induction and first note that $\mathcal{L}_j\mathcal{L}_2 = \mathcal{L}_{j+2}$ for all j, which follows trivially from the definitions of the $\mathcal{L}_j$. Consider a general integral

$$I = \int_0^\pi F(\theta)\sin\theta \mathcal{L}_2\big(G(\theta)\big)d\theta = \int_0^\pi [F(\theta)\sin\theta\big(\tfrac{\partial^2}{\partial\theta^2}G(\theta)\big) + F(\theta)\cos\theta\big(\tfrac{\partial}{\partial\theta}G(\theta)\big)]d\theta,$$

which, integrating by parts, gives

$$\begin{aligned} I &= \big[F(\theta)\sin\theta\big(\tfrac{\partial}{\partial\theta}G(\theta)\big) + F(\theta)\cos\theta G(\theta)\big]_0^\pi \\ &\quad - \int_0^\pi \tfrac{\partial}{\partial\theta}\big(F(\theta)\sin\theta\big)\big(\tfrac{\partial}{\partial\theta}G(\theta)\big)d\theta - \int_0^\pi \tfrac{\partial}{\partial\theta}\big(F(\theta)\cos\theta\big)G(\theta)d\theta. \end{aligned} \tag{7.71}$$

Integrating the first integral by parts once again gives

$$\begin{aligned} I &= \big[F(\theta)\sin\theta\big(\tfrac{\partial}{\partial\theta}G(\theta)\big) + F(\theta)\cos\theta G(\theta) - \tfrac{\partial}{\partial\theta}\big(F(\theta)\sin\theta\big)G(\theta)\big]_0^\pi \\ &\quad + \int_0^\pi \big[\tfrac{\partial^2}{\partial\theta^2}\big(F(\theta)\sin\theta\big) - \tfrac{\partial}{\partial\theta}\big(F(\theta)\cos\theta\big)\big]G(\theta)d\theta \\ &= \big[F(\theta)\sin\theta\big(\tfrac{\partial}{\partial\theta}G(\theta)\big) - \big(\tfrac{\partial}{\partial\theta}F(\theta)\big)\sin\theta G(\theta)\big]_0^\pi \\ &\quad + \int_0^\pi \big[\big(\tfrac{\partial}{\partial\theta} + \tfrac{\cos\theta}{\sin\theta}\big)\tfrac{\partial}{\partial\theta}F(\theta)\big]G(\theta)\sin\theta d\theta. \end{aligned}$$

It is therefore clear that the result holds for $J = 1$ with the more general functions $f(\theta) = F(\theta)$ and $\Phi_\ell(\theta) = G(\theta)$.

Assume the more general result holds for J and so we have

$$\int_0^\pi F(\theta)\big(\mathcal{L}_{2(J+1)}G(\theta)\big)\sin\theta d\theta = \Big[\sin\theta\sum_{j=0}^{2J-1}(-1)^j\big(\mathcal{L}_jf(\theta)\big)\mathcal{L}_{2J-1-j}\big(\mathcal{L}_2G(\theta)\big)\Big]_0^\pi + \int_0^\pi\big(\mathcal{L}_{2J}f(\theta)\big)\big(\mathcal{L}_2G(\theta)\big)\sin\theta d\theta.$$

Using $\mathcal{L}_j\mathcal{L}_2 = \mathcal{L}_{j+2}$ and the inductive hypothesis we find

$$\int_0^\pi F(\theta)\big(\mathcal{L}_{2(J+1)}G(\theta)\big)\sin\theta d\theta = \int_0^\pi\big(\mathcal{L}_{2(J+1)}f(\theta)\big)G(\theta)\sin\theta d\theta$$
$$+\Big[\sin\theta\sum_{j=0}^{2J-1}(-1)^j\big(\mathcal{L}_jf(\theta)\big)\big(\mathcal{L}_{2J+1-j}G(\theta)\big) + \big(\mathcal{L}_{2J}f(\theta)\big)\sin\theta\big(\tfrac{\partial}{\partial\theta}G(\theta)\big)$$
$$-\big(\tfrac{\partial}{\partial\theta}\mathcal{L}_{2J}f(\theta)\big)\sin\theta G(\theta)\Big]_0^\pi,$$

recalling that $\frac{\partial}{\partial\theta}\mathcal{L}_{2J} = \mathcal{L}_{2J+1}$, and setting $F(\theta) = f(\theta)$, $G(\theta) = \Phi_\ell(\theta)$, gives the result. ☐

We now note that as long as $(\mathcal{L}_jf(\theta))(\mathcal{L}_{2J-1-j}\Phi_\ell(\theta))$ is bounded for all $j = 0,\dots,2J-1$ then the boundary terms in (7.71) are zero. In our case, $f(\theta) = R_J(r_<, r_>, r_{12})$, $I = R_{J,\ell}(r_<, r_>)$ and we have the following lemmas:

Lemma 7.8.2. *Suppose that*

$$\big|\big(\mathcal{L}_jR_{2J-1}(r_<, r_>, r_{12})\big)\big(\mathcal{L}_{2J-1-j}\Phi_\ell(\theta)\big)\big| < \infty$$

for $j = 0,\dots,2J-1$, then

$$\sum_{\ell=0}^\infty \ell^{4J}|R_{2J-1,\ell}(r_<, r_>)|^2 \le \int_{r_>-r_<}^{r_>+r_<}|\mathcal{L}_{2J}R_{2J-1}(r_<, r_>, r_{12})|^2\frac{r_{12}}{r_<r_>}dr_{12}. \qquad (7.72)$$

Further, if the integral on the right hand side is finite then

$$\lim_{\ell\to\infty}\ell^{2J}R_{2J-1,\ell}(r_<, r_>) = 0$$

and if, in addition,

$$\int_0^\infty\int_0^{r_>}\int_{r_>-r_<}^{r_>+r_<}|\mathcal{L}_{2J}R_{2J-1}(r_<, r_>, r_{12})|^2\frac{r_{12}}{r_<r_>}dr_{12}r_<^2r_>^2dr_<dr_> < \infty \qquad (7.73)$$

then

$$\lim_{\ell\to\infty} \ell^{4J} \int_0^\infty \int_0^{r_>} |R_{2J-1,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_> = 0.$$

Proof Using (7.70) with $f(\theta) = R_{2J-1}(r_<, r_>, r_{12})$ (i.e. $I = R_{2J-1,\ell}(r_<, r_>)$), along with the assumption that $(\mathcal{L}_j R_{2J-1}(r_<, r_>, r_{12}))(\mathcal{L}_{2J-1-j}\Phi_\ell(\theta))$ is finite, Lemma 7.8.1 shows that

$$[-\ell(\ell+1)]^J R_{2J-1,\ell}(r_<, r_>) = \int_0^\pi \big(\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})\big)\Phi_\ell(\theta)\sin\theta d\theta. \quad (7.74)$$

Since $\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})$ is a function of $r_<, r_>$ and θ it may be expanded as

$$\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12}) = \sum_{i=0}^\infty \alpha_i(r_<, r_>)\Phi_i(\theta).$$

Hence

$$\int_0^\pi \big(\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})\big)\Phi_\ell(\theta)\sin\theta d\theta = \alpha_\ell(r_<, r_>)$$

and so

$$[\ell(\ell+1)]^{2J}|R_{2J-1,\ell}(r_<, r_>)|^2 = |\alpha_\ell(r_<, r_>)|^2.$$

Summing over ℓ gives

$$\sum_{\ell=0}^\infty [\ell(\ell+1)]^{2J}|R_{2J-1,\ell}(r_<, r_>)|^2 = \sum_{\ell=0}^\infty |\alpha_\ell(r_<, r_>)|^2.$$

By definition,

$$|\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})|^2 = \sum_{i,j=1}^\infty \alpha_i(r_<, r_>)\alpha_j^*(r_<, r_>)\Phi_i(\theta)\Phi_j(\theta),$$

where we have used the fact that the $\Phi_i(\theta)$ are real. Using the orthonormality of the $\Phi_i(\theta)$ and the Dominated Convergence Theorem we have that, if the integral of the left hand side is bounded,

$$\int_0^\pi |\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})|^2 \sin\theta d\theta = \sum_{\ell=0}^\infty |\alpha_\ell(r_<, r_>)|^2.$$

Since $\ell^{4J} \leq [\ell(\ell+1)]^{2J}$, it follows from (7.74) that

$$\sum_{\ell=0}^\infty \ell^{4J}|R_{2J-1,\ell}(r_<, r_>)|^2 \leq \int_0^\pi |\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})|^2 \sin\theta d\theta.$$

Using that

$$\int_0^\pi f \sin\theta d\theta = \int_{r_>-r_<}^{r_<+r_>} f \tfrac{\partial\theta}{\partial r_{12}} \sin\theta dr_{12}$$

and $\frac{\partial r_{12}}{\partial\theta} = r_< r_> r_{12}^{-1} \sin\theta$ gives the first part of the result.

The second part follows trivially as if the sum is bounded then the summand must tend to zero and hence $\lim_{\ell\to\infty} \ell^2 R_{2J-1,\ell}(r_<, r_>) = 0$.

The final part is less trivial and requires the use of the Dominated Convergence Theorem. Let $f_n = \sum_{\ell=0}^{n} \ell^4 |R_{2J-1,\ell}(r_<, r_>)|^2$ then it is trivial that $f_n \to f := \sum_{\ell=0}^{\infty} \ell^4 |R_{2J-1,\ell}(r_<, r_>)|^2$ pointwise as $n \to \infty$. Further, by the positivity of the summand,

$$|f_n| \le f \le \int_0^\pi |\mathcal{L}_{2J} R_{2J-1}(r_<, r_>, r_{12})|^2 \sin\theta d\theta =: g.$$

By assumption $\int_0^\infty \int_0^{r_>} g\, r_<^2 r_>^2 dr_< dr_> < \infty$ and hence we may exchange the order of summation and integration in

$$\int_0^\infty \int_0^{r_>} \sum_{\ell=0}^{\infty} \ell^{4J} |R_{2J-1,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_> < \infty,$$

which gives the final part of the result. □

Lemma 7.8.3. *Suppose that*

$$\left|\left(\mathcal{L}_j R_{2J}(r_<, r_>, r_{12})\right)\left(\mathcal{L}_{2J-j} \Phi_\ell(\theta)\right)\right| < \infty$$

for $j = 0, \dots, 2J-1$, *that* $\frac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)$ *and* $\frac{\partial}{\partial r_\gamma}(\mathcal{L}_{2J} R_{2J}(r_<, r_>, r_{12}))$ *exist for* $r_\gamma \in \{r_<, r_>\}$, *and that* $\frac{\partial}{\partial r_\gamma}(\mathcal{L}_{2J} R_{2J}(r_<, r_>, r_{12}))$ *is a continuous function of* r_γ. *Then*

$$\sum_{\ell=0}^{\infty} \ell^{4J} \left|\tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)\right|^2 \le \int_{r_>-r_<}^{r_>+r_<} \left|\tfrac{\partial}{\partial r_\gamma}\left(\mathcal{L}_{2J} R_{2J}(r_<, r_>, r_{12})\right)\right|^2 \frac{r_{12}}{r_< r_>} dr_{12}. \quad (7.75)$$

Further, if the integral on the right hand side is finite then

$$\lim_{\ell\to\infty} \ell^{2J} \tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>) = 0$$

and if, in addition,

$$\int_0^\infty \int_0^{r_>} \int_{r_>-r_<}^{r_>+r_<} \left|\tfrac{\partial}{\partial r_\gamma}\left(\mathcal{L}_{2J} R_{2J}(r_<, r_>, r_{12})\right)\right|^2 \frac{r_{12}}{r_< r_>} dr_{12} r_<^2 r_>^2 dr_< dr_> < \infty \quad (7.76)$$

then

$$\lim_{\ell\to\infty} \ell^{4J} \int_0^\infty \int_0^{r_>} \left|\tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)\right|^2 r_<^2 r_>^2 dr_< dr_> = 0.$$

Proof As in the proof of Lemma 7.8.2 we find

$$[-\ell(\ell+1)]^J R_{2J,\ell}(r_<,r_>) = \int_0^\pi \big(\mathcal{L}_{2J} R_{2J}(r_<,r_>,r_{12})\big)\Phi_\ell(\theta)\sin\theta\,d\theta.$$

Differentiating both sides with respect to r_γ and, by the continuity of $\mathcal{L}_{2J}R_{2J}(r_<,r_>,r_{12})$ and $\frac{\partial}{\partial r_\gamma}(\mathcal{L}_{2J}R_{2J}(r_<,r_>,r_{12}))$, using Leibniz's rule [AS72] to take the derivative inside the integral gives

$$[-\ell(\ell+1)]^J \tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<,r_>) = \int_0^\pi \big[\tfrac{\partial}{\partial r_\gamma}\big(\mathcal{L}_{2J} R_{2J}(r_<,r_>,r_{12})\big)\big]\Phi_\ell(\theta)\sin\theta\,d\theta.$$

The rest of the proof is analogous to that of Lemma 7.8.2. □

7.8.2 Validity of Assumptions

It is clear from the statement of Lemma 7.8.2 that we wish to show that, for $J' \in \{2J-1, 2J\}$,

$$\big|\big(\mathcal{L}_j R_{J'}(r_<,r_>,r_{12})\big)\big(\mathcal{L}_{J'-j}\Phi_\ell(\theta)\big)\big| < \infty,\ j = 0,\dots,2J-1,$$

$$\int_{r_>-r_<}^{r_>+r_<} \big|\mathcal{L}_{2J}R_{2J-1}(r_<,r_>,r_{12})\big|^2 \frac{r_{12}}{r_<r_>} dr_{12} < \infty$$

and

$$\int_{r_>-r_<}^{r_>+r_<} \big|\tfrac{\partial}{\partial r_\gamma}(\mathcal{L}_{2J}R_{2J}(r_<,r_>,r_{12}))\big|^2 \frac{r_{12}}{r_<r_>} dr_{12} < \infty$$

for $r_\gamma \in \{r_<, r_>\}$. The other necessary assumptions will turn out to be more technical assumptions on the wavefunction.

For convenience we follow Hill in defining the more general function

$$h_m^n(r_<,r_>,r_{12}) := \int_{r_>-r_<}^{r_{12}} \frac{(r_{12}-t)^m}{m!}\frac{\partial^n}{\partial r_{12}^n}\psi(r_<,r_>,t)dt,\ m \geq 0 \tag{7.77}$$

where, in particular, $R_J(r_<,r_>,r_{12}) = h_J^{J+1}(r_<,r_>,r_{12})$. Using Leibniz's rule,

$$\tfrac{\partial}{\partial r_{12}} h_m^n(r_<,r_>,r_{12}) = h_{m-1}^n(r_<,r_>,r_{12}) \tag{7.78}$$

and so we also define

$$h_{-1}^n(r_<,r_>,r_{12}) := \frac{\partial^n}{\partial r_{12}^n}\psi(r_<,r_>,r_{12}). \tag{7.79}$$

The aim of the next step is to write the operator $\mathcal{L}_j$, which is in terms of $\frac{\partial}{\partial\theta}$, in terms of $\frac{\partial}{\partial r_{12}}$ allowing us to make use of (7.78). The following lemma follows Hill's Appendix C:

Lemma 7.8.4. *The operator $\mathcal{L}_j$, as defined recursively in (7.69), may be rewritten as*

$$\mathcal{L}_j = \sum_{k=1}^{j} u_k^j(r_<, r_>, r_{12}) \frac{\partial^k}{\partial r_{12}^k}$$

where the $u_k^j(r_<, r_>, r_{12})$ are given recursively by

$$\begin{aligned} u_k^j(r_<, r_>, r_{12}) &:= 0, \text{ if } k < 1 \text{ or } k > j, \\ u_1^1(r_<, r_>, r_{12}) &:= \tfrac{\partial r_{12}}{\partial \theta}, \\ u_k^{2j}(r_<, r_>, r_{12}) &:= \left(\tfrac{\partial}{\partial \theta} + \tfrac{\cos\theta}{\sin\theta}\right) u_k^{2j-1}(r_<, r_>, r_{12}) + \tfrac{\partial r_{12}}{\partial \theta} u_{k-1}^{2j-1}(r_<, r_>, r_{12}), \quad (7.80) \\ u_k^{2j+1}(r_<, r_>, r_{12}) &:= \tfrac{\partial}{\partial \theta} u_k^{2j}(r_<, r_>, r_{12}) + \tfrac{\partial r_{12}}{\partial \theta} u_{k-1}^{2j}(r_<, r_>, r_{12}). \quad (7.81) \end{aligned}$$

Proof The proof proceeds by induction, the first case of $j = 1$ is given by $\mathcal{L}_1 = \frac{\partial}{\partial \theta} = \frac{\partial r_{12}}{\partial \theta} \frac{\partial}{\partial r_{12}}$ and the result holds. Now assume the identities hold up to $2j - 1$ and construct $\mathcal{L}_{2j}$ which, by definition, is equal to $(\frac{\partial}{\partial \theta} + \frac{\cos\theta}{\sin\theta})\mathcal{L}_{2j-1}$. Inserting the expansion

$$\mathcal{L}_{2j-1} = \sum_{k=1}^{2j-1} u_k^{2j-1}(r_<, r_>, r_{12}) \frac{\partial^k}{\partial r_{12}^k}$$

we find

$$\begin{aligned} \mathcal{L}_{2j} &= \left(\tfrac{\partial}{\partial \theta} + \tfrac{\cos\theta}{\sin\theta}\right) \sum_{k=1}^{2j-1} u_k^{2j-1}(r_<, r_>, r_{12}) \frac{\partial^k}{\partial r_{12}^k} \\ &= \sum_{k=1}^{2j-1} \left(\tfrac{\partial}{\partial \theta} u_k^{2j-1}(r_<, r_>, r_{12})\right) \frac{\partial^k}{\partial r_{12}^k} + u_k^{2j-1}(r_<, r_>, r_{12}) \frac{\partial r_{12}}{\partial \theta} \frac{\partial}{\partial r_{12}} \frac{\partial^k}{\partial r_{12}^k} \\ &\qquad + \tfrac{\cos\theta}{\sin\theta} u_k^{2j-1}(r_<, r_>, r_{12}) \frac{\partial^k}{\partial r_{12}^k} \end{aligned}$$

and equating powers of $\frac{\partial}{\partial r_{12}}$ gives the result. The formula for progression from even to odd follows analogously, using $\mathcal{L}_{2j+1} = \frac{\partial}{\partial \theta} \mathcal{L}_{2j}$. □

It follows directly from this lemma that we may write

$$\mathcal{L}_j h_m^n(r_<, r_>, r_{12}) = \sum_{k=1}^{j} h_{m-k}^n(r_<, r_>, r_{12}) u_k^j(r_<, r_>, r_{12}) \qquad (7.82)$$

and hence we are interested in the boundedness of the $u_k^j(r_<, r_>, r_{12})$. At present we have an expression for $u_k^j(r_<, r_>, r_{12})$ with terms in both r_{12} and θ, which can can be related in terms of $r_<$ and $r_>$. Our next step is therefore to write $u_k^j(r_<, r_>, r_{12})$ in only $r_<, r_>$ and r_{12}, this again follows Appendix C of [Hil85].

Lemma 7.8.5. *The $u_k^j(r_<, r_>, r_{12})$ defined in Lemma 7.8.4 are expressible in the form*

$$u_k^j(r_<, r_>, r_{12}) = \sum_{i=0}^{j-k} \frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}} v_{k,i}^j(\theta) \tag{7.83}$$

where the $v_{k,i}^j(\theta)$ are given recursively by

$$\begin{aligned} v_{k,i}^j(\theta) &= 0, \quad \textit{if} \;\; i < 0 \;\textit{or}\; i > j-k, \\ v_{1,0}^1(\theta) &= \sin\theta, \\ v_{k,i}^{2j}(\theta) &= \left(\tfrac{\partial}{\partial\theta} + \tfrac{\cos\theta}{\sin\theta}\right) v_{k,i}^{2j-1}(\theta) + \sin\theta \left[v_{k-1,i}^{2j-1}(\theta) - (k+2i-2) v_{k,i-1}^{2j-1}(\theta)\right], \qquad (7.84) \\ v_{k,i}^{2j+1}(\theta) &= \tfrac{\partial}{\partial\theta} v_{k,i}^{2j}(\theta) + \sin\theta\left[v_{k-1,i}^{2j}(\theta) - (k+2i-2) v_{k,i}^{2j-1}(\theta)\right]. \qquad (7.85) \end{aligned}$$

Proof Again we proceed by induction and for the case of $j = 1$ we have $u_1^1(r_<, r_>, r_{12}) = \frac{\partial r_{12}}{\partial\theta} = r_< r_> \sin\theta / r_{12}$, hence $v_{1,0}^1(\theta) = \sin\theta$ and the result holds. Now assume the result holds up to $2j-1$ and insert the expansions into the formula for $u_k^{2j}(r_<, r_>, r_{12})$ given in (7.80) giving

$$\begin{aligned} u_k^{2j}(r_<, r_>, r_{12}) = &\left(\frac{\partial}{\partial\theta} + \frac{\cos\theta}{\sin\theta}\right) \sum_{i=0}^{2j-1-k} \frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}} v_{k,i}^{2j-1}(\theta) \\ &+ \sum_{i=0}^{2j-1-(k-1)} \frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}} \sin\theta\, v_{k-1,i}^{2j-1}(\theta). \end{aligned}$$

Investigating the first term further, we have that

$$\begin{aligned} \frac{\partial}{\partial\theta} \sum_{i=0}^{2j-1-k} \frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}} v_{k,i}^{2j-1}(\theta) = &\sum_{i=0}^{2j-1-k} \frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}} \frac{\partial}{\partial\theta} v_{k,i}^{2j-1}(\theta) \\ &+ (-k-2i) \frac{(r_< r_>)^{k+i+1}}{r_{12}^{k+2i+1}} \sin\theta\, v_{k,i}^{2j-1}(\theta). \end{aligned}$$

Inserting this back into the original expression and comparing powers of $\frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i}}$ leads to the result. The result for progression from $2j$ to $2j+1$ is proved in an analogous manner. □

The final stage necessary to derive a form for $u_k^j(r_<, r_>, r_{12})$ sufficient to prove the boundedness of $h_{m-k}^n(r_<, r_>, r_{12}) u_k^j(r_<, r_>, r_{12})$, and hence by (7.82) the boundedness of $\mathcal{L}_j h_m^n(r_<, r_>, r_{12})$, is to derive a more explicit form for $v_{k,i}^j(\theta)$. For this stage of the proof we deviate somewhat from [Hil85] and derive the minimal degree of the $v^j_{k,i}(\theta)$ as polynomials in $\sin\theta$ more directly. In particular we wish to show

Lemma 7.8.6. *Let*

$$\begin{aligned} \mu^{2j}_{k,i} &:= \max\{0, 2(k+i-j)\}, \\ \mu^{2j+1}_{k,i} &:= \max\{1, 2(k+i-j)-1\}, \end{aligned}$$

then for $1 \le k \le j$ *and* $0 \le i \le j-k$, $v^j_{k,i}(\theta)$ *are of the form*

$$v^j_{k,i}(\theta) = \begin{cases} \displaystyle\sum_{n=0}^{\frac{1}{2}(k+i-\mu^j_{k,i})} a^j_{k,i,n}(\sin\theta)^{2n+\mu^j_{k,i}} & \text{if } j+k+i \text{ even} \\ \displaystyle\sum_{n=0}^{\frac{1}{2}(k+i-1-\mu^j_{k,i})} a^j_{k,i,n}(\sin\theta)^{2n+\mu^j_{k,i}}\cos\theta & \text{if } j+k+i \text{ odd,} \end{cases}$$

where $a^j_{k,i,n} \in \mathbb{R}$, *and else* $v^j_{k,i}(\theta) = 0$.

Proof The cases where $v^j_{k,i}(\theta) = 0$ follow immediately from their definition in Lemma 7.8.5. It is also worth noting that the upper limits of the sums are integers since $\mu^j_{k,i}$ is of the same parity as j.

For the other cases, proceeding by induction, for $j = 1$ we need only consider the case $v^1_{1,0}(\theta) = \sin\theta$ where $\mu^1_{1,0} = 1$, $j+k+i = 2$, $k+i-\mu^1_{1,0} = 0$ and so the result holds with $a^1_{1,0,0} = 1$.

Now, assuming that the result holds up to $2j-1$, we see from (7.84) that

$$v^{2j}_{k,i}(\theta) = \left(\tfrac{\partial}{\partial\theta} + \tfrac{\cos\theta}{\sin\theta}\right) v^{2j-1}_{k,i}(\theta) + \sin\theta\left[v^{2j-1}_{k-1,i}(\theta) - (k+2i-2)v^{2j-1}_{k,i-1}(\theta)\right], \qquad (7.86)$$

so to construct $v^{2j}_{k,i}(\theta)$ we require $v^{2j-1}_{k,i}(\theta)$, $v^{2j-1}_{k-1,i}(\theta)$ and $v^{2j-1}_{k,i-1}(\theta)$. It is useful at this stage to compare

$$\begin{aligned} \mu^{2j-1}_{k,i} &= \max\{1, 2(k+i-j)+1\}, \\ \mu^{2j-1}_{k-1,i} = \mu^{2j-1}_{k,i-1} &= \max\{1, 2(k+i-j)-1\}, \text{ and} \\ \mu^{2j}_{k,i} &= \max\{0, 2(k+i-j)\}, \end{aligned}$$

from which it is clear that

$$\mu^{2j-1}_{k,i} = \mu^{2j}_{k,i} + 1.$$

We now consider two cases, firstly, if $\mu_{k,i}^{2j-1} = 1$ then we have

$$\mu_{k,i}^{2j-1} = \mu_{k-1,i}^{2j-1} = \mu_{k,i-1}^{2j-1} = 1, \quad \mu_{k,i}^{2j} = 0. \tag{7.87}$$

Secondly, if $\mu_{k,i}^{2j-1} > 1$ then, by its oddness, $\mu_{k,i}^{2j-1} \geq 3$ and hence $2(k+i-j)-1 \geq 1$, giving

$$\mu_{k,i}^{2j-1} = \mu_{k,i}^{2j} + 1, \quad \mu_{k-1,i}^{2j-1} = \mu_{k,i-1}^{2j-1} = \mu_{k,i}^{2j-1} - 2 = \mu_{k,i}^{2j} - 1. \tag{7.88}$$

Recalling that we are constructing $v_{k,i}^{2j}(\theta)$ from $v_{k,i}^{2j-1}(\theta)$, there are two cases to consider depending on whether $2j-1+k+i$ is odd or even. Suppose first that $2j-1+k+i$ is odd (so $2j-k+i-2$ is even). Therefore we have

$$\begin{aligned}
v_{k,i}^{2j-1}(\theta) &= \sum_{n=0}^{\frac{1}{2}(k+i-1-\mu_{k,i}^{2j-1})} a_{k,i,n}^{2j-1}(\sin\theta)^{2n+\mu_{k,i}^{2j-1}}\cos\theta, \\
v_{k-1,i}^{2j-1}(\theta) &= \sum_{n=0}^{\frac{1}{2}(k+i-1-\mu_{k-1,i}^{2j-1})} a_{k-1,i,n}^{2j-1}(\sin\theta)^{2n+\mu_{k-1,i}^{2j-1}}, \\
v_{k,i-1}^{2j-1}(\theta) &= \sum_{n=0}^{\frac{1}{2}(k+i-1-\mu_{k,i-1}^{2j-1})} a_{k,i-1,n}^{2j-1}(\sin\theta)^{2n+\mu_{k,i-1}^{2j-1}},
\end{aligned}$$

which, for the first term of (7.86) gives, up to a constant,

$$\begin{aligned}
&\left(\tfrac{\partial}{\partial\theta} + \tfrac{\cos\theta}{\sin\theta}\right)(\sin\theta)^{2n+\mu_{k,i}^{2j-1}}\cos\theta \\
&= (2n+\mu_{k,i}^{2j-1}+1)(\sin\theta)^{2n+\mu_{k,i}^{2j-1}-1}\cos^2\theta - (\sin\theta)^{2n+\mu_{k,i}^{2j-1}+1} \\
&= (2n+\mu_{k,i}^{2j-1}+1)(\sin\theta)^{2n+\mu_{k,i}^{2j-1}-1} - (2n+\mu_{k,i}^{2j-1}+2)(\sin\theta)^{2n+\mu_{k,i}^{2j-1}+1}.
\end{aligned}$$

Using that $\mu_{k-1,i}^{2j-1} = \mu_{k,i-1}^{2j-1}$, the second term of (7.86) contributes, up to a constant, $(\sin\theta)^{2n+\mu_{k-1,i}^{2j-1}+1}$. These two results lead to an expression for $v_{k,i}^{2j}(\theta)$ of the form

$$\begin{aligned}
v_{k,i}^{2j}(\theta) = &\sum_{n=0}^{\frac{1}{2}(k+i-1-\mu_{k,i}^{2j-1})} \alpha_n(\sin\theta)^{2n+\mu_{k,i}^{2j-1}-1} + \beta_n(\sin\theta)^{2n+\mu_{k,i}^{2j-1}+1} \\
&+ \sum_{n=0}^{\frac{1}{2}(k+i-1-\mu_{k-1,i}^{2j-1})} \gamma_n(\sin\theta)^{2n+\mu_{k-1,i}^{2j-1}+1}.
\end{aligned}$$

If $\mu_{k,i}^{2j-1} = 1$ then, using (7.87) we see that this becomes

$$\begin{aligned} v_{k,i}^{2j}(\theta) &= \sum_{n=0}^{\frac{1}{2}(k+i-2)} \alpha_n(\sin\theta)^{2n} + (\beta_n + \gamma_n)(\sin\theta)^{2n+2} \\ &=: \sum_{n=0}^{\frac{1}{2}(k+i-\mu_{k,i}^{2j})} a_{k,i,n}^{2j}(\sin\theta)^{2n+\mu_{k,i}^{2j}} \end{aligned}$$

for some $a_{k,i,n}^{2j}$, and $\mu_{k,i}^{2j} = 0$, which is of the required form.

If $\mu_{k,i}^{2j-1} > 1$ we have, using (7.88),

$$\begin{aligned} v_{k,i}^{2j}(\theta) = &\sum_{n=0}^{\frac{1}{2}(k+i-2-\mu_{k,i}^{2j})} \alpha_n(\sin\theta)^{2n+\mu_{k,i}^{2j}} + \beta_n(\sin\theta)^{2n+\mu_{k,i}^{2j}+2} + \sum_{n=0}^{\frac{1}{2}(k+i-\mu_{k,i}^{2j})} \gamma_n(\sin\theta)^{2n+\mu_{k,i}^{2j}} \\ =: &\sum_{n=0}^{\frac{1}{2}(k+i-\mu_{k,i}^{2j})} a_{k,i,n}^{2j}(\sin\theta)^{2n+\mu_{k,i}^{2j}} \end{aligned}$$

for some $a_{k,i,n}^{2j}$, which is once again of the required form. This proves the case for $2j-1 \to 2j$ with $2j-1+k+i$ even. The other results follow with almost identical proofs and the use of (7.85). □

The important point of Lemma 7.8.6 is that it shows that the lowest power of $\sin\theta$ in $v_{k,i}^j(\theta)$ is given by $\mu_{k,i}^j$. Writing $\mu_{k,i}^{2j} = \max\{0, 2(k+i)-2j\}$ and $\mu_{k,i}^{2j+1} = \max\{1, 2(k+i)-(2j+1)\}$ it is clear that, for some positive constants $C_{k,i}^j$,

$$v_{k,i}^j(\theta) \le \begin{cases} C_{k,i}^j & \text{if } 2k+2i-j \le 0 \\ C_{k,i}^j(\sin\theta)^{2k+2i-j} & \text{if } 2k+2i-j > 0. \end{cases} \tag{7.89}$$

We wish to use this result to prove a bound on $u_k^j(r_<, r_>, r_{12})$:

Lemma 7.8.7. *For $u_k^j(r_<, r_>, r_{12})$ as defined in Lemma 7.8.4 and positive constants $C_{k,i}^j$ as defined in (7.89),*

$$|u_k^j(r_<, r_>, r_{12})| \le C_k^j\Big(\frac{r_<r_>}{r_{12}}\Big)^{1/2}(r_< + r_>)^{j-1} r_{12}^{k-j+1/2} \tag{7.90}$$

where $C_k^j = \sum_{i=0}^{j-k} C_{k,i}^j$.

Proof Consider first the case $2k+2i-j \le 0$. Inserting the bound in (7.89) into the summand of the expression for $u_k^j(r_<, r_>, r_{12})$ in terms of $v_{k,i}^j(\theta)$ as given in

(7.83), and using the trivial inequality $r_{12} \le r_> + r_<$, we see that

$$\begin{aligned}\Big|\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}v^j_{k,i}(\theta)\Big| &\le C^j_{k,i}\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}\Big(\frac{r_<+r_>}{r_{12}}\Big)^{j-2k-2i}\\ &= C^j_{k,i}(r_<r_>)^{k+i}(r_<+r_>)^{j-2k-2i}r_{12}^{k-j}.\end{aligned}$$

By the arithmetic-geometric inequality

$$(r_<r_>)^{1/2} \le r_< + r_>, \tag{7.91}$$

we have

$$\Big|\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}v^j_{k,i}(\theta)\Big| \le C^j_{k,i}\Big(\frac{r_<r_>}{r_{12}}\Big)^{1/2}(r_<+r_>)^{j-1}r_{12}^{k-j+1/2} \tag{7.92}$$

where $(r_<r_>/r_{12})^{1/2}$ is kept so as to cancel with the analogous term in (7.72).

For the case where $2k+2i-j>0$ we use the inequality

$$\sin\theta \le r_{12}(r_<r_>)^{-1/2}, \tag{7.93}$$

which is easily derived from the sine rule, to give

$$\begin{aligned}\Big|\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}v^j_{k,i}(\theta)\Big| &\le C^j_{k,i}\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}(\sin\theta)^{2k+2i-j}\\ &\le C^j_{k,i}\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}\frac{r_{12}^{2k+2i-j}}{(r_<r_>)^{k+i-j/2}} = C^j_{k,i}(r_<r_>)^{j/2}r_{12}^{k-j}.\end{aligned}$$

Once again using (7.91) gives

$$\Big|\frac{(r_<r_>)^{k+i}}{r_{12}^{k+2i}}v^j_{k,i}(\theta)\Big| \le C^j_{k,i}\Big(\frac{r_<r_>}{r_{12}}\Big)^{1/2}(r_<+r_>)^{j-1}r_{12}^{k-j+1/2}$$

which is identical to (7.92).

Noting that this bound is independent of i and using the summation formula for $u^j_k(r_<,r_>,r_{12})$ given in (7.83) we have

$$|u^j_k(r_<,r_>,r_{12})| \le C^j_k\Big(\frac{r_<r_>}{r_{12}}\Big)^{1/2}(r_<+r_>)^{j-1}r_{12}^{k-j+1/2}$$

where $C^j_k = \sum_{i=0}^{j-k} C^j_{k,i}$, proving the result. □

We now prove an analogous result for the derivatives of $u^j_k(r_<,r_>,r_{12})$.

Lemma 7.8.8. *For* $\frac{\partial}{\partial r_\gamma}u^j_k(r_<,r_>,r_{12})$ *with* $u^j_k(r_<,r_>,r_{12})$ *as defined in Lemma*

7.8.4 and $r_\gamma \in \{r_<, r_>\}$ *there exist positive constants* $\tilde{C}^j_{k,\gamma}$ *such that*

$$|\tfrac{\partial}{\partial r_\gamma} u^j_k(r_<, r_>, r_{12})| \leq \tilde{C}^j_{k,\gamma} \Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2} (r_< + r_>)^{j-1} r_{12}^{k-j-\frac{1}{2}}.$$

Proof We start with the case $r_\gamma = r_<$. From (7.83) we see that

$$\begin{aligned}\tfrac{\partial}{\partial r_<} u^j_k(r_<, r_>, r_{12}) &= \sum_{i=0}^{j-k} \Big[\frac{\partial r_{12}}{\partial r_<}(-k-2i)\frac{(r_< r_>)^{k+i}}{r_{12}^{k+2i+1}} \\ &\qquad\qquad + (k+i) r_> \frac{(r_< r_>)^{k+i-1}}{r_{12}^{k+2i}} \Big] v^j_{k,i}(\theta) \\ &= \sum_{i=0}^{j-k} \Big[(-k-2i)\frac{\partial r_{12}}{\partial r_<} r_< r_> + (k+i) r_{12} r_> \Big] \frac{(r_< r_>)^{k+i-1}}{r_{12}^{k+2i+1}} v^j_{k,i}(\theta).\end{aligned}$$

Now note that $|\partial r_{12}/\partial r_<| \leq 1$, which follows trivially from the geometric interpretation of r_{12}, the triangle inequality and the definition of the derivative (or from considering the algebraic form of $(\partial r_{12}/\partial r_<)^2$).

Using the inequalities $r_> \leq r_< + r_>$, $r_{12} \leq r_< + r_>$ and $(r_< r_>)^{1/2} \leq r_> + r_<$ we therefore have that

$$|\tfrac{\partial}{\partial r_<} u^j_k(r_<, r_>, r_{12})| \leq \sum_{i=0}^{j-k} \tilde{C}^j_{k,i} \Big| \frac{(r_< r_>)^{k+i-1}}{r_{12}^{k+2i+1}} (r_< + r_>)^2 \Big| |v^j_{k,i}(\theta)|$$

for some positive constants $\tilde{C}^j_{k,i}$. The rest of the proof is virtually identical to that of Lemma 7.8.7. The proof for $r_\gamma = r_>$ is analogous. ☐

These bounds on $u^j_k(r_<, r_>, r_{12})$ and $\frac{\partial}{\partial r_\gamma} u^j_k(r_<, r_>, r_{12})$ are sufficient to show that the assumptions necessary for Lemma 7.8.2 hold, namely that:

Lemma 7.8.9. *For* $j = 0, \ldots, m$ *there exist constants* $0 < C^j < \infty$ *such that*

$$|\mathcal{L}_j h^n_m(r_<, r_>, r_{12})| \leq C^j I_n^{1/2} \Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2} (r_< + r_>)^{j-1} r_{12}^{m+1-j}$$

and $0 < C < \infty$ *such that*

$$\int_{r_> - r_<}^{r_> + r_<} |\mathcal{L}_{m+1} h^n_m(r_<, r_>, r_{12})|^2 \frac{r_{12}}{r_< r_>} dr_{12} \leq C r_>^{2m+1} I_n \tag{7.94}$$

where

$$I_n := \int_{r_> - r_<}^{r_< + r_>} \Big| \frac{\partial^n}{\partial r_{12}^n} \psi(r_<, r_>, t) \Big|^2 dt.$$

Furthermore, there exists $0 < \tilde{C} < \infty$ *such that*

$$\int_{r_>-r_<}^{r_>+r_<} \left|\tfrac{\partial}{\partial r_\gamma}\big(\mathcal{L}_m h_m^n(r_<, r_>, r_{12})\big)\right|^2 \frac{r_{12}}{r_< r_>} dr_{12} \le \tilde{C} r_>^{2m-1} I_n. \tag{7.95}$$

Proof We begin by using the identity (7.82) which implies

$$|\mathcal{L}_j h_m^n(r_<, r_>, r_{12})| \le \sum_{k=1}^{j} |h_{m-k}^n(r_<, r_>, r_{12}) u_k^j(r_<, r_>, r_{12})| \tag{7.96}$$

and hence the bound on $u_k^j(r_<, r_>, r_{12})$ given in (7.90) shows that

$$|\mathcal{L}_j h_m^n(r_<, r_>, r_{12})| \le \sum_{k=1}^{j} |h_{m-k}^n(r_<, r_>, r_{12})| \Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2} (r_< + r_>)^{j-1} r_{12}^{k-j+1/2}.$$

Recalling the definition of $h_m^n(r_<, r_>, r_{12})$ from (7.77) and using the Cauchy-Schwarz inequality we have, for $m \ge 0$,

$$\begin{aligned}
|h_m^n(r_<, r_>, r_{12})| &\le \bigg(\int_{r_>-r_<}^{r_{12}} \Big|\frac{(r_{12}-t)^m}{m!}\Big|^2 dt \int_{r_>-r_<}^{r_{12}} \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|^2 dt\bigg)^{1/2} \\
&= \frac{(r_{12} - (r_> - r_<))^{m+1/2}}{m!\sqrt{2m+1}} \bigg(\int_{r_>-r_<}^{r_<+r_>} \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|^2 dt\bigg)^{1/2} \\
&\le \frac{r_{12}^{m+1/2}}{m!\sqrt{2m+1}} \bigg(\int_{r_>-r_<}^{r_<+r_>} \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|^2 dt\bigg)^{1/2}.
\end{aligned} \tag{7.97}$$

Hence, denoting the integral

$$I_n := \int_{r_>-r_<}^{r_<+r_>} \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|^2 dt, \tag{7.98}$$

we have, for $1 \le k \le 2J-1$,

$$\begin{aligned}
&|h_{m-k}^n(r_<, r_>, r_{12}) u_k^j(r_<, r_>, r_{12})| \\
&\qquad \le I_n^{1/2} \frac{C_k^j}{(m-k)!\sqrt{2m-2k+1}} (r_< r_>)^{1/2} (r_< + r_>)^{j-1} r_{12}^{m-j+1/2}.
\end{aligned} \tag{7.99}$$

We note that, since $j \le m$, the power of r_{12} is positive, preventing a singularity at $r_{12} = 0$. The first part of the result, the bound on $\mathcal{L}_j h_m^n(r_<, r_>, r_{12})$ for $1 \le j \le m$ follows from the expansion (7.96) and the bound (7.99) (which, apart from the constant is independent of k). In particular,

$$C^j := \sum_{k=1}^{j} \frac{C_k^j}{(m-k)!\sqrt{2m-2k+1}}. \tag{7.100}$$

The case $j = 0$ follows directly from the bound (7.97). This proves the first part of the result.

The above bound does not apply for $k = m+1$ as the Cauchy-Schwarz estimate is not valid. However, since $u_1^1(r_<, r_>, r_{12}) = \frac{\partial r_{12}}{\partial \theta}$ and $u_{j+1}^j(r_<, r_>, r_{12}) = 0$, using (7.80) and (7.81), we have

$$u_j^j(r_<, r_>, r_{12}) = (\tfrac{\partial r_{12}}{\partial \theta})^j.$$

Hence, using the definition of $h_{-1}^n(r_<, r_>, r_{12})$ in (7.79),

$$\begin{aligned}
|h_{-1}^n(r_<, r_>, r_{12}) u_{m+1}^{m+1}(r_<, r_>, r_{12})| &= \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|\Big(\frac{\partial r_{12}}{\partial \theta}\Big)^{m+1} \\
&= \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|\Big(\frac{r_< r_> \sin\theta}{r_{12}}\Big)^{m+1} \\
&\leq \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, t)\Big|(r_< r_>)^{(2m+1)/4}\Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2}
\end{aligned} \tag{7.101}$$

where the last inequality uses (7.93) and $|\sin\theta| \leq 1$.

To obtain the bound on the integral we use the identity

$$\mathcal{L}_{m+1} h_m^n(r_<, r_>, r_{12}) = \sum_{k=1}^{m+1} h_{m-k}^n(r_<, r_>, r_{12}) u_k^{m+1}(r_<, r_>, r_{12})$$

and the bounds (7.99) and (7.101) to obtain

$$\begin{aligned}
&|\mathcal{L}_{m+1} h_m^n(r_<, r_>, r_{12})| \\
&\quad \leq \sum_{k=1}^{m} I_n^{1/2} \frac{C_k^{m+1}}{(m-k)!\sqrt{2m-2k+1}}\Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2}(r_< + r_>)^m \\
&\qquad\qquad + \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, r_{12})\Big|(r_< r_>)^{(2m+1)/4}\Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2} \\
&\quad \leq \Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2}\Big[I_n^{1/2}\tilde{C} r_>^m + \Big|\frac{\partial^n}{\partial r_{12}^n}\psi(r_<, r_>, r_{12})\Big| r_>^{(2m+1)/2}\Big]
\end{aligned}$$

where the second inequality uses that $r_< \leq r_>$ and the constant $\tilde{C}$ is given explicitly by

$$\tilde{C} := 2^m \sum_{k=1}^{m} \frac{C_k^{m+1}}{(m-k)!\sqrt{2m-2k+1}}.$$

Hence we see that

$$\begin{aligned}
&\int_{r_>-r_<}^{r_>+r_<} |\mathcal{L}_{m+1} h_m^n(r_<, r_>, r_{12})|^2 \frac{r_{12}}{r_< r_>} dr_{12} \\
&\quad \leq \int_{r_>-r_<}^{r_>+r_<} \Big[I_n^{1/2} \tilde{C} r_>^m + \Big| \frac{\partial^n}{\partial r_{12}^n} \psi(r_<, r_>, r_{12}) \Big| r_>^{(2m+1)/2} \Big]^2 dr_{12} \\
&\quad \leq I_n \tilde{C}^2 r_>^{2m} \int_{r_>-r_<}^{r_>+r_<} dr_{12} + 2 I_n \tilde{C} r_>^{2m+1/2} \Big(\int_{r_>-r_<}^{r_>+r_<} dr_{12} \Big)^{1/2} + r_>^{2m+1} I_n \\
&\quad \leq C I_n r_>^{2m+1}
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality, the definition of I_n from (7.98), the fact that

$$\int_{r_>-r_<}^{r_>+r_<} dr_{12} = 2r_< \leq 2r_>$$

and the definition

$$C := 2\tilde{C}^2 + 2\sqrt{2}\tilde{C} + 1 = (\sqrt{2}\tilde{C} + 1)^2.$$

Finally we move on to consider the derivatives of $\mathcal{L}_m h_m^n(r_<, r_>, r_{12})$ and using

$$\mathcal{L}_m h_m^n(r_<, r_>, r_{12}) = \sum_{k=1}^{m} h_{m-k}^n(r_<, r_>, r_{12}) u_k^m(r_<, r_>, r_{12})$$

we see that

$$\begin{aligned}
\tfrac{\partial}{\partial r_\gamma} \big(\mathcal{L}_m h_m^n(r_<, r_>, r_{12}) \big) = \sum_{k=1}^{m} \big[& h_{m-k}^n(r_<, r_>, r_{12}) \tfrac{\partial}{\partial r_\gamma} u_k^m(r_<, r_>, r_{12}) \\
& + u_k^m(r_<, r_>, r_{12}) \tfrac{\partial}{\partial r_\gamma} h_{m-k}^n(r_<, r_>, r_{12}) \big].
\end{aligned}$$

We also have that

$$|\tfrac{\partial}{\partial r_\gamma} h_m^n(r_<, r_>, r_{12})| = |\tfrac{\partial r_{12}}{\partial r_\gamma} \tfrac{\partial}{\partial r_{12}} h_m^n(r_<, r_>, r_{12})| \leq |h_{m-1}^n(r_<, r_>, r_{12})|,$$

where we have once again used that $|\partial r_{12}/\partial r_\gamma| \leq 1$. Using the bound on $|h_m^n(r_<, r_>, r_{12})|$ given in (7.97), along with the bounds on $u_k^j(r_<, r_>, r_{12})$ and

$\frac{\partial}{\partial r_\gamma} u_k^j(r_<, r_>, r_{12})$ from Lemmas 7.8.7 and 7.8.8, we have

$$|\tfrac{\partial}{\partial r_\gamma}\mathcal{L}_m h_m^n(r_<, r_>, r_{12})| \leq C I_n^{1/2}(r_< + r_>)^{m-1}\Big(\frac{r_< r_>}{r_{12}}\Big)^{1/2},$$

for some positive constant C. Hence

$$\begin{aligned}\int_{r_>-r_<}^{r_>+r_<} \big|\tfrac{\partial}{\partial r_\gamma}\mathcal{L}_m h_m^n(r_<, r_>, r_{12})\big|^2 \frac{r_{12}}{r_< r_>} dr_{12} &\leq C^2 I_n (r_< + r_>)^{2m-2} \int_{r_>-r_<}^{r_>+r_<} dr_{12} \\ &\leq \tilde{C} r_>^{2m-1} I_n,\end{aligned}$$

giving the result. □

7.8.3 Remainder Decay Rates

We now have the following result for the remainder terms:

Corollary 7.8.10. *Suppose that*

$$\int_{r_>-r_<}^{r_<+r_>} \left|\frac{\partial^{2J}}{\partial r_{12}^{2J}}\psi(r_<, r_>, t)\right|^2 dt < \infty \tag{7.102}$$

then the ℓ-projection of the remainder term

$$\begin{aligned}R_{2J-1,\ell}(r_<, r_>) &= \int_0^\pi R_{2J-1}(r_<, r_>, r_{12})\Phi_\ell(\theta)\sin\theta d\theta \\ &= \int_0^\pi \int_{r_>-r_<}^{r_{12}} \frac{(r_{12}-t)^{2J-1}}{(2J-1)!}\frac{\partial^{2J}}{\partial r_{12}^{2J}}\psi(r_<, r_>, t) dt \Phi_\ell(\theta)\sin\theta d\theta,\end{aligned}$$

satisfies

$$\lim_{\ell\to\infty} \ell^{2J} R_{2J-1,\ell}(r_<, r_>) = 0.$$

If, in addition,

$$\int_0^\infty \int_0^{r_>} r_<^{4J-1} \int_{r_>-r_<}^{r_<+r_>} \left|\frac{\partial^{2J}}{\partial r_{12}^{2J}}\psi(r_<, r_>, t)\right|^2 dt r_<^2 r_>^2 dr_< dr_> < \infty \tag{7.103}$$

then

$$\lim_{\ell\to\infty} \ell^{4J}\int_0^\infty \int_0^{r_>} |R_{2J-1,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_> = 0.$$

Proof We begin by recalling that $R_{2J-1}(r_<, r_>, r_{12}) = h_{2J-1}^{2J}(r_<, r_>, r_{12})$ and so, using (7.94) with $m = 2J-1$ and $n = 2J$ shows that (7.102) and (7.103) are stronger than the assumption on the right hand side of (7.72) and the assumption (7.73) respectively.

Furthermore, Lemma 7.8.9 and the fact that $\mathcal{L}_j\Phi_\ell(\theta) < \infty$ for all j removes the need for the assumption on the boundedness of $|\mathcal{L}_j R_{2J}(r_<, r_>, r_{12})|$ in Lemma 7.8.2.

Hence all assumptions in Lemma 7.8.2 are satisfied and the result holds. □

Corollary 7.8.11. *Suppose that*

$$\int_{r_>-r_<}^{r_<+r_>} \left| \frac{\partial^{2J+1}}{\partial r_{12}^{2J+1}} \psi(r_<, r_>, t) \right|^2 dt < \infty \tag{7.104}$$

then the ℓ-projection of the derivative of the remainder term

$$\tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>) = \tfrac{\partial}{\partial r_\gamma} \int_0^\pi R_{2J}(r_<, r_>, r_{12})\Phi_\ell(\theta)\sin\theta d\theta$$

satisfies

$$\lim_{\ell\to\infty} \ell^{2J} \tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>) = 0.$$

If, in addition,

$$\int_0^\infty \int_0^{r_>} r_<^{4J-1} \int_{r_>-r_<}^{r_<+r_>} \left| \frac{\partial^{2J+1}}{\partial r_{12}^{2J+1}} \psi(r_<, r_>, t) \right|^2 dt r_<^2 r_>^2 dr_< dr_> < \infty \tag{7.105}$$

then

$$\lim_{\ell\to\infty} \ell^{4J} \int_0^\infty \int_0^{r_>} |\tfrac{\partial}{\partial r_\gamma} R_{2J,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 dr_< dr_> = 0.$$

Proof As for the previous result, we have $R_{2J}(r_<, r_>, r_{12}) = h_{2J}^{2J+1}(r_<, r_>, r_{12})$ and so using using (7.95) with $m = 2J$ and $n = 2J + 1$ shows that (7.104) and (7.105) are stronger than the assumption on the right hand side of (7.75) and the assumption (7.76) respectively.

Furthermore, Lemma 7.8.9 removes the need for the assumption on the boundedness of $|\mathcal{L}_j R_{2J}(r_<, r_>, r_{12})|$ in Lemma 7.8.3. The continuity assumption in Lemma 7.8.3 is shown to be true by the explicit forms derived above.

Hence all assumptions in Lemma 7.8.3 are satisfied and the result holds. □

We are now in a position to prove Theorems 7.5.1 and 7.5.2:

Proof of Theorem 7.5.1 The expansion of $\psi_\ell(r_<, r_>)$ given in (7.17) follows from (7.11) for $J = 2J$. The decay rates of the norms given in (7.18) follow from the bound (7.67) in Corollary 7.7.6 with $g = A(r_<, r_>)$. Finally the results on the decay of the remainder in (7.19) and (7.20) follow directly from Corollary 7.8.10. □

Proof of Theorem 7.5.2 The expansion of $\frac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)$ given in (7.21) follows directly from (7.11) for $J = 2J + 1$. The decay rates of the norms given in (7.22) follow from the bound (7.68) in Corollary 7.7.6 with $g = A(r_<, r_>)$. Finally the results on the decay of the remainder in (7.23) and (7.24) follow directly from Corollary 7.8.11. ☐

7.9 Decay of the Energy Error

We are now in a position to prove the rate of decay of the energy error. In order to do this we prove the following results on the norms from (7.8) and (7.9):

Theorem 7.9.1. *Let* $\psi_L^\perp := \sum_{\ell=L+1}^\infty \psi_\ell(r_<, r_>)\Phi_\ell(\theta)$ *be the part of the wavefunction not captured by projection onto the first* L *angular momentum eigenfunctions* $\Phi_\ell(\theta)$*. Suppose that* $\frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, r_{12})$ *exist for* $r_> - r_< \leq r_{12} \leq r_< + r_>$, $0 \leq j \leq 4$, *that* $A_j(r_<, r_>)$ *and* $\frac{\partial}{\partial r_\gamma}A_j(r_<, r_>)$, $j = 0 \ldots 4$ *satisfy the conditions on* $A(r_<, r_>)$ *in Theorem 7.5.1 for* $N = 1$, *and that*

$$\int_0^\infty r^5 |\psi(r, r, 0)|^2 \, dr < \infty, \quad \textit{and} \quad \int_0^\infty r^7 |\psi(r, r, 0)|^2 \, dr < \infty.$$

Define

$$I_j(r_<, r_>) := \int_{r_> - r_<}^{r_< + r_>} \left| \frac{\partial^j}{\partial r_{12}^j}\psi(r_<, r_>, t) \right|^2 dt < \infty,$$

and suppose that $I_3, I_4 < \infty$,

$$\int_0^\infty \int_0^{r_>} r_<^3 I_3(r_<, r_>) r_<^2 r_>^2 \, dr_< dr_> < \infty,$$

and

$$\int_0^\infty \int_0^{r_>} r_<^7 I_4(r_<, r_>) r_<^2 r_>^2 \, dr_< dr_> < \infty.$$

Then

$$\|\psi_L^\perp\|^2 = 5\pi^2 \sum_{\ell=L+1}^\infty \Big(\ell^{-6} \int_0^\infty r^7 |\psi(r, r, 0)|^2 \, dr + \mathcal{O}(\ell^{-7}) \Big), \tag{7.106}$$

$$\|\psi_L^\perp\|_e^2 = 6\pi^2 \sum_{\ell=L+1}^\infty \Big(\ell^{-4} \int_0^\infty r^5 |\psi(r, r, 0)|^2 \, dr + o(\ell^{-4}) \Big), \tag{7.107}$$

and

$$E_L - E = 2\pi^2 L^{-3} \int_0^\infty r^5 |\psi(r, r, 0)|^2 \, dr + o(L^{-3}). \tag{7.108}$$

Before we prove the result we note that the existence of the fourth derivative is only necessary for the decay of $\|\psi_L^\perp\|^2$ given in (7.8), and in particular it is required so that the remainder term decays faster than the first term in the expansion. If we are only interested in the error in the energy then the existence of the third derivative is sufficient to show (7.107).

Proof We begin with the case of $\|\psi_L^\perp\|^2$ and by (7.8) we see that

$$\begin{aligned}\|\psi_L^\perp\|^2 &= 16\pi^2 \sum_{\ell=L+1}^{\infty} \int_0^\infty \int_0^{r_>} |\psi_\ell(r_<, r_>)|^2 r_<^2 r_>^2 \, dr_< dr_> \\ &= 16\pi^2 \sum_{\ell=L+1}^{\infty} \|\psi_\ell(r_<, r_>)\|_{2,2}^2.\end{aligned}$$

Equation (7.14) shows that, with $J = 4$,

$$\begin{aligned}\|\psi_\ell(r_<, r_>)\|_{2,2}^2 \le &\sum_{i,j=1}^{3} \|A_i(r_<, r_>)\xi_\ell^i(r_<, r_>)\|_{2,2}\|A_j(r_<, r_>)\xi_\ell^j(r_<, r_>)\|_{2,2} \\ &+ 2\sum_{k=1}^{3} \|A_k(r_<, r_>)\xi_\ell^k(r_<, r_>)\|_{2,2}\, \|R_{3,\ell}(r_<, r_>)\|_{2,2} \\ &+ \|R_{3,\ell}(r_<, r_>)\|_{2,2}^2. \end{aligned} \tag{7.109}$$

By Theorem 7.5.1 with $J = 2$, and in particular by (7.20), we have that

$$\lim_{\ell\to\infty} \ell^8 \|R_{3,\ell}(r_<, r_>)\|_{2,2}^2 = \lim_{\ell\to\infty} \ell^8 \int_0^\infty \int_0^{r_>} |R_{3,\ell}(r_<, r_>)|^2 r_<^2 r_>^2 \, dr_< dr_> = 0$$

and hence

$$\lim_{\ell\to\infty} \ell^4 \|R_{3,\ell}(r_<, r_>)\|_{2,2} = 0,$$

i.e. $\|R_{3,\ell}(r_<, r_>)\|_{2,2}^2 = o(l^{-8})$ and $\|R_{3,\ell}(r_<, r_>)\|_{2,2} = o(l^{-4})$.

The same theorem shows (choosing $A(r_<, r_>) = A_j(r_<, r_>)$ and $a = b = 2$ in (7.18)) that $\|A_j(r_<, r_>)\xi_\ell^j(r_<, r_>)\|_{2,2} = \mathcal{O}(\ell^{-j-2})$ and it is therefore clear that the slowest decaying term comes from the $i = j = 1$ term in the first sum of (7.109), with all other terms decaying at least one order of ℓ faster. Theorem 7.5.1 also shows that (choosing $A(r_<, r_>) = A_1(r_<, r_>)$, $M = 0$, $a = b = 2$ and $j = 1$ in (7.18))

$$\|A_1(r_<, r_>)\xi_\ell^1(r_<, r_>)\|_{2,2}^2 = \mathcal{O}(\ell^{-6}) \int_0^\infty r_>^7 \left[|A_1(r_<, r_>)|^2\right]_{r_<=r_>} dr_> + \mathcal{O}(\ell^{-7}).$$

By the definition of $A_j(r_<, r_>)$ in (7.16), we have that

$$\begin{aligned}\left[|A_1(r_<, r_>)|^2\right]_{r_<=r_>} &= \left[\left[\frac{\partial}{\partial r_{12}}\psi(r_<, r_>, r_{12})\right]_{r_{12}=r_>-r_<}\right]_{r_<=r_>} \\ &= \left[\frac{\partial}{\partial r_{12}}\psi(r_>, r_>, r_{12})\right]_{r_{12}=0}.\end{aligned}$$

Using the Kato cusp condition without spherical averaging [Kat57] (see also Appendix E of Hill)

$$\left[\frac{\partial}{\partial r_{12}}\psi(r_>, r_>, r_{12})\right]_{r_{12}=0} = \frac{1}{2}\psi(r_>, r_>, 0), \tag{7.110}$$

gives (7.106) up to a constant, which will be computed later.

The proof of (7.107) is analogous, noting that (7.9) shows that

$$\begin{aligned}\|\psi_L^\perp\|_e^2 = 8\pi^2 \sum_{\ell=L+1}^{\infty} \Big(&\|\tfrac{\partial}{\partial r_<}\psi_\ell(r_<, r_>)\|_{2,2} + \|\tfrac{\partial}{\partial r_>}\psi_\ell(r_<, r_>)\|_{2,2} \\ &+ \ell(\ell+1)\left[\|\psi_\ell(r_<, r_>)\|_{2,0} + \|\psi_\ell(r_<, r_>)\|_{0,2}\right]\Big).\end{aligned} \tag{7.111}$$

For the terms involving $\psi_\ell(r_<, r_>)$ we use Theorem 7.5.1 with $J = 2$ and in particular (7.18) with $M = 0$, $A(r_<, r_>) = A_1(r_<, r_>)$, $j = 1$ and either $a = 2, b = 0$ or $a = 0, b = 2$, giving

$$\|A_1(r_<, r_>)\xi_\ell^1(r_<, r_>)\|_{2,0}^2 = \mathcal{O}(\ell^{-6}) \int_0^\infty r_>^5 \left[|A_1(r_<, r_>)|^2\right]_{r_<=r_>} dr_> + \mathcal{O}(\ell^{-7}), \tag{7.112}$$

$$\|A_1(r_<, r_>)\xi_\ell^1(r_<, r_>)\|_{0,2}^2 = \mathcal{O}(\ell^{-6}) \int_0^\infty r_>^5 \left[|A_1(r_<, r_>)|^2\right]_{r_<=r_>} dr_> + \mathcal{O}(\ell^{-7}). \tag{7.113}$$

For the other terms in (7.111), we use (7.15) with $J = 3$ to obtain

$$\begin{aligned}&\|\tfrac{\partial}{\partial r_\gamma}\psi_\ell(r_<, r_>)\|_{2,2}^2 \\ &\leq \sum_{i,j=1}^{2} \Big[\|\xi_\ell^i(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}A_i(r_<, r_>)\|_{2,2}\,\|\xi_\ell^j(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}A_j(r_<, r_>)\|_{2,2} \\ &\qquad + 2\|\xi_\ell^i(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}A_i(r_<, r_>)\|_{2,2}\,\|A_j(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<, r_>)\|_{2,2} \\ &\qquad + \|A_i(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^i(r_<, r_>)\|_{2,2}\,\|A_j(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^j(r_<, r_>)\|_{2,2}\Big]\end{aligned} \tag{7.114}$$

$$\begin{aligned}
&+ 2\|\tfrac{\partial}{\partial r_\gamma} R_{J-1,\ell}(r_<, r_>)\|_{2,2} \sum_{k=1}^{2} \|A_k(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^k(r_<, r_>)\|_{2,2}\\
&+ 2\|\tfrac{\partial}{\partial r_\gamma} R_{J-1,\ell}(r_<, r_>)\|_{2,2} \sum_{k=1}^{2J-1} \|\xi_\ell^k(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}A_k(r_<, r_>)\|_{2,2}\\
&+ \|\tfrac{\partial}{\partial r_\gamma} R_{2,\ell}(r_<, r_>)\|_{2,2}^2.
\end{aligned}$$

By Theorem 7.5.2 with $J = 1$ (and in particular (7.24)) we find $\|\frac{\partial}{\partial r_\gamma} R_{2,\ell}(r_<, r_>)\|_{2,2}^2 = o(\ell^{-4})$. Furthermore, Theorem 7.5.1 shows that, for $J = 2$, $A(r_<, r_>) = \frac{\partial}{\partial r_\gamma} A_j(r_<, r_>)$, $a = b = 2$, and $j = 1, 2$ in (7.18),

$$\begin{aligned}
&\|\xi_\ell^1(r_<, r_>)\tfrac{\partial}{\partial r_\gamma} A_1(r_<, r_>)\|_{2,2}^2 = \mathcal{O}(\ell^{-6}), \text{ and}\\
&\|\xi_\ell^2(r_<, r_>)\tfrac{\partial}{\partial r_\gamma} A_2(r_<, r_>)\|_{2,2}^2 = \mathcal{O}(\ell^{-8}).
\end{aligned}$$

Using (7.22) of Theorem 7.5.2 with $J = 1$, $A(r_<, r_>) = A_2(r_<, r_>)$, $j = 2$, and $a = b = 2$ gives

$$\|A_2(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^2(r_<, r_>)\|_{2,2}^2 = \mathcal{O}(\ell^{-6}).$$

Hence it is only necessary to consider the term $\|A_1(r_<, r_>)\frac{\partial}{\partial r_\gamma}\xi_\ell^1(r_<, r_>)\|_{2,2}$ in (7.114). Once again using (7.22) with $J = 1$, $A(r_<, r_>) = A_1(r_<, r_>)$, $j = 1$, $a = b = 2$ and $M = 0$ gives

$$\|A_1(r_<, r_>)\tfrac{\partial}{\partial r_\gamma}\xi_\ell^1(r_<, r_>)\|_{2,2}^2 = \mathcal{O}(\ell^{-6}) \int_0^\infty r_>^5 \left[|A_1(r_<, r_>)|^2\right]_{r_<=r_>} dr_> + \mathcal{O}(\ell^{-7}).$$

As before, using the definition of $A_j(r_<, r_>)$ and the cusp condition gives (7.107) up to a constant.

Calculation of Constants We begin by noting that

$$\xi_\ell^1(r_<, r_>) = \int_0^\pi \big(r_{12} - (r_> - r_<)\big)\Phi_\ell(\theta)\sin\theta d\theta$$

which, for $\ell > 0$, is given by

$$\xi_\ell^1(r_<, r_>) = (\ell + \tfrac{1}{2})^{-1/2} f_\ell^1(r_<, r_>) = (\ell + \tfrac{1}{2})^{-1/2} r_<^\ell r_<^{-\ell-1}\Big(\frac{r_<^2}{2\ell+3} - \frac{r_>^2}{2\ell-1}\Big),$$

and hence

$$\tfrac{\partial}{\partial r_<}\xi_\ell^1(r_<, r_>) = (\ell + \tfrac{1}{2})^{-1/2} r_>^{-\ell-1}\Big(\frac{r_<^{\ell+1}(\ell+2)}{2\ell+3} - \frac{r_>^2 r_<^{\ell-1}\ell}{2\ell-1}\Big)$$

and

$$\tfrac{\partial}{\partial r_>}\xi_\ell^1(r_<,r_>) = (\ell+\tfrac{1}{2})^{-1/2} r_<^\ell \Big(\frac{r_<^2 r_>^{-\ell-2}(-\ell-1)}{2\ell+3} - \frac{r_>^{-\ell}(-\ell+1)}{2\ell-1}\Big).$$

It follows that

$$\begin{aligned}\int_0^{r_>} r_<^a |\xi_\ell^1(r_<,r_>)|^2 dr_< &= \frac{1}{\ell+\frac{1}{2}} r_>^{2\ell+5+a}\Bigg[\frac{1}{(2\ell+3)^2(2\ell+5+a)} \\ &\qquad -\frac{2}{(2\ell+3)(2\ell-1)(2\ell+3+a)} + \frac{1}{(2\ell-1)^2(2\ell+1+a)}\Bigg] \\ &= r_>^{2\ell+5+a}\frac{16(20\ell^2+(44+12a)\ell+2a^2+14a+21)}{(2\ell+3)^2(2\ell-1)^2(2\ell+5+a)(2\ell+3+a)(2\ell+1+a)(2\ell+1)}\end{aligned}$$

and so

$$\begin{aligned}\|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{2,2}^2 &= \frac{16(10\ell+19)}{(2\ell+3)^2(2\ell-1)^2(2\ell+7)(2\ell+5)(2\ell+1)} \\ &\qquad \times \int_0^\infty r_>^7 |A_1(r_<,r_>)|^2\, dr_>, \\ \|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{2,0}^2 &= \frac{16(10\ell+19)}{(2\ell+3)^2(2\ell-1)^2(2\ell+7)(2\ell+5)(2\ell+1)} \\ &\qquad \times \int_0^\infty r_>^5 |A_1(r_<,r_>)|^2\, dr_>, \\ \|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{0,2}^2 &= \frac{16(10\ell+7)}{(2\ell+3)^2(2\ell-1)^2(2\ell+5)(2\ell+1)^2} \\ &\qquad \times \int_0^\infty r_>^5 |A_1(r_<,r_>)|^2\, dr_>.\end{aligned}$$

In a similar fashion, we find that

$$\begin{aligned}\|A_1(r_<,r_>)\tfrac{\partial}{\partial r_<}\xi_\ell^1(r_<,r_>)\|_{2,2}^2 &= \frac{8(4\ell^3+10\ell^2+4\ell+1)}{(2\ell+3)^2(2\ell-1)^2(2\ell+5)(2\ell+1)^2} \\ &\qquad \times \int_0^\infty r_>^5 |A_1(r_<,r_>)|^2\, dr_>, \\ \|A_1(r_<,r_>)\tfrac{\partial}{\partial r_>}\xi_\ell^1(r_<,r_>)\|_{2,2}^2 &= \frac{8(4\ell^3-6\ell^2-8\ell+17)}{(2\ell+3)^2(2\ell-1)^2(2\ell+5)(2\ell+1)^2} \\ &\qquad \times \int_0^\infty r_>^5 |A_1(r_<,r_>)|^2\, dr_>.\end{aligned}$$

The leading term in $\|\psi_L^\perp\|^2$ is $16\pi^2\|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{2,2}^2$ and hence the coefficient of the leading order term is

$$\ell^{-6}16\pi^2\frac{160}{2^7}\int_0^\infty r_>^7|A_1(r_<,r_>)|^2\,dr_> = 5\pi^2\ell^{-6}\int_0^\infty r^7|\psi(r,r,0)|^2dr,$$

where we have used the cusp condition (7.110). Similarly, the leading term in $\|\psi_L^\perp\|_e^2$ is

$$8\pi^2\Big(\|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{2,2}^2 + \|A_1(r_<,r_>)\tfrac{\partial}{\partial r_>}\xi_\ell^1(r_<,r_>)\|_{2,2}^2 \\ + \ell(\ell+1)\Big[\|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{2,0}^2 + \|A_1(r_<,r_>)\xi_\ell^1(r_<,r_>)\|_{0,2}^2\Big]\Big)$$

which, combining the above, is given by

$$8\pi^2\frac{192(2\ell^3+7\ell^2+6\ell+1)}{(2\ell+3)^2(2\ell-1)^2(2\ell+7)(2\ell+5)(2\ell+1)}\int_0^\infty r_>^5|A_1(r_<,r_>)|^2\,dr_>,$$

and the leading order term is thus

$$8\pi^2\ell^{-4}\frac{192\cdot 2}{2^7}\int_0^\infty r_>^5|A_1(r_<,r_>)|^2\,dr_> = 6\pi^2\ell^{-4}\int_0^\infty r^5|\psi(r,r,0)|^2dr,$$

where we have once again used the cusp condition (7.110). We have therefore shown (7.106) and (7.107) and it remains to show (7.108). For this we note that

$$\frac{1}{n}(L+1)^{-n+1} = \int_{L+1}^\infty \ell^{-n} \le \sum_{\ell=L+1}^\infty \ell^{-n} \le \int_{L+1}^\infty (\ell-1)^{-n} = \frac{1}{n}L^{-n+1}$$

and applying this to the leading term of E_L, which by Lemma 7.2.2 and the above decay rate estimates, is given by the leading term of $\|\psi_L^\perp\|_e^2$, gives the result. □

Note that Theorem 7.9.1 gives the same leading order result as [Hil85].

7.10 Conclusions and Open Problems

The first thing of note is that the regularity of the wavefunction required by Theorem 7.9.1 is not physically realistic. For example, [FHOHOS05] have shown that any eigenfunction of the atomic Schrödinger equation can be expected to be at most $C^{1,1}$ which is far short of the required third derivatives necessary for the result to hold.

However, the similar results of [Sch62] and [KM92] (both of which use a perturbation expansion in $1/Z$), along with the numerical simulations such as [BM06] suggest that the derived decay rate is indeed correct. In order to rigorously prove this result using only the known properties of the wavefunction it is clear that a different method must be used.

One possibility would be to consider not the wavefunction but the density matrix. One advantage of this is that it would remove the need for an analogous coordinate system for many-electron atoms, which one would envisage leading to a much more difficult problem. Further, the density matrix can be thought of as the kernel of an integral operator, for which more standard methods may be applicable.

In order to utilise the density matrices it would be necessary to derive an asymptotic decay rate for the occupation numbers of the natural orbitals. (see, for example, [Löw55] for an introduction to density matrices).

Another possible extension would be to derive an asymptotic decay rate in terms of the total number of orbitals N. As a first step we may use the result for L to derive an approximate rate for N. We begin by noting that an angular momentum eigenspace with eigenvalue $L = \ell(\ell+1)$ is of dimension $2\ell+1 \sim \sqrt{L}$. Hence we see that, in order to span the whole space, we must have $N \sim L\sqrt{L}$. It follows that $L^{-3} \sim N^{-2}$ and we propose a naïve estimate that the energy error should decay as N^{-2}.

Appendix A

Proof of Lemma 2.3.4

We wish to prove Lemma 2.3.4 and begin by considering the modified integral

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{e^{-\lambda|x-y|}}{|x-y|}f^*(x)g(y),\ \lambda > 0. \tag{A.1}$$

Now, $e^{-\lambda|x|}/|x| \in L^1(\mathbb{R}^3)$ and hence we may compute its Fourier transform. Either by following a similar argument to that of Lemma 2.5.1, or by finding the anti-derivative of the result with respect to λ, we have

$$\widehat{\frac{e^{-\lambda|x|}}{|x|}} = \frac{4\pi}{\lambda^2+|k|^2}.$$

Firstly we note that $e^{-\lambda|x|} < 1$ and hence $|\frac{e^{-\lambda|x-y|}}{|x-y|}f^*(x)g(y)| < \frac{1}{|x-y|}|f^*(x)||g(y)|$. Secondly, the right hand side of this inequality is in $L^1(\mathbb{R}^6)$:

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{1}{|x-y|}|f^*(x)||g(y)| = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dzdy\frac{1}{|z|}|f^*(z+y)||g(y)|$$
$$= \int_{\{|z|\leq 1\}}\int_{\mathbb{R}^3} dzdy\frac{1}{|z|}|f^*(z+y)||g(y)| + \int_{\{|z|>1\}}\int_{\mathbb{R}^3} dzdy\frac{1}{|z|}|f^*(z+y)||g(y)|.$$

For the first integral we use that $f \in L^\infty(\mathbb{R}^3)$ and $g \in L^1(\mathbb{R}^3)$ and hence

$$\begin{aligned}\int_{\{|z|\leq 1\}}\int_{\mathbb{R}^3} dzdy\frac{1}{|z|}|f^*(z+y)||g(y)| &\leq \|f\|_\infty \int_{\{|z|\leq 1\}}\frac{1}{|z|}dz\int_{\mathbb{R}^3}|g(y)|dy \\ &= 2\pi\|f\|_\infty\|g\|_1 < \infty.\end{aligned}$$

For the second integral we use that both f and g are in $L^1(\mathbb{R}^3)$ and so

$$\begin{aligned}\int_{\{|z|>1\}}\int_{\mathbb{R}^3} dzdy\frac{1}{|z|}|f^*(z+y)||g(y)| &\le \int_{\{|z|>1\}}\int_{\mathbb{R}^3} dzdy|f^*(z+y)||g(y)| \\ &\le \|f\|_1\|g\|_1 < \infty,\end{aligned}$$

and the claim follows.

Now, since $e^{-\lambda|x|} \to 1$ as $\lambda \to 0$, the Dominated Convergence theorem implies

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{e^{-\lambda|x-y|}}{|x-y|}f^*(x)g(y) \to \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{1}{|x-y|}f^*(x)g(y),\ \lambda \to 0. \quad \text{(A.2)}$$

Next we rewrite (A.1) as the integral of a convolution:

Definition A.1 ([Fri07]). *Let* $f, h \in L^1(\mathbb{R}^n)$. *Define the* convolution *by*

$$(f * h)(x) = \int_{\mathbb{R}^n} f(y)h(x-y)dy.$$

Note by Fubini's theorem that $f * h \in L^1(\mathbb{R}^n)$. Setting $h := e^{-\lambda|x|}/|x|$ we have that

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{e^{-\lambda|x-y|}}{|x-y|}f^*(x)g(y) = \int_{\mathbb{R}^3}\big(f * h\big)(y)g(y)dy.$$

Further we have that, for $f, h \in L^1(\mathbb{R}^n)$, $\widehat{f * h}(k) = \hat{f}(k)\hat{h}(k)$ [Fri07]. Since $\hat{f} \in L^1(\mathbb{R}^3)$ and $\hat{h} \in L^\infty(\mathbb{R}^3)$ we have that $\widehat{f * h} \in L^1(\mathbb{R}^3)$ and hence we may apply Plancherel's Theorem:

Lemma A.2 (Plancherel's Theorem, [Fri07]). *Suppose that* $f, g \in L^1(\mathbb{R}^n)$ *and* $\hat{f}, \hat{g} \in L^1(\mathbb{R}^n)$ *then*

$$\int_{\mathbb{R}^n} f^*(y)g(y)dy = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{f}^*(k)\hat{g}(k)dk.$$

to give

$$\begin{aligned}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy\frac{e^{-\lambda|x-y|}}{|x-y|}f^*(x)g(y) &= \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\Big(\widehat{\frac{e^{-\lambda|\cdot|}}{|\cdot|} * f^*}\Big)(k)\hat{g}(k)dk \\ &= \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\frac{4\pi}{\lambda^2+|k|^2}\hat{f}^*(k)\hat{g}(k)dk. \qquad \text{(A.3)}\end{aligned}$$

It is clear that

$$\frac{4\pi}{\lambda^2+|k|^2}\hat{f}^*(k)\hat{g}(k) \to \frac{4\pi}{|k|^2}\hat{f}^*(k)\hat{g}(k),\ \lambda \to 0$$

for almost every k. Further we see that, for $\lambda > 0$, $1/(\lambda^2 + |k|^2) < 1/|k|^2$ and hence, in order to apply the Dominated Convergence theorem and show that

$$\frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{4\pi}{\lambda^2 + |k|^2} \hat{f}^*(k)\hat{g}(k)dk \rightarrow \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{4\pi}{|k|^2} \hat{f}^*(k)\hat{g}(k)dk, \ \lambda \rightarrow 0, \tag{A.4}$$

it remains to show that $\int_{\mathbb{R}^3} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk < \infty$.

Consider

$$\int_{\mathbb{R}^3} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk = \int_{|k|\geq 1} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk + \int_{|k|<1} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk,$$

the first term of which is finite since

$$\int_{|k|\geq 1} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk \leq 4\pi \int_{|k|\geq 1} |\hat{f}^*(k)||\hat{g}(k)|dk \leq 4\pi \|\hat{f}\|_\infty \|\hat{g}\|_1 < \infty,$$

and the second term is also finite since

$$\int_{|k|<1} \frac{4\pi}{|k|^2} |\hat{f}^*(k)||\hat{g}(k)|dk \leq (4\pi)^2 \|\hat{f}\|_\infty \|\hat{g}\|_\infty < \infty,$$

and hence (A.4) holds.

Finally, using (A.2), (A.3) and (A.4) gives

$$\begin{array}{ccc} \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy \frac{e^{-\lambda|x-y|}}{|x-y|} f^*(x)g(y) & = & \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} \frac{4\pi}{\lambda^2+|k|^2} \hat{f}^*(k)\hat{g}(k)dk \\ \downarrow (\lambda \rightarrow 0) & & \downarrow (\lambda \rightarrow 0) \\ \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} dxdy \frac{1}{|x-y|} f^*(x)g(y) & & \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} \frac{4\pi}{|k|^2} \hat{f}^*(k)\hat{g}(k)dk \end{array},$$

and the result holds.

Appendix B

Coulomb and Exchange Integrals

The following Coulomb and exchange integrals are evaluated using (2.5) along with the explicit Fourier transforms given in Table 2.2. As mentioned in Section 2.5.2, a number of standard angular integrals are needed to evaluate the Coulomb and exchange integrals. These are

$$\int_0^\pi \cos^4\theta \sin\theta d\theta = \frac{2}{5}, \quad \int_0^\pi \cos^2\theta \sin\theta d\theta = \frac{2}{3}, \quad \int_0^\pi \sin\theta d\theta = 2,$$

$$\int_0^\pi \sin^3\theta d\theta = \frac{4}{3}, \quad \int_0^\pi \sin^3\theta \cos^2\theta d\theta = \frac{4}{15},$$

$$\text{and} \quad \int_0^{2\pi} \sin^2\phi d\phi = \int_0^{2\pi} \cos^2\phi d\phi = \pi.$$

All angular integrals necessary are evaluated with these standard forms, while the radial integrals are evaluated using Maple. Alternatively they may be evaluated via contour integration.

We begin with the integrals that do not contain any angular terms:

$$\begin{aligned}
&(11|22) \\
&= \frac{1}{2\pi^2}\int_{\mathbb{R}^3} \frac{1}{|k|^2}\frac{16Z^4}{(4Z^2+|k|^2)^2}\left(\frac{2Z^4}{(Z^2+|k|^2)^2} - \frac{7Z^6}{(Z^2+|k|^2)^3} + \frac{6Z^8}{(Z^2+|k|^2)^4}\right)dk \\
&= \frac{2}{\pi}\int_0^\infty \frac{16Z^4}{(4Z^2+r^2)^2}\left(\frac{2Z^4}{(Z^2+r^2)^2} - \frac{7Z^6}{(Z^2+r^2)^3} + \frac{6Z^8}{(Z^2+r^2)^4}\right)dr \\
&= \frac{2}{\pi}\frac{17Z\pi}{162} = \frac{17}{81}Z.
\end{aligned}$$

$$
\begin{aligned}
(12|21) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\left(\frac{4\sqrt{2}Z^4}{\left((\frac{3}{2}Z)^2+|k|^2\right)^2}-\frac{9\sqrt{2}Z^6}{\left((\frac{3}{2}Z)^2+|k|^2\right)^3}\right)^2 dk\\
&= \frac{2}{\pi}\int_0^\infty\left(\frac{4\sqrt{2}Z^4}{\left((\frac{3}{2}Z)^2+r^2\right)^2}-\frac{9\sqrt{2}Z^6}{\left((\frac{3}{2}Z)^2+r^2\right)^3}\right)^2 dr\\
&= \frac{2}{\pi}\frac{8Z\pi}{729} = \frac{16}{729}Z.
\end{aligned}
$$

$$
\begin{aligned}
(22|22) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\left(\frac{2Z^4}{(Z^2+|k|^2)^2}-\frac{7Z^6}{(Z^2+|k|^2)^3}+\frac{6Z^8}{(Z^2+|k|^2)^4}\right)^2 dk\\
&= \frac{2}{\pi}\int_{\mathbb{R}^3}\left(\frac{2Z^4}{(Z^2+r^2)^2}-\frac{7Z^6}{(Z^2+r^2)^3}+\frac{6Z^8}{(Z^2+r^2)^4}\right)^2 dr\\
&= \frac{2}{\pi}\frac{77Z\pi}{1024} = \frac{77}{512}Z.
\end{aligned}
$$

For the integrals containing only one type of p-orbital, we note that, by rotational symmetry, each of the p-orbitals is equivalent and hence we may consider just p_3, with $k_3 = r\cos\theta$. The integrals are given by

$$
\begin{aligned}
(11|33) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\frac{16Z^4}{(4Z^2+r^2)^2}\left(\frac{Z^6}{(Z^2+|k|^2)^3}-\frac{6Z^6k_3^2}{(Z^2+|k|^2)^4}\right)dk\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\frac{16Z^4}{(4Z^2+r^2)^2}\left(\frac{Z^6}{(Z^2+r^2)^3}-\frac{6Z^6r^2\cos^2\theta}{(Z^2+r^2)^4}\right)\sin\theta d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\frac{16Z^4}{(4Z^2+r^2)^2}\left(\frac{2Z^6}{(Z^2+r^2)^3}-\frac{4Z^6r^2}{(Z^2+r^2)^4}\right)dr\\
&= \frac{1}{\pi}\frac{59Z\pi}{243} = \frac{59}{243}Z.
\end{aligned}
$$

$$
\begin{aligned}
(13|31) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\left(-\frac{6\sqrt{2}Z^5k_3}{((\frac{3}{2}Z)^2+|k|^2)^3}\right)^2 dk\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\frac{72Z^{10}r^2\cos^2\theta\sin\theta}{\left((\frac{3}{2}Z)^2+r^2\right)^6}d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\frac{48r^2}{\left((\frac{3}{2}Z)^2+r^2\right)^6}dr = \frac{112}{6561}Z.
\end{aligned}
$$

$$\begin{aligned}
(22|33) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\bigg(\frac{2Z^4}{(Z^2+|k|^2)^2}-\frac{7Z^6}{(Z^2+|k|^2)^3}+\frac{6Z^8}{(Z^2+|k|^2)^4}\bigg)\\
&\qquad\times\bigg(\frac{Z^6}{(Z^2+|k|^2)^3}-\frac{6Z^6k_3^2}{(Z^2+|k|^2)^4}\bigg)dk\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\bigg(\frac{2Z^4}{(Z^2+r^2)^2}-\frac{7Z^6}{(Z^2+r^2)^3}+\frac{6Z^8}{(Z^2+r^2)^4}\bigg)\\
&\qquad\times\bigg(\frac{Z^6}{(Z^2+r^2)^3}-\frac{6Z^6r^2\cos^2\theta}{(Z^2+r^2)^4}\bigg)\sin\theta d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\bigg(\frac{2Z^4}{(Z^2+r^2)^2}-\frac{7Z^6}{(Z^2+r^2)^3}+\frac{6Z^8}{(Z^2+r^2)^4}\bigg)\\
&\qquad\times\bigg(\frac{2Z^6}{(Z^2+r^2)^3}-\frac{4Z^6r^2}{(Z^2+r^2)^4}\bigg)dr = \frac{83}{512}Z.
\end{aligned}$$

$$\begin{aligned}
(23|32) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\bigg(\frac{6Z^7k_3}{(Z^2+|k|^2)^4}-\frac{3Z^5k_3}{(Z^2+|k|^2)^3}\bigg)^2dk\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\bigg(\frac{6Z^7r\cos\theta}{(Z^2+r^2)^4}-\frac{3Z^5r\cos\theta}{(Z^2+r^2)^3}\bigg)^2\sin\theta d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\frac{2}{3}\bigg(\frac{6Z^7r}{(Z^2+r^2)^4}-\frac{3Z^5r}{(Z^2+r^2)^3}\bigg)^2dr\\
&= \frac{2}{3\pi}\frac{45Z\pi}{1024} = \frac{15}{512}Z.
\end{aligned}$$

$$\begin{aligned}
(33|33) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\bigg(\frac{Z^6}{(Z^2+|k|^2)^3}-\frac{6Z^6k_j^2}{(Z^2+|k|^2)^4}\bigg)^2dk\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\bigg(\frac{Z^6}{(Z^2+r^2)^3}-\frac{6Z^6r^2\cos^2\theta}{(Z^2+r^2)^4}\bigg)^2\sin\theta d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\int_0^\pi\bigg(\frac{Z^{12}}{(Z^2+r^2)^6}-\frac{12Z^{12}r^2\cos^2\theta}{(Z^2+r^2)^7}+\frac{36Z^{12}r^4\cos^4\theta}{(Z^2+r^2)^8}\bigg)\sin\theta d\theta dr\\
&= \frac{1}{\pi}\int_0^\infty\frac{2Z^{12}}{(Z^2+r^2)^6}-\frac{8Z^{12}r^2}{(Z^2+r^2)^7}+\frac{72Z^{12}r^4}{5(Z^2+r^2)^8}dr = \frac{501}{2560}Z.
\end{aligned}$$

The remaining integral involves two different p-orbitals and, again by rotational symmetry, we may choose these to be p_3 and p_1, as in the case of (33|44) calculated in Section 2.5.2, giving

$$\begin{aligned}
(34|43) &= \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\frac{1}{|k|^2}\left(-\frac{6k_3k_1Z^6}{(Z^2+|k|^2)^4}\right)^2 dk \\
&= \frac{1}{2\pi^2}\int_0^\infty\int_0^\pi\int_0^{2\pi}\frac{36Z^{12}r^4\cos^2\theta\sin^3\theta\sin^2\phi}{(Z^2+r^2)^8}d\phi d\theta dr \\
&= \frac{1}{2\pi}\int_0^\infty\int_0^\pi\frac{36Z^{12}r^4\cos^2\theta\sin^3\theta}{(Z^2+r^2)^8}d\theta dr \\
&= \frac{1}{2\pi}\frac{4}{15}\int_0^\infty\frac{36Z^{12}r^4}{(Z^2+r^2)^8}dr \\
&= \frac{1}{2\pi}\frac{4}{15}\frac{81Z\pi}{1024} = \frac{27}{2560}Z.
\end{aligned}$$

Appendix C

$\underline{L}^2$-$\underline{S}^2$-L_3-S_3 Eigenspaces

In the following we use the fact that the $1s\overline{1s}$ orbitals present in every Slater determinant do not contribute to the formation of angular momentum and spin eigenspaces. This is due to them being a spin pair of rotationally invariant orbitals. Also recall that we need only consider $s \geq 0$.

C.1 Lithium and Fluorine

We begin with the case of Lithium (which by Theorem 3.1.10 will also give us the case of Fluorine) and see that there are only two S_3 eigenspaces, with eigenvalues $\frac{1}{2}$ and $-\frac{1}{2}$. The $s = \frac{1}{2}$ space is spanned by

$$V_{1/2} := \{|s\rangle, |p_1\rangle, |p_2\rangle, |p_3\rangle\}$$

and since these are single orbitals we need only (3.7) and (3.4), which give $\underline{L}^2 s = 0$, $\underline{L}^2 p_i = 2 \cdot p_i$, for $i = 1, 2, 3$, and $\underline{S}^2 \psi = \frac{3}{4}\psi$ for all $\psi \in V_{1/2}$.

It is trivial that each of the Slater determinants with $\underline{L}^2|\Psi\rangle = 0$ (i.e. $|s\rangle$) also satisfy $L_3|\Psi\rangle = 0$. For the space with $\underline{L}^2|\Psi\rangle = 2$, with an ordered basis $\{|p_1\rangle, |p_2\rangle, |p_3\rangle\}$, we have an L_3 matrix given by

$$\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{C.1}$$

from which it is clear that $|p_3\rangle$ is the only eigenfunction with L_3 eigenvalue zero, the other two eigenfunctions being $|p_1\rangle + i|p_2\rangle$ and $|p_1\rangle - i|p_2\rangle$ with eigenvalues 1 and -1 respectively.

C.2 Beryllium and Oxygen

The case of Beryllium is more complicated as we need to consider the two-electron terms. We see that $s \in \{-1, 0, 1\}$ and so we consider first the case $s = 0$ which is a 16-dimensional space spanned by

$$V_0 := \{|s\overline{s}\rangle, |s\overline{p_i}\rangle, |\overline{s}p_i\rangle, |p_i\overline{p_j}\rangle \mid i, j = 1, 2, 3\}.$$

We start with $\underline{S}^2$ and find that, for $\psi \neq \phi$,

$$\begin{aligned} \underline{S}^2|\psi\overline{\psi}\rangle &= 0, \\ \underline{S}^2|\psi\overline{\phi}\rangle &= |\psi\overline{\phi}\rangle + |\overline{\psi}\phi\rangle, \end{aligned}$$

and hence the $\underline{S}^2$ eigenfunctions are of the form $|\psi\overline{\psi}\rangle$, $|\psi\overline{\phi}\rangle - |\overline{\psi}\phi\rangle$ and $|\psi\overline{\phi}\rangle + |\overline{\psi}\phi\rangle$, the first two having eigenvalue 0 and the last having eigenvalue 2. We therefore write $V_0 = V_{0,0} \cup V_{0,2}$ where

$$\begin{aligned} V_{0,0} &:= \{|\psi\overline{\psi}\rangle, |\psi\overline{\phi}\rangle - |\overline{\psi}\phi\rangle \mid \psi, \phi \in \{s, p_1, p_2, p_3\}, \psi \neq \phi\}, \\ V_{0,2} &:= \{|\psi\overline{\phi}\rangle + |\overline{\psi}\phi\rangle \mid \psi, \phi \in \{s, p_1, p_2, p_3\}, \psi \neq \phi\}. \end{aligned}$$

We now wish to partition each of these spaces into their $\underline{L}^2$ eigenspaces. Beginning with $V_{0,0}$ we see that

$$\begin{aligned} \underline{L}^2|s\overline{s}\rangle &= 0, \\ \underline{L}^2|p_i\overline{p_i}\rangle &= 4|p_i\overline{p_i}\rangle - 2|p_{i-1}\overline{p_{i-1}}\rangle - 2|p_{i+1}\overline{p_{i+1}}\rangle, \\ \underline{L}^2(|s\overline{p_i}\rangle - |\overline{s}p_i\rangle) &= 2(|s\overline{p_i}\rangle - |\overline{s}p_i\rangle), \text{ and} \\ \underline{L}^2(|p_i\overline{p_j}\rangle - |\overline{p_i}p_j\rangle) &= 6(|p_i\overline{p_j}\rangle - |\overline{p_i}p_j\rangle). \end{aligned}$$

The only subspace that remains to be diagonalized is $\{|p_1\overline{p_1}\rangle, |p_2\overline{p_2}\rangle, |p_3\overline{p_3}\rangle\}$, which is equivalent to diagonalizing the matrix

$$\begin{pmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{pmatrix} \tag{C.2}$$

leading to the eigenfunction $|p_1\overline{p_1}\rangle + |p_2\overline{p_2}\rangle + |p_3\overline{p_3}\rangle$, with eigenvalue 0, and the space spanned by $\{a|p_1\overline{p_1}\rangle + b|p_2\overline{p_2}\rangle + c|p_3\overline{p_3}\rangle \mid a + b + c = 0\}$, with eigenvalue 6.

Next we consider $V_{0,2}$ and see that

$$\begin{aligned}\underline{L}^2(|s\overline{p_i}\rangle + |\overline{s}p_i\rangle) &= 2(|s\overline{p_i}\rangle + |\overline{s}p_i\rangle),\\ \underline{L}^2(|p_i\overline{p_j}\rangle + |\overline{p_i}p_j\rangle) &= 2(|p_i\overline{p_j}\rangle + |\overline{p_i}p_j\rangle).\end{aligned}$$

and thus all eigenfunctions in $V_{0,2}$ have $\underline{L}^2$ eigenvalue 2.

We now move on to the case where $s = 1$, which has a 6-dimensional space spanned by

$$V_1 := \{|sp_i\rangle, |p_ip_j\rangle \mid i, j = 1, 2, 3, i \neq j\}.$$

Since all the spins are aligned these are all eigenfunctions of $\underline{S}^2$ and have eigenvalue 2. Moving on to $\underline{L}^2$ we see that

$$\begin{aligned}\underline{L}^2|sp_i\rangle &= 2|sp_i\rangle,\\ \underline{L}^2|p_ip_j\rangle &= 2|p_ip_j\rangle,\end{aligned}$$

and hence each Slater determinant of V_1 is an eigenfunction of $\underline{L}^2$ and $\underline{S}^2$ with both eigenvalues equal to 2.

Moving on to construct the L_3 eigenfunctions, we begin with the eigenspace with $\underline{L}^2\Psi = 0$ and $\underline{S}^2\Psi = 0$ (which must have $s = 0$). This is of dimension 2 and has minimal dimension 1 and thus has degeneracy 2. It is therefore clear that both $\underline{L}^2$-$\underline{S}^2$-S_3 in this space are eigenfunctions have L_3 eigenvalue zero.

For the space with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = 0$ and $s = 0$ the degeneracy is 1 and we note that each Slater determinant has a single p-orbital, leading to an L_3 matrix of the form $(C.1)$. It follows that $|s\overline{p_3}\rangle - |\overline{s}p_3\rangle$ is an $\underline{L}^2$-$\underline{S}^2$-S_3-L_3 eigenfunction with L_3 eigenvalue zero.

The space with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = 2\Psi$ and $s = 1$ has degeneracy 2, and with ordered basis $\{|sp_1\rangle, |sp_2\rangle, |sp_3\rangle, |p_1p_2\rangle, |p_1p_3\rangle, |p_2p_3\rangle\}$ the L_3 matrix is given by

$$\begin{pmatrix} 0 & -i & 0 & 0 & 0 & 0\\ i & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & -i\\ 0 & 0 & 0 & 0 & i & 0 \end{pmatrix}, \tag{C.3}$$

from which it is clear that the $\underline{L}^2$-$\underline{S}^2$-S_3-L_3 simultaneous eigenfunctions with L_3

eigenvalue zero are $|sp_3\rangle$ and $|p_1p_2\rangle$.

The fact that (C.3) is block diagonal with 3×3 blocks follows from the fact that L_3 commutes with parity.

The final space has $\underline{L}^2\Psi = 6\Psi$ and $\underline{S}^2\Psi = 0$ and hence $s = 1$. The dimension is 5, which is the same as the minimal dimension, and thus the degeneracy is 1. Taking an ordered basis of

$$\{|p_1\overline{p_2}\rangle - |\overline{p_1}p_2\rangle, |p_1\overline{p_3}\rangle - |\overline{p_1}p_3\rangle, |p_2\overline{p_3}\rangle - |\overline{p_2}p_3\rangle, |p_1\overline{p_1}\rangle - |p_2\overline{p_2}\rangle, |p_2\overline{p_2}\rangle - |p_3\overline{p_3}\rangle\},$$

the L_3 matrix is given by

$$\begin{pmatrix} 0 & 0 & 0 & 2i & -i \\ 0 & 0 & -i & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ -2i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{C.4}$$

which, with the obvious ordering of basis elements, has eigenfunctions $\Psi_4 + 2\Psi_5$, $i\Psi_2 + \Psi_3$, $-i\Psi_2 + \Psi_3$, $i\Psi_1 + \Psi_4$, and $-i\Psi_1 + \Psi_4$, with eigenvalues $0, 1, -1, 2$ and -2 respectively. Hence the eigenfunction with eigenvalue 0 is given by $2|p_3\overline{p_3}\rangle - |p_1\overline{p_1}\rangle - |p_2\overline{p_2}\rangle$. This completes the analysis of the Beryllium (and hence Oxygen) eigenspaces.

C.3 Boron and Nitrogen

For these two cases we see that the possible eigenvalues of S_3 are $s \in \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$ and so we must consider the cases $s = \frac{1}{2}$ and $s = \frac{3}{2}$. Analogously to the previous cases we form the 24-dimensional space

$$\begin{aligned} V_{1/2} := \{&|s\overline{s}p_i\rangle, |s\overline{p_i}p_i\rangle, |s\overline{p_i}p_j\rangle, |p_i\overline{p_i}p_j\rangle, |p_i\overline{p_j}p_k\rangle \mid i \neq j \neq k, i < k\} \\ &\cup \{|p_i\overline{s}p_j\rangle \mid i < j\}. \end{aligned}$$

There are two different kinds of spatial-spin arrangements, namely $|\psi_1\overline{\psi_1}\psi_2\rangle$ and $|\psi_1\overline{\psi_2}\psi_3\rangle$ where $\psi_1 \neq \psi_2 \neq \psi_3$. We have that

$$\begin{aligned} \underline{S}^2|\psi_1\overline{\psi_1}\psi_2\rangle &= \tfrac{3}{4}|\psi_1\overline{\psi_1}\psi_2\rangle, \\ \underline{S}^2|\psi_1\overline{\psi_2}\psi_3\rangle &= \tfrac{7}{4}|\psi_1\overline{\psi_2}\psi_3\rangle - |\psi_2\overline{\psi_1}\psi_3\rangle - |\psi_1\overline{\psi_3}\psi_2\rangle, \end{aligned}$$

and to find the eigenfunctions for the second case we need to diagonalize the matrix (in the ordered basis $\{|\psi_1\overline{\psi_2}\psi_3\rangle, |\psi_2\overline{\psi_1}\psi_3\rangle, |\psi_1\overline{\psi_3}\psi_2\rangle\}$)

$$\begin{pmatrix} \frac{7}{4} & -1 & -1 \\ -1 & \frac{7}{4} & 1 \\ -1 & 1 & \frac{7}{4} \end{pmatrix},$$

where the sign differences on the off-diagonal terms come from maintaining the correct cyclic ordering of the orbitals within the determinants. This gives a one-dimensional eigenspace with eigenvalue $\frac{15}{4}$ spanned by $\psi_1 - \psi_2 - \psi_3$ and a two-dimensional eigenspace with eigenvalue $\frac{3}{4}$ spanned by $a\psi_1 + b\psi_2 + c\psi_3$, $a = b + c$.

We now have a splitting of $V_{1/2} = V_{1/2,3/4} \cup V_{1/2,15/4}$ where

$$\begin{aligned} V_{1/2,3/4} :&= \{|s\overline{s}p_i\rangle, |s\overline{p_i}p_i\rangle, |p_i\overline{p_i}p_j\rangle \mid i \neq j\} \\ &\quad \cup \{a|s\overline{p_i}p_j\rangle + b|p_i\overline{s}p_j\rangle + c|s\overline{p_j}p_i\rangle, \\ &\qquad\qquad a|p_1\overline{p_2}p_3\rangle + b|p_1\overline{p_3}p_2\rangle + c|p_2\overline{p_1}p_3\rangle \mid i < j,\, a = b + c\} \\ V_{1/2,15/4} :&= \{|s\overline{p_i}p_j\rangle - |p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle, |p_1\overline{p_2}p_3\rangle - |p_1\overline{p_3}p_2\rangle - |p_2\overline{p_1}p_3\rangle \mid i < j\}, \end{aligned}$$

where $V_{1/2,3/4}$ is 20-dimensional and $V_{1/2,15/4}$ is 4-dimensional. We now consider the action of $\underline{L}^2$ on these eigenfunctions:

$$\begin{aligned} \underline{L}^2|s\overline{s}p_i\rangle &= 2|s\overline{s}p_i\rangle, \\ \underline{L}^2|s\overline{p_i}p_i\rangle &= 4|s\overline{p_i}p_i\rangle - 2|s\overline{p_j}p_j\rangle - 2|s\overline{p_k}p_k\rangle, \text{ and} \\ \underline{L}^2|p_i\overline{p_i}p_j\rangle &= 4|p_i\overline{p_i}p_j\rangle - 2|p_k\overline{p_k}p_j\rangle, \end{aligned}$$

the second case of which has the same matrix (C.2), and leads to the eigenfunction $|sp_1\overline{p_1}\rangle + |sp_2\overline{p_2}\rangle + |sp_3\overline{p_3}\rangle$ with eigenvalue 0 and the space spanned by $\{a|sp_1\overline{p_1}\rangle + b|sp_2\overline{p_2}\rangle + c|sp_3\overline{p_3}\rangle \mid a + b + c = 0\}$ with eigenvalue 6.

The third case of the above corresponds to diagonalizing the matrix

$$\begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}, \tag{C.5}$$

which leads to eigenfunctions $|p_i\overline{p_i}p_j\rangle + |p_k\overline{p_k}p_j\rangle$ (with eigenvalue 2) and $|p_i\overline{p_i}p_j\rangle - |p_k\overline{p_k}p_j\rangle$ (with eigenvalue 6).

We treat the remaining two-dimensional spaces by noting that $\{a\psi_1 + b\psi_2 + c\psi_3 \mid a = b+c\}$ is spanned by the two functions $\psi_1 + \psi_2$ and $\psi_2 - \psi_3$. We therefore see that

$$\begin{aligned}
\underline{L}^2(|s\overline{p_i}p_j\rangle + |p_i\overline{s}p_j\rangle) &= 4|s\overline{p_i}p_j\rangle + 2|p_i\overline{s}p_j\rangle + 2|s\overline{p_j}p_i\rangle,\\
&= 4(|s\overline{p_i}p_j\rangle + |p_i\overline{s}p_j\rangle) - 2(|p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle),\\
\underline{L}^2(|p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle) &= -4|s\overline{p_i}p_j\rangle + 2|p_i\overline{s}p_j\rangle - 2|s\overline{p_i}p_j\rangle,\\
&= 4(|p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle) - 2(|s\overline{p_i}p_j\rangle + |p_i\overline{s}p_j\rangle),\\
\underline{L}^2(|p_1\overline{p_2}p_3\rangle + |p_1\overline{p_2}p_3\rangle) &= 6(|p_1\overline{p_2}p_3\rangle + |p_1\overline{p_2}p_3\rangle),\\
\underline{L}^2(|p_1\overline{p_3}p_2\rangle - |p_2\overline{p_1}p_3\rangle) &= 6(|p_1\overline{p_3}p_2\rangle - |p_2\overline{p_1}p_3\rangle).
\end{aligned}$$

The first two cases once again lead to the matrix (C.5) and hence to eigenfunctions

$$\begin{aligned}
(|s\overline{p_i}p_j\rangle + |p_i\overline{s}p_j\rangle) + (|p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle) &= 2|p_i\overline{s}p_j\rangle + |p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle,\\
(|s\overline{p_i}p_j\rangle + |p_i\overline{s}p_j\rangle) - (|p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle) &= |s\overline{p_i}p_j\rangle + |s\overline{p_j}p_i\rangle,
\end{aligned}$$

the first of which has eigenvalue 2 and the second eigenvalue 6. This completes the analysis of $V_{1/2,3/4}$ and we now move onto the simpler case of $V_{1/2,15/4}$. Recall that

$$V_{1/2,15/4} := \{|s\overline{p_i}p_j\rangle - |p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle, |p_1\overline{p_2}p_3\rangle - |p_1\overline{p_3}p_2\rangle - |p_2\overline{p_1}p_3\rangle \mid i < j\}$$

and hence we calculate

$$\begin{aligned}
\underline{L}^2(|s\overline{p_i}p_j\rangle - |p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle) &= 2(|s\overline{p_i}p_j\rangle - |p_i\overline{s}p_j\rangle - |s\overline{p_j}p_i\rangle),\\
\underline{L}^2|p_1\overline{p_2}p_3\rangle - |p_1\overline{p_3}p_2\rangle - |p_2\overline{p_1}p_3\rangle &= 0
\end{aligned}$$

which completes the analysis of $V_{1/2}$ for $\underline{L}^2$.

We now consider the case where $s = \frac{3}{4}$ and have the 4-dimensional space

$$V_{3/2} := \{|sp_ip_j\rangle, |p_1p_2p_3\rangle \mid i < j\}.$$

Applying $\underline{S}^2$ we find that $\underline{S}^2|\psi_1\psi_2\psi_3\rangle = \frac{15}{4}|\psi_1\psi_2\psi_3\rangle$ and hence all four Slater determinants are eigenfunctions of $\underline{S}^2$ with eigenvalue $\frac{15}{4}$. Considering angular momentum we see that

$$\underline{L}^2|sp_ip_j\rangle = 2|sp_ip_j\rangle \text{ and } \underline{L}^2|p_1p_2p_3\rangle = 0,$$

which completes the analysis of the Boron (and, by duality, Nitrogen) $\underline{L}^2$-$\underline{S}^2$-S_3 eigenspaces.

We move on to consider the L_3 eigenspaces. For the case with $\underline{L}^2\Psi = 0$, $\underline{S}^2\Psi = \frac{3}{4}\Psi$ and $s = \frac{1}{2}$ the degeneracy is one and the diagonalization is trivial For the space with $\underline{L}^2\Psi = 0$, $\underline{S}^2\Psi = \frac{15}{4}\Psi$ and $s = \frac{3}{2}$ the degeneracy is also one and the diagonalization is once again trivial.

For the case with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = \frac{3}{4}\Psi$ and $s = \frac{1}{2}$, the degeneracy is three and, since L_3 commutes with the number of p-orbitals (direct calculation or Lemma 3.3.10), L_3 does not couple Slater determinants with different numbers of p-orbitals. We therefore have three submatrices, which, with ordered bases

$$\{|s\overline{s}p_1\rangle, |s\overline{s}p_2\rangle, |s\overline{s}p_3\rangle\},$$

$$\{|p_1p_2\overline{p_2}\rangle + |p_1p_3\overline{p_3}\rangle, |p_2p_1\overline{p_1}\rangle + |p_2p_3\overline{p_3}\rangle, |p_3p_1\overline{p_1}\rangle + |p_3p_2\overline{p_2}\rangle\},$$

$$\{2|\overline{s}p_2p_3\rangle - |s\overline{p_2}p_3\rangle - |sp_2\overline{p_3}\rangle, 2|\overline{s}p_1p_3\rangle - |s\overline{p_1}p_3\rangle - |sp_1\overline{p_3}\rangle, 2|\overline{s}p_1p_2\rangle - |s\overline{p_1}p_2\rangle - |sp_1\overline{p_2}\rangle\},$$

are each of the form $(C.1)$ and hence the three eigenfunctions with L_3 eigenvalue zero are $|s\overline{s}p_3\rangle$, $|p_3p_1\overline{p_1}\rangle + |p_3p_2\overline{p_2}\rangle$, and $2|\overline{s}p_1p_2\rangle - |s\overline{p_1}p_2\rangle - |sp_1\overline{p_2}\rangle$.

For the space with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = \frac{15}{4}\Psi$ and $s = \frac{3}{2}$, the degeneracy is one, and with the ordered basis $\{|sp_1p_3\rangle, |sp_2p_3\rangle, |sp_1p_2\rangle\}$, the L_3 matrix is given by $(C.1)$ and $|sp_1p_2\rangle$ has eigenvalue zero.

For the final case with $\underline{L}^2\Psi = 6\Psi$, $\underline{S}^2\Psi = \frac{3}{4}\Psi$ and $s = \frac{1}{2}$, the degeneracy is two, and once again considering the two submatrices of Slater determinants with one two or three p-orbitals respectively, we have the two ordered bases

$$\begin{aligned}\{|sp_1\overline{p_2}\rangle - |s\overline{p_1}p_2\rangle, |sp_1\overline{p_3}\rangle - |s\overline{p_1}p_3\rangle, |sp_2\overline{p_3}\rangle - |s\overline{p_2}p_3\rangle,& \\ |sp_1\overline{p_1}\rangle - |sp_2\overline{p_2}\rangle, |sp_2\overline{p_2}\rangle - |sp_3\overline{p_3}\rangle\},&\end{aligned}$$

and

$$\begin{aligned}\{|p_3p_2\overline{p_2}\rangle - |p_3p_1\overline{p_1}\rangle, |p_1p_2\overline{p_2}\rangle - |p_1p_3\overline{p_3}\rangle, |p_2p_1\overline{p_1}\rangle - |p_2p_3\overline{p_3}\rangle,& \\ |p_3p_1\overline{p_2}\rangle - |p_3\overline{p_1}p_2\rangle, |\overline{p_3}p_1p_2\rangle - |p_3p_1\overline{p_2}\rangle\},&\end{aligned}$$

both of which result in the L_3 matrix (C.4). It follows that $2|sp_3\overline{p_3}\rangle - |sp_1\overline{p_1}\rangle - |sp_2\overline{p_2}\rangle$ and $2|\overline{p_3}p_1p_2\rangle - |p_3\overline{p_4}p_5\rangle - |p_3p_4\overline{p_5}\rangle$ are eigenfunctions with L_3 eigenvalue zero. This completes the analysis of the Boron (and Nitrogen) eigenspaces.

C.4 Carbon

We see that there are five possible S_3 eigenvalues for Slater determinants forming Carbon wavefunctions, namely $-2, -1, 0, 1$ and 2 and we need only consider $0, 1$ and 2. The $s = 0$ space is 36-dimensional and is given by

$$\begin{aligned} V_0 := \{&|s\overline{s}p_i\overline{p_i}\rangle, |s\overline{s}p_i\overline{p_j}\rangle, |s\overline{p_i}p_j\overline{p_j}\rangle, |p_i\overline{s}p_j\overline{p_j}\rangle, |p_i\overline{p_i}p_j\overline{p_k}\rangle \mid i \neq j \neq k\} \\ &\cup \{|s\overline{p_i}p_j\overline{p_k}\rangle, |p_i\overline{s}p_k\overline{p_j}\rangle, |p_i\overline{p_i}p_k\overline{p_k}\rangle \mid i \neq j \neq k, i < k\}. \end{aligned}$$

As in the previous cases we begin by considering the total spin and for $\psi_i \neq \psi_j$ we have three cases to consider, namely $|\psi_1\overline{\psi_1}\psi_2\overline{\psi_2}\rangle$, $|\psi_1\overline{\psi_1}\psi_2\overline{\psi_3}\rangle$ and $|\psi_1\overline{\psi_2}\psi_3\overline{\psi_4}\rangle$. Applying $\underline{S}^2$ we have

$$\begin{aligned} \underline{S}^2|\psi_1\overline{\psi_1}\psi_2\overline{\psi_2}\rangle &= 0, \\ \underline{S}^2|\psi_1\overline{\psi_1}\psi_2\overline{\psi_3}\rangle &= |\psi_1\overline{\psi_1}\psi_2\overline{\psi_3}\rangle - |\psi_1\overline{\psi_1}\psi_3\overline{\psi_2}\rangle, \text{ and} \\ \underline{S}^2|\psi_1\overline{\psi_2}\psi_3\overline{\psi_4}\rangle &= 2|\psi_1\overline{\psi_2}\psi_3\overline{\psi_4}\rangle - |\psi_2\overline{\psi_1}\psi_3\overline{\psi_4}\rangle - |\psi_1\overline{\psi_2}\psi_4\overline{\psi_3}\rangle \\ &\qquad - |\psi_3\overline{\psi_1}\psi_4\overline{\psi_2}\rangle - |\psi_1\overline{\psi_3}\psi_2\overline{\psi_4}\rangle. \end{aligned}$$

The second case leads to the matrix

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

which clearly has eigenvectors $|\psi_1\overline{\psi_1}\psi_2\overline{\psi_3}\rangle + |\psi_1\overline{\psi_1}\psi_3\overline{\psi_2}\rangle$ (with eigenvalue 0) and $|\psi_1\overline{\psi_1}\psi_2\overline{\psi_3}\rangle - |\psi_1\overline{\psi_1}\psi_3\overline{\psi_2}\rangle$ (with eigenvalue 2). The third case leads to the 6-dimensional matrix

$$\begin{pmatrix} 2 & -1 & -1 & 0 & -1 & -1 \\ -1 & 2 & 0 & -1 & +1 & +1 \\ -1 & 0 & 2 & -1 & +1 & +1 \\ 0 & -1 & -1 & 2 & -1 & -1 \\ -1 & +1 & +1 & -1 & 2 & 0 \\ -1 & +1 & +1 & -1 & 0 & 2 \end{pmatrix}$$

which has been given in the ordered basis

$$\begin{aligned} &\{\Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5, \Psi_6\} := \\ &\{|\psi_1\overline{\psi_2}\psi_3\overline{\psi_4}\rangle, |\psi_1\overline{\psi_2}\psi_4\overline{\psi_3}\rangle, |\psi_2\overline{\psi_1}\psi_3\overline{\psi_4}\rangle, |\psi_2\overline{\psi_1}\psi_4\overline{\psi_3}\rangle, |\psi_1\overline{\psi_3}\psi_2\overline{\psi_4}\rangle, |\psi_3\overline{\psi_1}\psi_4\overline{\psi_2}\rangle\}. \end{aligned}$$

Diagonalizing this matrix leads to a splitting into three eigenspaces, the first has eigenvalue 6 and has the single eigenvector $\Psi_1 - \Psi_2 - \Psi_3 + \Psi_4 - \Psi_5 - \Psi_6$. There is also a 2-dimensional eigenspace with eigenvalue 0 spanned by

$$\{a(\Psi_1+\Psi_4)+b(\Psi_2+\Psi_3)+c(\Psi_5+\Psi_6)\mid a=b+c\}.$$

Finally, there is a 3-dimensional eigenspace with eigenvalue 2 spanned by $\{\Psi_1-\Psi_4,\Psi_2-\Psi_3,\Psi_5-\Psi_6\}$.

As before, we may now split V_0 into $V_0=V_{0,0}\cup V_{0,2}\cup V_{0,6}$ where the spaces are given by

$$\begin{aligned}V_{0,0}:=\{&|s\overline{s}p_i\overline{p_i}\rangle,|p_i\overline{p_i}p_j\overline{p_j}\rangle,|s\overline{s}p_i\overline{p_j}\rangle+|s\overline{s}p_j\overline{p_i}\rangle,\\&|p_k\overline{p_k}p_i\overline{p_j}\rangle+|p_k\overline{p_k}p_j\overline{p_i}\rangle,|p_k\overline{p_k}p_i\overline{s}\rangle+|p_k\overline{p_k}s\overline{p_i}\rangle,\\&a(|s\overline{p_1}p_2\overline{p_3}\rangle+|p_1\overline{s}p_3\overline{p_2}\rangle)+b(|s\overline{p_1}p_3\overline{p_2}\rangle+|p_1\overline{s}p_2\overline{p_3}\rangle)\\&\qquad+c(|s\overline{p_2}p_1\overline{p_3}\rangle+|p_2\overline{s}p_3\overline{p_1}\rangle)\mid i\neq j\neq k,i<j,a=b+c\}\end{aligned}$$

which is of dimension 20,

$$\begin{aligned}V_{0,2}:=\{&|s\overline{s}p_i\overline{p_j}\rangle-|s\overline{s}p_j\overline{p_i}\rangle,|p_k\overline{p_k}p_i\overline{p_j}\rangle+|p_k\overline{p_k}p_j\overline{p_i}\rangle,\\&|p_k\overline{p_k}p_i\overline{s}\rangle-|p_k\overline{p_k}s\overline{p_i}\rangle,|s\overline{p_i}p_k\overline{p_j}\rangle-|p_i\overline{s}p_j\overline{p_k}\rangle\mid i\neq j\neq k,i<j\},\end{aligned}$$

which is of dimension 15 and the remaining one-dimensional space

$$\begin{aligned}V_{0,6}:=\{&(|s\overline{p_1}p_2\overline{p_3}\rangle+|p_1\overline{s}p_3\overline{p_2}\rangle)-(|s\overline{p_1}p_3\overline{p_2}\rangle+|p_1\overline{s}p_2\overline{p_3}\rangle)\\&-(|s\overline{p_2}p_1\overline{p_3}\rangle+|p_2\overline{s}p_3\overline{p_1}\rangle)\}.\end{aligned}$$

We now move on to consider angular momentum, and beginning with $V_{0,0}$ we have

$$\begin{aligned}\underline{L}^2|s\overline{s}p_i\overline{p_i}\rangle&=4|s\overline{s}p_i\overline{p_i}\rangle-2|s\overline{s}p_j\overline{p_j}\rangle-2|s\overline{s}p_k\overline{p_k}\rangle,\\\underline{L}^2|p_i\overline{p_i}p_j\overline{p_j}\rangle&=4|p_i\overline{p_i}p_j\overline{p_j}\rangle-2|p_i\overline{p_i}p_k\overline{p_k}\rangle-2|p_j\overline{p_j}p_k\overline{p_k}\rangle,\\\underline{L}^2(|s\overline{s}p_i\overline{p_j}\rangle+|s\overline{s}p_j\overline{p_i}\rangle)&=6(|s\overline{s}p_i\overline{p_j}\rangle+|s\overline{s}p_j\overline{p_i}\rangle),\\\underline{L}^2(|p_k\overline{p_k}p_i\overline{p_j}\rangle+|p_k\overline{p_k}p_j\overline{p_i}\rangle)&=6(|p_k\overline{p_k}p_i\overline{p_j}\rangle+|p_k\overline{p_k}p_j\overline{p_i}\rangle),\text{ and}\\\underline{L}^2(|pi\overline{p_i}p_j\overline{s}\rangle+|p_i\overline{p_i}s\overline{p_j}\rangle)&=4(|p_i\overline{p_i}p_j\overline{s}\rangle+|p_i\overline{p_i}s\overline{p_j}\rangle)\\&\qquad-2(|p_k\overline{p_k}p_j\overline{s}\rangle+|p_k\overline{p_k}s\overline{p_j}\rangle).\end{aligned}$$

The first two cases are both of the form $|\psi_1\overline{\psi_1}\psi_2\overline{\psi_2}\rangle$ and both lead to the matrix (C.2) (in the ordered basis $\{|\psi_1\overline{\psi_1}\psi_2\overline{\psi_2}\rangle,|\psi_1\overline{\psi_1}\psi_3\overline{\psi_3}\rangle,|\psi_2\overline{\psi_2}\psi_3\overline{\psi_3}\rangle\}$), leading to one eigenfunction with eigenvalue zero, namely $|s\overline{s}p_1\overline{p_1}\rangle+|s\overline{s}p_2\overline{p_2}\rangle+|s\overline{s}p_3\overline{p_3}\rangle$ or $|p_1\overline{p_1}p_2\overline{p_2}\rangle+|p_1\overline{p_1}p_3\overline{p_3}\rangle+|p_2\overline{p_2}p_3\overline{p_3}\rangle$. The other eigenspace is 2-dimensional with eigenvalue 6 and is given by $a|s\overline{s}p_1\overline{p_1}\rangle + b|s\overline{s}p_2\overline{p_2}\rangle + c|s\overline{s}p_3\overline{p_3}\rangle$ or $a|p_1\overline{p_1}p_2\overline{p_2}\rangle + b|p_1\overline{p_1}p_3\overline{p_3}\rangle + c|p_2\overline{p_2}p_3\overline{p_3}\rangle$ where in both cases $a+b+c=0$.

The final case leads to the matrix (C.5) giving the two eigenfunctions $(|p_i\overline{p_i}p_j\overline{s}\rangle + |p_i\overline{p_i}s\overline{p_j}\rangle) + (|p_k\overline{p_k}p_j\overline{s}\rangle + |p_k\overline{p_k}s\overline{p_j}\rangle)$ (with eigenvalue 2) and $(|p_i\overline{p_i}p_j\overline{s}\rangle + |p_i\overline{p_i}s\overline{p_j}\rangle) - (|p_k\overline{p_k}p_j\overline{s}\rangle + |p_k\overline{p_k}s\overline{p_j}\rangle)$ (with eigenvalue 6).

It remains to consider the two-dimensional subspace of $V_{0,0}$ for which we choose the basis functions $(|s\overline{p_1}p_2\overline{p_3}\rangle + |p_1\overline{s}p_3\overline{p_2}\rangle) + (|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle)$ and $(|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle) - (|s\overline{p_2}p_1\overline{p_3}\rangle + |p_2\overline{s}p_3\overline{p_1}\rangle)$, and applying $\underline{L}^2$ gives

$$\begin{aligned}
&\underline{L}^2\big((|s\overline{p_1}p_2\overline{p_3}\rangle + |p_1\overline{s}p_3\overline{p_2}\rangle) + (|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle)\big)\\
&\quad = 6\big((|s\overline{p_1}p_2\overline{p_3}\rangle + |p_1\overline{s}p_3\overline{p_2}\rangle) + (|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle)\big),\\
&\underline{L}^2\big((|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle) - (|s\overline{p_2}p_1\overline{p_3}\rangle + |p_2\overline{s}p_3\overline{p_1}\rangle)\big)\\
&\quad = 6\big((|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle) - (|s\overline{p_2}p_1\overline{p_3}\rangle + |p_2\overline{s}p_3\overline{p_1}\rangle)\big),
\end{aligned}$$

showing that the whole of the 2-dimensional eigenspace has eigenvalue 6.

We now move on to consider the fifteen spin eigenfunctions in $V_{0,2}$ and applying $\underline{L}^2$ gives

$$\begin{aligned}
\underline{L}^2(|s\overline{s}p_i\overline{p_j}\rangle - |s\overline{s}p_i\overline{p_j}\rangle) &= 2(|s\overline{s}p_i\overline{p_j}\rangle - |s\overline{s}p_i\overline{p_j}\rangle),\\
\underline{L}^2(|p_k\overline{p_k}p_i\overline{p_j}\rangle + |p_k\overline{p_k}p_j\overline{p_i}\rangle) &= 2(|p_k\overline{p_k}p_i\overline{p_j}\rangle + |p_k\overline{p_k}p_j\overline{p_i}\rangle),
\end{aligned}$$

both of which are already angular momentum eigenfunctions. The remaining spin eigenfunctions give

$$\begin{aligned}
&\underline{L}^2(|p_k\overline{p_k}p_i\overline{s}\rangle - |p_k\overline{p_k}s\overline{p_i}\rangle)\\
&\qquad = 4(|p_k\overline{p_k}p_i\overline{s}\rangle - |p_k\overline{p_k}s\overline{p_i}\rangle) - 2(|p_j\overline{p_j}p_i\overline{s}\rangle - |p_j\overline{p_j}s\overline{p_i}\rangle),\\
&\underline{L}^2(|s\overline{p_i}p_j\overline{p_k}\rangle - |p_i\overline{s}p_k\overline{p_j}\rangle)\\
&\qquad = 4(|s\overline{p_i}p_j\overline{p_k}\rangle - |p_i\overline{s}p_k\overline{p_j}\rangle) - 2(|s\overline{p_j}p_i\overline{p_k}\rangle - |p_j\overline{s}p_k\overline{p_i}\rangle)\\
&\qquad\qquad - 2(|s\overline{p_i}p_k\overline{p_j}\rangle - |p_i\overline{s}p_j\overline{p_k}\rangle).
\end{aligned}$$

The first case has a six dimensional matrix, which splits into three submatrices, all of the form (C.5, and hence has eigenfunctions $(|p_k\overline{p_k}p_i\overline{s}\rangle - |p_k\overline{p_k}s\overline{p_i}\rangle) - (|p_j\overline{p_j}p_i\overline{s}\rangle - |p_j\overline{p_j}s\overline{p_i}\rangle)$ (with eigenvalue 6) and $(|p_k\overline{p_k}p_i\overline{s}\rangle - |p_k\overline{p_k}s\overline{p_i}\rangle) + (|p_j\overline{p_j}p_i\overline{s}\rangle - |p_j\overline{p_j}s\overline{p_i}\rangle)$ (with eigenvalue 2). The second case has matrix (C.2), which once again leads to one eigenfunction with eigenvalue 0,

$$\begin{aligned}&-\left(|s\overline{p_1}p_2\overline{p_3}\rangle - |p_1\overline{s}p_3\overline{p_2}\rangle\right) + \left(|s\overline{p_2}p_1\overline{p_3}\rangle - |p_2\overline{s}p_3\overline{p_1}\rangle\right)\\ &\qquad + \left(|s\overline{p_1}p_3\overline{p_2}\rangle - |p_1\overline{s}p_2\overline{p_3}\rangle\right),\end{aligned}$$

and a 2-dimensional space with eigenvalue 6 given by

$$\begin{aligned}&-a\left(|s\overline{p_1}p_2\overline{p_3}\rangle - |p_1\overline{s}p_3\overline{p_2}\rangle\right) + b\left(|s\overline{p_1}p_3\overline{p_2}\rangle - |p_1\overline{s}p_2\overline{p_3}\rangle\right)\\ &\quad + c\left(|s\overline{p_2}p_1\overline{p_3}\rangle + |p_2\overline{s}p_3\overline{p_1}\rangle\right)\end{aligned}$$

where $a + b + c = 0$.

Finally we need to consider the space $V_{0,6}$ and find that

$$\begin{aligned}&\underline{L}^2[(|s\overline{p_1}p_2\overline{p_3}\rangle + |p_1\overline{s}p_3\overline{p_2}\rangle) - (|s\overline{p_1}p_3\overline{p_2}\rangle + |p_1\overline{s}p_2\overline{p_3}\rangle)\\ &\qquad - (|s\overline{p_2}p_1\overline{p_3}\rangle + |p_2\overline{s}p_3\overline{p_1}\rangle)] = 0.\end{aligned}$$

We now move on to consider the case where $s = 1$, which has a 16-dimensional basis given by

$$V_1 := \{|s\overline{s}p_ip_j\rangle, |\overline{p_k}sp_ip_j\rangle, |p_i\overline{p_i}sp_k\rangle, |p_k\overline{p_k}p_ip_j\rangle, |\overline{s}p_1p_2p_3\rangle \mid i \neq j \neq k, i < j\}.$$

Again, starting by forming the spin eigenfunctions, we see that there are two types of spin Slater determinants, $|\psi_1\overline{\psi_1}\psi_2\psi_3\rangle$ and $|\psi_1\overline{\psi_2}\psi_3\psi_4\rangle$ for $\psi_i \neq \psi_j$. We have

$$\begin{aligned}\underline{S}^2|\psi_1\overline{\psi_1}\psi_2\psi_3\rangle &= 2|\psi_1\overline{\psi_1}\psi_2\psi_3\rangle\\ \underline{S}^2|\overline{\psi_1}\psi_2\psi_3\psi_4\rangle &= 3|\overline{\psi_1}\psi_2\psi_3\psi_4\rangle - |\overline{\psi_2}\psi_1\psi_3\psi_4\rangle + |\overline{\psi_3}\psi_1\psi_2\psi_4\rangle - |\overline{\psi_4}\psi_1\psi_2\psi_3\rangle.\end{aligned}$$

The second case gives the matrix

$$\begin{pmatrix} 3 & -1 & 1 & -1\\ -1 & 3 & -1 & 1\\ 1 & -1 & 3 & -1\\ -1 & 1 & -1 & 3 \end{pmatrix}$$

in the ordered basis

$$\{|\overline{\psi_1}\psi_2\psi_3\psi_4\rangle, |\overline{\psi_2}\psi_1\psi_3\psi_4\rangle, |\overline{\psi_3}\psi_1\psi_2\psi_4\rangle, |\overline{\psi_4}\psi_1\psi_2\psi_3\rangle\},$$

which gives a one-dimensional eigenspace with eigenvalue 6, namely

$$V_{1,6} := \{|\overline{s}p_1p_2p_3\rangle - |\overline{p_1}sp_2p_3\rangle + |\overline{p_2}sp_1p_3\rangle - |\overline{p_3}sp_1p_2\rangle\}$$

and a 3-dimensional space with eigenvalue 2 spanned by

$$\{|\overline{s}p_1p_2p_3\rangle + |\overline{p_1}sp_2p_3\rangle, |\overline{s}p_1p_2p_3\rangle - |\overline{p_2}sp_1p_3\rangle, |\overline{s}p_1p_2p_3\rangle + |\overline{p_3}sp_1p_2\rangle\}.$$

Combining this eigenspace with the other eigenfunctions with eigenvalue 2 we have the 15-dimensional space

$$\begin{aligned} V_{1,2} : = \{&|s\overline{s}p_ip_j\rangle, |p_i\overline{p_i}sp_k\rangle, |p_k\overline{p_k}p_ip_j\rangle, |\overline{s}p_1p_2p_3\rangle + |\overline{p_1}sp_2p_3\rangle, \\ &|\overline{s}p_1p_2p_3\rangle - |\overline{p_2}sp_1p_3\rangle, |\overline{s}p_1p_2p_3\rangle + |\overline{p_3}sp_1p_2\rangle \mid i \neq j \neq k, i < j\} \end{aligned}$$

For $V_{1,6}$ we find that

$$\underline{L}^2(|\overline{s}p_1p_2p_3\rangle - |\overline{p_1}sp_2p_3\rangle + |\overline{p_2}sp_1p_3\rangle - |\overline{p_3}sp_1p_2\rangle) = 0,$$

and for $V_{1,2}$ we have

$$\begin{aligned} \underline{L}^2|s\overline{s}p_ip_j\rangle &= 2|s\overline{s}p_ip_j\rangle, \\ \underline{L}^2|p_i\overline{p_i}sp_j\rangle &= 4|p_i\overline{p_i}sp_j\rangle - 2|p_k\overline{p_k}sp_j\rangle, \text{ and} \\ \underline{L}^2|p_i\overline{p_i}p_jp_k\rangle &= 2|p_i\overline{p_i}p_jp_k\rangle. \end{aligned}$$

The middle case once again gives the matrix (C.5) and hence has eigenfunctions $|p_i\overline{p_i}sp_j\rangle + |p_k\overline{p_k}sp_j\rangle$ (with eigenvalue 2) and $|p_i\overline{p_i}sp_j\rangle - |p_k\overline{p_k}sp_j\rangle$ (with eigenvalue 6). We are left with the 3-dimensional subspace

$$\{|\overline{s}p_1p_2p_3\rangle + |\overline{p_1}sp_2p_3\rangle, |\overline{s}p_1p_2p_3\rangle - |\overline{p_2}sp_1p_3\rangle, |\overline{s}p_1p_2p_3\rangle + |\overline{p_3}sp_1p_2\rangle\}.$$

and find that

$$\begin{aligned} \underline{L}^2(|\overline{s}p_1p_2p_3\rangle + |\overline{p_1}sp_2p_3\rangle) &= 4|\overline{p_1}sp_2p_3\rangle + 2|\overline{p_2}sp_1p_3\rangle - 2|\overline{p_3}sp_1p_2\rangle, \\ \underline{L}^2(|\overline{s}p_1p_2p_3\rangle - |\overline{p_2}sp_1p_3\rangle) &= -4|\overline{p_2}sp_1p_3\rangle - 2|\overline{p_1}sp_2p_3\rangle - 2|\overline{p_3}sp_1p_2\rangle, \\ \underline{L}^2(|\overline{s}p_1p_2p_3\rangle + |\overline{p_3}sp_1p_2\rangle) &= 4|\overline{p_3}sp_1p_2\rangle - 2|\overline{p_1}sp_2p_3\rangle + 2|\overline{p_2}sp_1p_3\rangle. \end{aligned}$$

Defining $\Psi_1 := |\overline{s}p_1p_2p_3\rangle + |\overline{p_1}sp_2p_3\rangle$, $\Psi_2 := |\overline{s}p_1p_2p_3\rangle - |\overline{p_2}sp_1p_3\rangle$ and $\Psi_3 := |\overline{s}p_1p_2p_3\rangle + |\overline{p_3}sp_1p_2\rangle$ this gives the matrix (C.2), which once again results in a one-dimensional eigenspace with eigenvalue zero, $\Psi_1 + \Psi_2 + \Psi_3$ and a 2-dimensional eigenspace with eigenvalue 6 of the form $\{a\Psi_1 + b\Psi_2 + c\Psi_3 \mid a = b + c\}$.

Finally we need to consider the case $s = 2$ which is the single Slater determinant $|sp_1p_2p_3\rangle$ and is an eigenfunction of both $\underline{S}^2$ (with eigenvalue 6) and $\underline{L}^2$ (with eigenvalue 0).

Moving on to consider the L_3 eigenspaces, we begin with the case where $\underline{L}^2\Psi = 0$, $\underline{S}^2\Psi = 0$ and $s = 0$, the degeneracy is two and we see that both $\underline{L}^2$-$\underline{S}^2$-S_3 eigenfunctions have L_3 eigenvalue zero. The space where $\underline{L}^2\Psi = 0$, $\underline{S}^2\Psi = 2\Psi$ and $s = 1$ has degeneracy one and the $\underline{L}^2$-$\underline{S}^2$-S_3 eigenfunction has L_3 eigenvalue zero. Similarly, the case where $\underline{L}^2\Psi = 0$, $\underline{S}^2\Psi = 6\Psi$ and $s = 2$ has degeneracy one and the $\underline{L}^2$-$\underline{S}^2$-S_3 eigenfunction has L_3 eigenvalue zero.

For the space with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = 0$ and $s = 0$, the degeneracy is one and, with an ordered basis of $\{|s\overline{p_i}p_{i-1}\overline{p_{i-1}}\rangle - |\overline{s}p_ip_{i-1}\overline{p_{i-1}}\rangle + |s\overline{p_i}p_{i+1}\overline{p_{i+1}}\rangle - |\overline{s}p_i\overline{p_{i+1}}p_{i+1}\rangle, i = 1, 2, 3\}$, the L_3 matrix is given by (C.1) and the eigenfunction with $i = 1$ has L_3 eigenvalue zero.

For the space with $\underline{L}^2\Psi = 2\Psi$, $\underline{S}^2\Psi = 2\Psi$ and $s = 1$, the degeneracy is three. Using the fact that L_3 does not mix states with different numbers of p-orbitals, we have three ordered bases

$$\{|s\overline{s}p_1p_3\rangle, |s\overline{s}p_2p_3\rangle, |s\overline{s}p_1p_2\rangle\},$$

$$\{|sp_1p_2\overline{p_2}\rangle + |sp_1p_3\overline{p_3}\rangle, |sp_2p_1\overline{p_1}\rangle + |sp_2p_3\overline{p_3}\rangle, |sp_3p_1\overline{p_1}\rangle + |sp_3p_2\overline{p_2}\rangle\},$$

and

$$\{|p_1p_2p_3\overline{p_1}\rangle, |p_1p_2p_3\overline{p_2}\rangle, |p_1p_2p_3\overline{p_3}\rangle\},$$

each of which have L_3 matrix (C.1), and hence the three $\underline{L}^2$-$\underline{S}^2$-S_3-L_3 eigenfunctions with L_3 eigenvalue zero are $|s\overline{s}p_1p_2\rangle$, $|p_1p_2p_3\overline{p_3}\rangle$, and $|sp_3p_1\overline{p_1}\rangle + |sp_3p_2\overline{p_2}\rangle$.

For the space with $\underline{L}^2\Psi = 6\Psi$, $\underline{S}^2\Psi = 0$ and $s = 0$, the degeneracy is once again three, giving three bases with 2, 3 and 4 p-orbitals respectively:

$$\{|s\overline{s}p_1\overline{p_2}\rangle - |s\overline{s}\overline{p_1}p_2\rangle, |s\overline{s}p_1\overline{p_3}\rangle - |s\overline{s}\overline{p_1}p_3\rangle, |s\overline{s}p_2\overline{p_3}\rangle - |s\overline{s}\overline{p_2}p_3\rangle,$$
$$|s\overline{s}p_1\overline{p_1}\rangle - |s\overline{s}p_2\overline{p_2}\rangle, |s\overline{s}p_2\overline{p_2}\rangle - |s\overline{s}p_3\overline{p_3}\rangle\},$$

$$\begin{aligned}\{&|s\overline{p_3}p_2\overline{p_2}\rangle - |s\overline{p_3}p_1\overline{p_1}\rangle + |\overline{s}p_3p_2\overline{p_2}\rangle - |\overline{s}p_3p_1\overline{p_1}\rangle,\\ &|s\overline{p_1}p_2\overline{p_2}\rangle - |s\overline{p_1}p_3\overline{p_3}\rangle + |\overline{s}p_1p_2\overline{p_2}\rangle - |\overline{s}p_1p_3\overline{p_3}\rangle,\\ &|s\overline{p_2}p_1\overline{p_1}\rangle - |s\overline{p_2}p_3\overline{p_3}\rangle + |\overline{s}p_2p_1\overline{p_1}\rangle - |\overline{s}p_2p_3\overline{p_3}\rangle,\\ &|s\overline{p_3}p_1\overline{p_2}\rangle + |\overline{s}p_3p_1\overline{p_2}\rangle - |s\overline{p_3p_1}p_2\rangle - |\overline{s}p_3\overline{p_1}p_2\rangle,\\ &|sp_3\overline{p_1p_1}\rangle + |\overline{sp_3}p_1p_2\rangle - |s\overline{p_3}p_1\overline{p_2}\rangle - |\overline{s}p_3p_1\overline{p_2}\rangle\},\end{aligned}$$

and

$$\begin{aligned}\{&|p_3\overline{p_3}p_1\overline{p_2}\rangle - |p_3\overline{p_3p_1}p_2\rangle, |p_1\overline{p_1}p_2\overline{p_3}\rangle - |p_1\overline{p_1p_2}p_3\rangle, |p_2\overline{p_2}p_1\overline{p_3}\rangle - |p_2\overline{p_2p_1}p_3\rangle\\ &|p_3\overline{p_3}p_1\overline{p_1}\rangle - |p_3\overline{p_3}p_2\overline{p_2}\rangle, |p_1\overline{p_1}p_2\overline{p_2}\rangle - |p_3\overline{p_3}p_1\overline{p_1}\rangle\},\end{aligned}$$

each of which give an L_3 matrix of the form (C.4) and the $\underline{L}^2$-$\underline{S}^2$-S_3-L_3 eigenfunctions with L_3 eigenvalue zero are $2|s\overline{s}p_3\overline{p_3}\rangle - |s\overline{s}p_1\overline{p_1}\rangle - |s\overline{s}p_2\overline{p_2}\rangle$, $2|p_1\overline{p_1}p_2\overline{p_2}\rangle - |p_3\overline{p_3}p_1\overline{p_1}\rangle - |p_3\overline{p_3}p_2\overline{p_2}\rangle$, and $2|sp_3\overline{p_1p_2}\rangle - |s\overline{p_3}p_1\overline{p_2}\rangle - |s\overline{p_3p_1}p_2\rangle + 2|\overline{sp_3}p_1p_2\rangle - |\overline{s}p_3p_1\overline{p_2}\rangle - |\overline{s}p_3\overline{p_1}p_2\rangle$.

The final case has $\underline{L}^2\Psi = 6\Psi$ and $\underline{S}^2\Psi = 2\Psi$ and $s = 1/2$, which has degeneracy one and, with ordered basis

$$\begin{aligned}\{&|sp_3p_2\overline{p_2}\rangle - |sp_3p_1\overline{p_1}\rangle, |sp_1p_2\overline{p_2}\rangle - |sp_1p_3\overline{p_3}\rangle, |sp_2p_1\overline{p_1}\rangle - |sp_2p_3\overline{p_3}\rangle,\\ &|sp_3p_1\overline{p_2}\rangle - |sp_3\overline{p_1}p_2\rangle, |s\overline{p_3}p_1p_2\rangle - |sp_3p_1\overline{p_2}\rangle\},\end{aligned}$$

has L_3 matrix (C.4) and hence the $\underline{L}^2$-$\underline{S}^2$-S_3-L_3 eigenfunction with L_3 eigenvalue zero is $2|s\overline{p_3}p_1p_2\rangle - |sp_3p_1\overline{p_2}\rangle - |sp_3\overline{p_1}p_2\rangle$.

Appendix D

Validity of a Critical Point of ρ_2^C

In this appendix we will prove that the final critical point in Table 6.2 is not valid for the Carbon ground state pair density.

Lemma D.1. *Let* V *be an arbitrary space and let* $f, g : V \to \mathbb{R}$ *be two functions. Then, for* $v \in V$*, with* $f(v) \in [0,1]$ *and* $g(v) \in [0,1]$*, all four of* $f(v)g(v)$, $\big(1 - f(v)\big)g(v)$, $f(v)\big(1 - g(v)\big)$ *and* $\big(1 - f(v)\big)\big(1 - g(v)\big)$ *are also in* $[0,1]$.

In particular, if for each $v \in V$*, at least one of these four quantities lie outside* $[0,1]$*, then the the intersection of the preimages of* $[0,1]$ *under* f *and* g *is empty, i.e.* $f^{-1}([0,1]) \cap g^{-1}([0,1]) = \varnothing$.

Proof The first part is a trivial consequence that $a, b \in [0,1] \Rightarrow ab \in [0,1]$ and the fact that $a \in [0,1] \Leftrightarrow (1-a) \in [0,1]$. The second part is simply the contrapositive of the first part. □

We now move on to consider the specific functions t and p which are given by

$$t(r,\lambda,c) := \frac{\alpha_2\alpha_4 - 2\alpha_1\alpha_5}{4\alpha_1\alpha_3 - \alpha_2^2}, \quad p(r,\lambda,c) := \frac{2\alpha_3\alpha_4 - \alpha_2\alpha_5}{2\alpha_1\alpha_5 - \alpha_2\alpha_4}. \tag{D.1}$$

Inserting the values of α_i from the end of Section 6.7.1 and then inserting the eight combinations of the ϵ's there are six cases in which t is defined. We wish to show that, in each of these cases, at least one of tp, $(1-t)p$, $t(1-p)$ and $(1-t)(1-p)$ lie outside $[0,1]$ for all $r \in \mathbb{R}$, $\lambda \in [0,1/2]$ and for a suitable range of c.

We begin by labelling these six cases as t_i and p_i, $i = 1, \ldots, 6$, and note further than t_i and p_i have the general forms

$$t_i =: a_{i,1}\frac{A_1(r)}{A_3(r)} + a_{i,2}\frac{A_2(r)}{A_3(r)} + a_{i,3}, \tag{D.2}$$

$$p_i =: \frac{f_i}{g_i} =: \frac{b_{i,1}A_1(r) + b_{i,2}A_2(r) + b_{i,3}A_3(r)}{c_{i,1}A_1(r) + c_{i,2}A_2(r) + c_{i,3}A_3(r)}, \tag{D.3}$$

where $a_{i,j}$, $b_{i,j}$ and $c_{i,j}$ are coefficients in λ and c.

The general method used to treat the first two cases is the same, namely we want to show that $t_i \leq 1 \Rightarrow p_i > 1$.

Lemma D.2. *Suppose t_i and p_i are as given in (D.2) and (D.3) with $A_i(r) \geq 0$. Denote $d_{i,j} := b_{i,j} - c_{i,j}$, $j = 1, 2, 3$ and $e_{i,j} := d_{i,j} - \frac{d_{i,3}a_{i,j}}{a_{i,3}-1}$, $j = 1, 2$ Suppose further that $a_{i,3} < 1$, $c_{i,j} > 0$, and $d_{i,3} > 0$. Then if $e_{i,1} > 0$ and $e_{i,2} > 0$ it follows that t_i and p_i are never simultaneously in $[0, 1]$.*

Proof Suppose that $t_i \leq 1$ then we have, using that $A_3(r) \geq 0$,

$$\begin{aligned} t_i \leq 1 &\Leftrightarrow a_{i,1}A_1(r) + a_{i,2}A_2(r) + (a_{i,3} - 1)A_3(r) \leq 0 \\ &\Leftrightarrow A_3(r) \geq -\frac{a_{i,1}A_1(r) + a_{i,2}A_2(r)}{a_{i,3} - 1}, \end{aligned} \tag{D.4}$$

where the second line uses that $a_{i,3} < 1$. Similarly, since $c_{i,j} > 0$ we have $g_i > 0$, giving

$$p_i > 1 \Leftrightarrow \frac{f_i}{g_i} > 1 \Leftrightarrow f_i - g_i > 0.$$

From the definitions of f_i and g_i in (D.3) it is clear that

$$f_i - g_i = d_{i,1}A_1(r) + d_{i,2}A_2(r) + d_{i,3}A_3(r), \tag{D.5}$$

where $d_{i,j} := b_{i,j} - c_{i,j}$. Since $d_{i,3} \geq 0$, it follows from (D.4) and (D.5) that

$$\begin{aligned} f_i - g_i &\geq \left(d_{i,1} - \frac{d_{i,3}a_{i,1}}{a_{i,3} - 1}\right) A_1(r) + \left(d_{i,2} - \frac{d_{i,3}a_{i,2}}{a_{i,3} - 1}\right) A_2(r) \\ &=: e_{i,1}A_1(r) + e_{i,2}A_2(r). \end{aligned}$$

Hence, since $A_i(r) \geq 0$, it follows that, if $e_{i,1}, e_{i,2} > 0$ then $p_i > 1$, giving the result. □

It remains to check that the conditions required by the above Lemma hold for t_1, p_1, t_2 and p_2, where the first case has $\epsilon_{c\theta} = -1, \epsilon_{c\phi} = 1, \epsilon_{s\phi} = 1$ and the second has $\epsilon_{c\theta} = 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = -1$. For the first case we find that the coefficients for

t_1 are given by

$$a_{1,1} = 1 - \frac{1}{2c^2}, \quad a_{1,2} = \frac{1}{2} + \frac{\sqrt{1-c^2}}{c}, \quad a_{1,3} = \frac{3}{2} - \frac{1}{2c^2},$$

and those for p_1 by

$$\begin{aligned} b_{1,1} &= 10c^2(2c^2-1)\lambda^2 + 2(-10c^4 - 2 + 11c^2)\lambda, \\ b_{1,2} &= 2c^3(5c + 2\sqrt{1-c^2})\lambda^2 + 2c(-3c^3 + 2(2-c^2)\sqrt{1-c^2} + 2c)\lambda, \\ b_{1,3} &= 2c^2(5c^2-3)\lambda^2 + 2(7c^2 - 2 - 5c^4)\lambda, \end{aligned}$$

and

$$c_{1,1} = 2c^2(2c^2-1)\lambda^2, \quad c_{1,2} = 2c^3(c + 2\sqrt{1-c^2})\lambda^2, \quad c_{1,3} = 2c^2(3c^2-1)\lambda^2.$$

Further we see that

$$d_{1,3} = b_{1,3} - c_{1,3} = 2\lambda(c+1)(1-c)(5c^2 - 2\lambda c^2 - 2).$$

Checking the conditions in Lemma D.2 we see that $a_{1,3} - 1 = (c^2-1)/(2c^2)$ which is clearly negative when $c < 1$. It is also clear that, if $c > 1/\sqrt{2}$ then each of the $c_{1,j}$ is positive. Finally we have, using that $(5 - 2\lambda) \geq 4$, $d_{1,3} > 0$ if $c > 1/\sqrt{2}$. Hence it remains to check that $e_{1,1}$ and $e_{1,2}$ are both positive. We have that

$$e_{1,1} = 4\lambda c^2\big(\lambda(2c^2-1) + 1\big), \quad e_{1,2} = 4\lambda c^2\big(c^2(1+\lambda) + c\sqrt{1-c^2}(4-2\lambda)\big),$$

both of which are clearly positive for $\lambda > 0$ and $c > 1/\sqrt{2}$. If $\lambda = 0$ we find that $g_1 = 0$ and hence p_1 is not defined. Thus we see that t_1 and p_1 are not simultaneously in $[0,1]$.

For the second case we find that the coefficients for t_2 are given by

$$a_{2,1} = 3 - \frac{3}{2c^2}, \quad a_{2,2} = \frac{3}{2} - \frac{\sqrt{1-c^2}}{c}, \quad a_{2,3} = \frac{3}{2} - \frac{1}{2c^2},$$

and those for p_2 by

$$\begin{aligned} b_{2,1} &= 14c^2(2c^2-1)\lambda^2 + 2(-14c^4 - 6 + 17c^2)\lambda, \\ b_{2,2} &= 2c^3(7c - 2\sqrt{1-c^2})\lambda^2 + 2c(-9c^3 - 2(2-c^2)\sqrt{1-c^2} + 6c)\lambda, \\ b_{2,3} &= 2c^2(5c^2-3)\lambda^2 + 2(7c^2 - 2 - 5c^4)\lambda, \end{aligned}$$

and

$$c_{2,1} = 6c^2(2c^2 - 1)\lambda^2, \quad c_{2,2} = 2c^3(3c - 2\sqrt{1-c^2})\lambda^2, \quad c_{2,3} = 2c^2(3c^2 - 1)\lambda^2.$$

Note that $a_{2,3} = a_{1,3}$, $b_{2,3} = b_{1,3}$ and $c_{2,3} = c_{1,3}$ and hence it remains to check that $c_{2,1}$ and $c_{2,2}$ are positive. This is clearly the case if $c > 1/\sqrt{2}$ and in order to prove the result we must show that $e_{2,1}$ and $e_{2,2}$ are positive where

$$\begin{aligned} e_{2,1} &= 4c^2(1 - 2c^2)\lambda^2 + 8(5c^4 - 4c^2 + 1)\lambda, \\ e_{1,2} &= 4c^3(2\sqrt{1-c^2} - c)\lambda^2 + 8c(-c + 3c^3 + (2 - 3c^2)\sqrt{1-c^2})\lambda. \end{aligned}$$

Again, if $\lambda = 0$, $g_2 = 0$ and hence we need only consider $\lambda \in (0, 1/2]$. It is easy to show that, on $c \in [1/\sqrt{2}, 1]$, the coefficients of λ^2 in $e_{2,1}$ and $e_{2,2}$ are decreasing and those corresponding to λ are increasing. Hence we have, for $c < 1$,

$$e_{2,1} > 2\lambda(1 - 2\lambda) \geq 0, \quad e_{2,2} > 4\lambda(1 - \lambda) \geq 0.$$

where the above inequalities hold for $\lambda \in [0, 1/2]$. Hence Lemma D.2 applies and we see that t_2 and p_2 are never simultaneously in $[0, 1]$.

The next case we consider has $\epsilon_{c\theta} = 1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = 1$ for which apply Lemma D.1 and in particular wish to show that $p_3 t_3 < 0$. For the coefficients of t_3 we have

$$\begin{aligned} a_{3,1} &= \frac{2 - (6 + \lambda)c^2 + 2\lambda c^4}{2(1 - 2\lambda)c^4}, \\ a_{3,2} &= \frac{-2c^2 + (\lambda - 2)c^4 + 2(c^2 - 2)c\sqrt{1-c^2}}{2(1 - 2\lambda)c^4}, \\ a_{3,3} &= \frac{2 - (4 + \lambda)c^2 + (2 - \lambda)c^4}{2(1 - 2\lambda)c^4}. \end{aligned}$$

For the coefficients of f_3 and g_3 we have

$$\begin{aligned} b_{3,1} &= 4c^2(1 - 2c^2)\lambda^2 + 2(4 - 15c^2 + 3c^4)\lambda - 4 + 14c^2 - 4c^4, \\ b_{3,2} &= 4c^2(-c^2 - 2c\sqrt{1-c^2})\lambda^2 + 2(-4c^2 - c^4 + (6c^2 - 8)c\sqrt{1-c^2})\lambda \\ &\quad + 4c^2 + 2c^4 + 4(2 - c^2)c\sqrt{1-c^2}, \\ b_{3,3} &= 4c^2(1 - 3c^2)\lambda^2 + 2(4 - 11c^2 + 9c^4)\lambda - 4 + 10c^2 - 6c^4, \end{aligned}$$

and

$$\begin{aligned}
c_{3,1} &= 4c^2(2c^2-1)\lambda^2 + 2(4-11c^2-2c^4)\lambda - 4 + 12c^2,\\
c_{3,2} &= 4c^2(c^2+2c\sqrt{1-c^2})\lambda^2 + 2(-4c^2-5c^4-2(c^2+4)c\sqrt{1-c^2})\\
&\quad + 4c^2 + 4c^4 + 8c\sqrt{1-c^2},\\
c_{3,3} &= 4c^2(-1-c^2)\lambda^2 2(4-7c^2+5c^4)\lambda - 4 + 8c^2 - 4c^4.
\end{aligned}$$

From these coefficients it follows that

$$t_3p_3 = \frac{1}{2c^4(1-2\lambda)A_3(r)}\big(d_{3,1}A_1(r) + d_{3,2}A_2(r) + d_{3,3}A_3(r)\big)$$

where

$$\begin{aligned}
d_{3,1} &= 2 + (\lambda-7)c^2 + 2(1-\lambda)c^4,\\
d_{3,2} &= -2c^2 + (-1-\lambda)c^4 + (2c^2(1-\lambda)-4)c\sqrt{1-c^2},\\
d_{3,3} &= (2 + (\lambda-5)c^2 + (-3\lambda+3)c^4.
\end{aligned}$$

This has a simpler form than a general product of two functions of the forms given in (D.2) and (D.3) because the numerator of t_i and denominator of p_i are the same up to sign, as can be seen in (D.1). It follows that t_3p_3 has the same sign as $d_{3,1}A_1(r) + d_{3,2}A_2(r) + d_{3,3}A_3(r)$, and it is easy to see that, for $c > 1/\sqrt{2}$, each of the d_i are linearly decreasing in λ and hence

$$d_{3,1} \le 2 - 7c^2 + 2c^4, \quad d_{3,2} \le -2c^2 - c^4 + (2c^2-4)c\sqrt{1-c^2}, \quad d_{3,3} \le 2 - 5c^2 + 3c^4,$$

from which it is clear that each of the $d_{3,i}$ are negative if $c > \sqrt{2/3} > 1/\sqrt{2}$. Thus, by Lemma D.1, for $c > \sqrt{2/3}$, t_3 and p_3 are never simultaneously in $[0,1]$.

For the fourth case, which corresponds to $\epsilon_{c\theta} = -1, \epsilon_{c\phi} = -1, \epsilon_{s\phi} = 1$, we will use Lemma D.1 and in particular show that if $t_4p_4 \ge 0$ then $(1-p_4)t_4 > 0$. The following lemma summarizes the method used:

Lemma D.3. *Suppose* $A_i(r) \ge 0$, $i = 1,2,3$ *and let* $F = d_1A_1(r) + d_2A_2(r) + d_3A_3(r)$ *with* $d_3 \le 0$ *and* $G = e_1A_1(r) + e_2A_2(r) + e_3A_3(r)$ *with* $e_3 > 0$*. Define* $f_i := d_i - d_3e_i/e_3$ *for* $i = 1,2$*. If both* f_1 *and* f_2 *are negative then* $G \ge 0 \Rightarrow F < 0$

Proof The proof is analogous to that of Lemma D.2. ☐

In our case we will have $F = k_1(1-p_4)t_4$ and $G = k_2t_4p_4$, where k_i are positive functions determined later. It remains to show that the coefficients satisfy the necessary conditions of the lemma. We have, for t_4,

$$\begin{aligned}
a_{4,1} &= \frac{6-(10+7\lambda)c^2+14\lambda c^4}{2c^4(1-\lambda)},\\
a_{4,2} &= \frac{-6c^2+(2+7\lambda)c^4+2(2-c^2\lambda)c\sqrt{1-c^2}}{2c^4(1-\lambda)},\\
a_{4,1} &= \frac{4-(4+3\lambda)c^2+(2+3\lambda)c^4}{2c^4(1-\lambda)},
\end{aligned}$$

and for p_4,

$$\begin{aligned}
b_{4,1} &= 6c^2[(3c^2-1)\lambda^2+(2-4c^2)\lambda+2c^2-1],\\
b_{4,2} &= 2c^2[(3c^2-2c\sqrt{1-c^2})\lambda^2+(4c\sqrt{1-c^2}-6c^2)\lambda+3c^2-3c\sqrt{1-c^2}],\\
b_{4,3} &= 2c^2(c^2-1)(1-\lambda)^2,
\end{aligned}$$

and

$$\begin{aligned}
c_{4,1} &= 14c^2(2-c^2)\lambda^2+2(6-3c^2-14c^4)\lambda-12+20c^2,\\
c_{4,2} &= 2c^2(7c^2-2c\sqrt{1-c^2})\lambda^2+2(-6c^2-5c^4+2(c^2+2)c\sqrt{1-c^2})\lambda\\
&\quad +12c^2-4c^4-8c\sqrt{1-c^2},\\
c_{4,3} &= 6c^2(c^2-1)\lambda^2+2(2-c^2-c^4)\lambda-4+8c^2-4c^4.
\end{aligned}$$

Using these coefficients to evaluate $(1-p_4)t_4$ we find $2c^4A_3(r)(1-p_4)t_4 =: d_{4,1}A_1(r)+d_{4,2}A_2(r)+d_{4,3}A_3(r)$, where we note that $2c^4A_3(r)\geq 0$ and hence this expression has the same sign as $(1-p_4)t_4$ and will be our F. The $d_{4,i}$ are given explicitly by

$$\begin{aligned}
d_{4,1} &= 6+(-13-4\lambda)c^2+(6+8\lambda)c^4,\\
d_{4,2} &= -6c^2+2(2-c^2)c\sqrt{1-c^2}+(5+4\lambda)c^4,\\
d_{4,3} &= (c^2-1)(c^2(3+2\lambda)-2),
\end{aligned}$$

where we note that, for $c>\sqrt{2/3}$, $d_{4,3}$ is clearly negative.

We also have that $2c^2A_3(r)t_4p_4 =: e_{4_1}A_1(r)+e_{4,2}A_2(r)+e_{4,3}A_3(r)$, where this has the same sign as t_4p_4 and the coefficients are given by

$$e_{4,1}=3(1-2c^2),\quad e_{4,2}=2c\sqrt{1-c^2}-3c^2,\quad e_{4,3}=1-c^2.$$

It is clear that $e_{4,3} > 0$. Using the definition of $f_{4,i}$ from Lemma D.3, we have

$$\begin{aligned} f_{4,1} &= 2c^2(1-2c^2)\lambda + 4c^2(2-3c^2), \text{ and} \\ f_{4,2} &= 2c^2(2c\sqrt{1-c^2}-c^2)\lambda + 4c^2(c\sqrt{1-c^2}-c^2). \end{aligned}$$

It is clear that each of the above terms is negative when $c > 2/\sqrt{5} > \sqrt{2/3}$ and hence, for this range of c, the lemma holds and, by Lemma D.1, t_4 and p_4 are never simultaneous in $[0,1]$.

For the fifth case, which corresponds to $\epsilon_{c\theta} = 1, \epsilon_{c\phi} = 1, \epsilon_{s\phi} = -1$, we have, for the coefficients in t_5,

$$\begin{aligned} a_{5,1} &= \frac{2-(6+5\lambda)c^2+10\lambda c^4}{2c^4(1-\lambda)}, \\ a_{5,2} &= \frac{-2c^2+(5\lambda-2)c^4+2(c^2-2)c\sqrt{1-c^2}}{2c^4(1-\lambda)}, \\ a_{5,3} &= \frac{2-(4+3\lambda)c^2+(2+3\lambda)c^4}{2c^4(1-\lambda)}. \end{aligned}$$

For the coefficients in f_5 and g_5 we have

$$\begin{aligned} b_{5,1} &= 2c^2(1-\lambda)^2(2c^2-1), \\ b_{5,2} &= 2c^2(1-\lambda)^2(c^2+2c\sqrt{1-c^2}), \\ b_{5,3} &= 2c^2(1-\lambda)(c^2-1), \end{aligned}$$

and

$$\begin{aligned} c_{5,1} &= 2(1-\lambda)(-2+(6+5\lambda)c^2+10\lambda c^4), \\ c_{5,2} &= 2(1-\lambda)(2c^2+(2-5\lambda)c^4+2(2-c^2)c\sqrt{1-c^2}), \\ c_{5,3} &= 2(1-\lambda)(1-c^2)(-2+(2+3\lambda)c^2). \end{aligned}$$

From these expressions it follows that $t_5(1-p_5)$ has the same sign as $2A_3(r)(1-\lambda)c^4 t_5(1-p_5) =: d_{5,1}A_1(r) + d_{5,2}A_2(r) + d_{5,3}A_3(r)$ where

$$\begin{aligned} d_{5,1} &= 4c^2(2c^2-1)\lambda + 2 - 7c^2 + 2c^4, \\ d_{5,2} &= 4c^4\lambda - 2c^2 - c^4 + 2(c^2-2)c\sqrt{1-c^2}, \\ d_{5,3} &= (c^2-1)(-2+3c^2+2\lambda c^2). \end{aligned}$$

It is clear that all three of these expressions are negative when $c > \sqrt{2/3}$ and therefore, by Lemma D.1, for $c > \sqrt{2/3}$, t_5 and p_5 are never simultaneously in $[0,1]$.

For the final case, which corresponds to $\epsilon_{c\theta} = -1, \epsilon_{c\phi} = 1, \epsilon_{s\phi} = -1$, we use Lemma D.1 and in particular show that $t_6 p_6$ is always negative. The coefficients for t_6 are

$$\begin{aligned}
a_{6,1} &= \frac{6 - (10 + 2\lambda)c^2 + 6\lambda c^4}{(1 - 2\lambda)c^4}, \\
a_{6,2} &= \frac{-6c^2 + (2 + 3\lambda)c^4 + 2(2 - c^2)c\sqrt{1 - c^2}}{(1 - 2\lambda)c^4}, \\
a_{6,2} &= \frac{2 - (4 + \lambda)c^2 - \lambda c^4}{(1 - 2\lambda)c^4}.
\end{aligned}$$

Also, the coefficients for f_6 and g_6 are

$$\begin{aligned}
b_{6,1} &= 12c^2(1 - 2c^2)\lambda^2 + 2(6 - 29c^2 + 18c^4)\lambda - 12 + 26c^2 - 12c^4, \\
b_{6,2} &= 4c^2(2c\sqrt{1 - c^2} - 3c^2)\lambda^2 + 2(-12c^2 + 13c^4 + (2 - 6c^2)c\sqrt{1 - c^2})\lambda \\
&\quad + 12c^2 - 10c^2 + 4(c^2 - 2)c\sqrt{1 - c^2}, \\
b_{6,3} &= 4c^2(1 - 3c^2)\lambda^2 + 2(4 - 11c^2 + 9c^4)\lambda - 4 + 10c^2 - 6c^4,
\end{aligned}$$

and

$$\begin{aligned}
c_{6,1} &= 12c^2(2c^2 - 1)\lambda^2 + 2(12 - 17c^2 - 6c^4)\lambda - 12 + 20c^2, \\
c_{6,2} &= 4c^2(3c^2 - 2c\sqrt{1 - c^2})\lambda^2 + 2(-12c^2 + c^4 + (2c^2 + 8)c\sqrt{1 - c^2})\lambda \\
&\quad + 12c^2 - 4c^4 - 8c\sqrt{1 - c^2}, \\
c_{6,3} &= 4c^2(-1 - c^2)\lambda^2 + 2(4 - 7c^2 + 5c^4)\lambda - 4 + 8c^2 - 4c^4.
\end{aligned}$$

Using these coefficients, we have that $t_6 p_6$ has the same sign as $2A_3(r)c^4(1 - 2\lambda)t_6 p_6 =: d_{6,1}A_1(r) + d_{6,2}A_2(r) + d_{6,3}A - 3(r)$, where

$$\begin{aligned}
d_{6,1} &= 3c^2(1 - 2c^2)\lambda + 6 - 13c^2 + 6c^4, \\
d_{6,2} &= c^2(2c\sqrt{1 - c^2} - 3c^2)\lambda - 6c^2 + 5c^4 + 2(2 - 2c^2)c\sqrt{1 - c^2}, \\
d_{6,3} &= c^2(1 - 3c^2)\lambda + 2 - 5c^2 + 3c^4.
\end{aligned}$$

It is clear that, if $c > 1/\sqrt{2}$, then all the coefficients of λ are negative and hence each is bounded above by the value when $\lambda = 0$. This gives

$$d_{6,1} \leq 6 - 13c^2 + 6c^4,\ d_{6,2} \leq -6c^2 + 5c^4 + 2(2 - c^2)c\sqrt{1 - c^2},\ d_{6,3} \leq 2 - 5c^2 + 3c^4,$$

and it is easy to check that all of these are negative if $c > \sqrt{2/3} > 1/\sqrt{2}$. Hence, by Lemma D.1, for $c > \sqrt{2/3}$, t_6 and p_6 are never simultaneously in $[0, 1]$.

Hence we see that the strongest restriction on c in all of the above is $c > 2/\sqrt{5}$, and for values of c in this range the final critical point in Table 6.2 never exists. In particular it is worth noting that the value of c in our Carbon variational calculation is approximately $0.9857649926 > 2/\sqrt{5}$ and hence this result holds for our pair density.

Appendix E

Comparison of Critical Points of ρ_2^C

The are clearly four different types of comparison necessary, two for each of the cases of $\rho_2^{C,max}$, corresponding critical points from Table 6.2 with or without existence conditions. The following five lemmas give conditions under which $\rho_2^{C,max}$ is greater than each of the other critical points from the previous section.

The first lemma compares two critical points, both of which always exist:

Lemma E.1. *Suppose that* $A_i(r) \geq 0$ *and let*

$$\begin{aligned} \rho_2^{(1)} &= e_{1,1}A_1(r) + e_{1,2}A_2(r) + e_{1,3}A_3(r), \\ \rho_2^{(2)} &= e_{2,1}A_1(r) + e_{2,2}A_2(r) + e_{2,3}A_3(r) \text{ and} \\ f &= -\frac{1}{\alpha A_3(r)}(\beta A_1(r) + \gamma A_2(r) + \delta A_3(r)), \end{aligned}$$

with $-\alpha A_3(r) > 0$*,* $\delta > 0$*, and* $f < 0$*. Denote* $\tilde{e}_i := e_{1,i} - e_{2,i}$*,* $i = 1, 2, 3$*, and suppose* $\tilde{e}_3 < 0$*. If* $h_1 := \delta\tilde{e}_1 - \tilde{e}_3\beta \geq 0$ *and* $h_2 := \delta\tilde{e}_2 - \tilde{e}_3\gamma \geq 0$*, then* $\rho_2^1 > \rho_2^2$*.*

Proof Using the definition of $\tilde{e}_i$ in the statement of then lemma, it is clear that

$$\rho_2^{(1)} - \rho_2^{(2)} = \tilde{e}_1 A_1(r) + \tilde{e}_2 A_2(r) + \tilde{e}_3 A_3(r). \tag{E.1}$$

Since $f < 0$, $-\alpha A_3(r) > 0$ and $\delta > 0$, it follows that

$$A_3(r) < -\frac{\beta A_1(r) + \gamma A_2(r)}{\delta},$$

and, using $\tilde{e}_3 < 0$, we have

$$\tilde{e}_3 A_3(r) > -\tilde{e}_3 \frac{\beta A_1(r) + \gamma A_2(r)}{\delta}.$$

Inserting this bound into (E.1) we obtain

$$\rho_2^{(1)} - \rho_2^{(2)} > \left(\tilde{e}_1 - \tilde{e}_3\frac{\beta}{\delta}\right)A_1(r) + \left(\tilde{e}_2 - \tilde{e}_3\frac{\gamma}{\delta}\right)A_2(r),$$

and the result follows from the positivity of δ and non-negativity of $A_1(r)$ and $A_2(r)$. ☐

The second lemma will enable us to compare $\rho_2^{C,max} = A_0(r) + g(r)$ and a critical point $\rho_2^{(2)}$ with an existence condition f_2, under the condition that $f_3 < 0$. The restriction on f_2 gives an upper bound on $\rho_2^{(2)}$ which is linear in the $A_i(r)$ and the restriction on f_3 then allows us to reduce the problem to considering the sign of two coefficients.

Lemma E.2. *Suppose that* $A_i(r) \geq 0$ *and let*

$$\begin{aligned}
\rho_2^{(1)} &= e_{1,1}A_1(r) + e_{1,2}A_2(r) + e_{1,3}A_3(r),\\
\rho_2^{(2)} &= -\frac{(\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r))^2}{2\alpha_2 A_3(r)} + g_{2,1}A_1(r) + g_{2,2}A_2(r) + g_{2,3}A_3(r),\\
f_2 &= -\frac{(\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r))}{\alpha_2 A_3(r)}, \quad \textit{and}\\
f_3 &= -\frac{(\beta_3 A_1(r) + \gamma_3 A_2(r) + \delta_3 A_3(r))}{\alpha_3 A_3(r)}.
\end{aligned}$$

Suppose further that $-\alpha_2 A_3(r) > 0$, $-\alpha_3 A_3(r) > 0$, $\delta_3 > 0$, $f_2 \in [0, 1]$, $f_3 < 0$, *and* $2e_{1,3} - (\delta_2 + 2g_{2,3}) \leq 0$. *If* $h_1 \geq 0$ *and* $h_2 \geq 0$, *where*

$$\begin{aligned}
h_1 &= (2e_{1,1} - (\beta_2 + 2g_{2,1}))\delta_3 - (2e_{1,3} - (\delta_2 + 2g_{2,3}))\beta_3,\\
h_2 &= (2e_{1,1} - (\gamma_2 + 2g_{2,1}))\delta_3 - (2e_{1,3} - (\delta_2 + 2g_{2,3}))\gamma_3,
\end{aligned}$$

it follows that $\rho_2^{(1)} > \rho_2^{(2)}$.

Proof Since $f_2 \geq 0$ and $-\alpha_2 A_3(r) > 0$ it follows that $\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r) \geq 0$. Hence, since $f_2 \leq 1$, $\rho_2^{(2)}$ is bounded above by

$$\rho_2^{(2)} \leq \tfrac{1}{2}((\beta_2 + 2g_{2,1})A_1(r) + (\gamma_2 + 2g_{2,2})A_2(r) + (\delta_2 + 2g_{2,3})A_3(r)).$$

We then have

$$\begin{aligned}
2(\rho_2^{(1)} - \rho_2^{(2)}) \geq{}& [2e_{1,1} - (\beta_2 + 2g_{2,1})]A_1(r) + [2e_{1,2} - (\gamma_2 + 2g_{2,2})]A_2(r)\\
&+ [2e_{1,3} - (\delta_2 + 2g_{2,3})]A_3(r).
\end{aligned} \tag{E.2}$$

Now, since $f_3 < 0$, $-\alpha_3 A_3(r) > 0$, and $\delta_3 > 0$, we have

$$A_3(r) < -\frac{\beta_3 A_1(r) + \gamma_3 A_2(r)}{\delta_3},$$

and since $2e_{1,3} - (\delta_2 + 2g_{2,3}) < 0$ it follows that

$$[2e_{1,3} - (\delta_2 + 2g_{2,3})]A_3(r) > -[2e_{1,3} - (\delta_2 + 2g_{2,3})]\frac{\beta_3 A_1(r) + \gamma_3 A_2(r)}{\delta_3}.$$

Inserting this lower bound into δ_3 times (E.2), and recalling that $\delta_3 > 0$, gives that

$$2\delta_3(\rho_2^{(1)} - \rho_2^{(2)}) > h_1 A_1(r) + h_2 A_2(r),$$

where h_1 and h_2 are as in the statement of the lemma and the result follows. ☐

The third and fourth lemmas allow us to compare the first case of $\rho_2^{C,max}$ from Theorem 6.7.1 with a critical point from Table 6.2 that always exists. In this case it happens that there is no need to use the restriction that $f(r, \lambda) \in [0, 1]$.

Lemma E.3. *Suppose that* $A_i(r) \geq 0$ *and let*

$$\begin{aligned}\rho_2^{(1)} &= -\frac{(\beta A_1(r) + \gamma A_2(r) + \delta A_3(r))^2}{2\alpha A_3(r)} + g_1 A_1(r) + g_2 A_2(r) + g_3 A_3(r),\\ \rho_2^{(2)} &= e_1 A_1(r) + e_2 A_2(r) + e_3 A_3(r).\end{aligned}$$

Denote

$$f_{11} = \beta^2, \quad f_{22} = \gamma^2, \quad f_{33} = \delta^2 + 2\alpha(e_3 - g_3), \quad f_{12} = 2\beta\gamma,$$

$$f_{13} = 2\beta\delta + 2\alpha(e_1 - g_1), \text{ and } f_{23} = 2\gamma\delta + 2\alpha(e_2 - g_2).$$

Let

$$\chi_{ij} = \begin{cases} 1 & \text{if } f_{ij} < 0 \\ 0 & \text{else,} \end{cases}$$

and further denote

$$\tilde{f}_{11} = f_{11} + \tfrac{1}{2}(\chi_{12} f_{12} + \chi_{13} f_{13}), \quad \tilde{f}_{22} = f_{22} + \tfrac{1}{2}(\chi_{12} f_{12} + \chi_{23} f_{23}),$$

$$\tilde{f}_{33} = f_{33} + \tfrac{1}{2}(\chi_{13} f_{13} + \chi_{23} f_{23}), \quad \tilde{f}_{12} = (1 - \chi_{12}) f_{12},$$

$$\tilde{f}_{13} = (1 - \chi_{13}) f_{13}, \text{ and } \tilde{f}_{23} = (1 - \chi_{23}) f_{23}.$$

Suppose further that $-\alpha A_3(r) >$, *then if all of* $\tilde{f}_{ij}$ *are non-negative, it follows that* $\rho_2^{(1)} \geq \rho_2^{(2)}$, *with equality only if all* $\tilde{f}_{ij}$ *are all zero or* $A_i(r) = 0$, $i = 1, 2, 3$.

Proof Since $-\alpha A_3(r) > 0$ it follows that $-2\alpha A_3(r)(\rho_2^{(1)} - \rho_2^{(2)})$ has the same sign as $\rho_2^{(1)} - \rho_2^{(2)}$. Calculating this explicitly gives

$$\begin{aligned}-2\alpha A_3(r)(\rho_2^{(1)} - \rho_2^{(2)}) &= f_{11}A_1^2(r) + f_{22}A_2^2(r) + f_{33}A_3^2(r) \\ &\quad + f_{12}A_1(r)A_2(r) + f_{13}A_1(r)A_3(r) + f_{23}A_2(r)A_3(r),\end{aligned}$$

where the f_{ij} are as given in the statement of the lemma.

Since $(A_i(r) - A_j(r))^2 \geq 0$ then, if $f_{ij} < 0$, it follows that $f_{ij}A_i(r)A_j(r) \geq \frac{f_{ij}}{2}(A_i^2(r) + A_j^2(r))$. Inserting this inequality into the above for each of the f_{ij}, $i \neq j$ which are negative, gives

$$\begin{aligned}-\alpha A_3(r)(\rho_2^{(1)} - \rho_2^{(2)}) &\geq \tilde{f}_{11}A_1^2(r) + \tilde{f}_{22}A_2^2(r) + \tilde{f}_{33}A_3^2(r) \\ &\quad + \tilde{f}_{12}A_1(r)A_2(r) + \tilde{f}_{13}A_1(r)A_3(r) + \tilde{f}_{23}A_2(r)A_3(r),\end{aligned}$$

and the result follows. ☐

For cases where the bound $\rho_2^{(1)} - \rho_2^{(2)} \geq 0$ is tight the previous lemma may not apply and we need to find the critical points explicitly.

Lemma E.4. *Suppose that* $A_1(r), A_2(r) \geq 0$ *and* $A_3(r) > 0$. *Let*

$$\begin{aligned}f &= e_{11}A_1^2(r) + e_{22}A_2^2(r) + e_{33}A_3^2(r) \\ &\quad + e_{12}A_1(r)A_2(r) + e_{13}A_1(r)A_3(r) + e_{23}A_2(r)A_3(r)\end{aligned}$$

and suppose that $e_{11} > 0$, $e_{12} > 0$, $e_{22} > 0$, *and* $e_{33} > 0$. *Suppose further that* $4e_{11}e_{22} - e_{12}^2 = 0$. *Let* $r_1 = -\frac{1}{2}e_{12}e_{13} + e_{11}e_{23}$, $r_2 = -\frac{1}{4}e_{13}^2 + e_{11}e_{33}$, $q_1 := -r_1e_{13} + r_2e_{12}$ *and* $q_2 := -e_{23}^2 + 4e_{22}e_{33}$. *If* $r_1 \leq 0$, $q_1 \geq 0$ *and* $q_2 \geq 0$, *it follows that* $f \geq 0$, *with equality only when* $q_1 = 0$ *or* $q_2 = 0$.

Proof Since $A_3(r) > 0$, dividing throughout by $A_3^2(r)$, we see that f has the same sign as

$$\tilde{f} = e_{11}\alpha^2 + e_{22}\beta^2 + e_{12}\alpha\beta + e_{13}\alpha + e_{23}\beta + e_{33},$$

where $\alpha = A_1(r)/A_3(r)$ and $\beta = A_2(r)/A_3(r)$, $\alpha, \beta \geq 0$.

Fixing β and differentiating with respect to α gives that the critical point lies

$$\alpha_0 = -\frac{e_{12}\beta + e_{13}}{e_{11}}.$$

Since $e_{11} > 0$ it is clear that this must be a minimum and hence the minimum for $\alpha \in [0, \infty)$ lies at $\alpha = \max\{\alpha_0, 0\}$. If $\alpha = \alpha_0$ we find that

$$\tilde{f} = \left(-\frac{e_{12}^2}{4e_{11}} + e_{22}\right)\beta^2 + \left(-\frac{e_{12}e_{13}}{2e_{11}} + e_{23}\right)\beta - \frac{e_{13}^2}{4e_{11}} + e_{33},$$

and using that $4e_{11}e_{22} - e_{12}^2 = 0$, $e_{11} > 0$ and the definitions of r_1 and r_2 we see that

$$\tilde{f} \geq 0 \Leftrightarrow r_1\beta + r_2 \geq 0. \tag{E.3}$$

Since we have assumed that $-\frac{e_{12}\beta+e_{13}}{e_{11}} \geq 0$, using $e_{11}, e_{12} \geq 0$, gives that $\beta \leq -e_{13}/e_{12}$, and since $r_1 \leq 0$ it follows that $r_1\beta \geq -r_1e_{13}/e_{12}$. Inserting this lower bound into (E.3) we have that

$$-\frac{r_1e_{13}}{e_{12}} + r_2 \geq 0 \Rightarrow \tilde{f} \geq 0,$$

and since $e_{12} \geq 0$ this is equivalent to $-r_1e_{13} + r_2e_{12} \geq 0$.

The second case is when $\alpha_0 < 0$, which gives a minimum at $\alpha = 0$ and we obtain $\tilde{f} = e_{22}\beta^2 + e_{23}\beta + e_{33}$. Differentiating with respect to β and using that $e_{22} > 0$ gives a minimum at $\beta_0 = -e_{23}/(2e_{22})$, and so the minimum on $[0, \infty)$ is at $\beta = \max\{\beta_0, 0\}$.

If $\beta = \beta_0$ we have

$$\tilde{f} = -\frac{e_{23}^2}{4e_{22}} + e_{33},$$

and using that $e_{22} > 0$ it follows that

$$\tilde{f} \geq 0 \Leftrightarrow -e_{23}^2 + 4e_{22}e_{33} \geq 0.$$

Finally if the minimum is at $\alpha = 0$, $\beta = 0$ then $f = e_{33} > 0$ and the result follows. □

The final two lemmas allow us to compare the first case of $\rho_2^{C,max}$, denoted by $\rho_2^{(1)}$ with a critical point from Table 6.2, denoted by $\rho_2^{(2)}$, which has an existence condition $f_2 \in [0, 1]$.

There are now two existence conditions which need to be met and it should be possible to derive a joint existence condition. However, this is non-trivial and we instead use the restrictions $f_1 \geq 0$ and $f_2 \leq 1$ and show one of the two lemmas holds for all values of λ. It is worth noting that we could also prove a similar lemma with the restriction $f_2 \geq 0$.

Lemma E.5. *Suppose that* $A_i(r) \geq 0$ *and let*

$$\begin{aligned}
\rho_2^{(1)} &= -\frac{(\beta_1 A_1(r) + \gamma_1 A_2(r) + \delta_1 A_3(r))^2}{2\alpha_1 A_3(r)} + g_{1,1}A_1(r) + g_{1,2}A_2(r) + g_{1,3}A_3(r),\\
\rho_2^{(2)} &= -\frac{(\beta_2 A_1(r) + \beta_2 A_2(r) + \delta_2 A_3(r))^2}{2\alpha_2 A_3(r)} + g_{2,1}A_1(r) + g_{2,2}A_2(r) + g_{2,3}A_3(r),\\
f_1 &= -\frac{(\beta_1 A_1(r) + \gamma_1 A_2(r) + \delta_1 A_3(r))}{\alpha_1 A_3(r)}, \text{ and}\\
f_2 &= -\frac{(\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r))}{\alpha_2 A_3(r)}.
\end{aligned}$$

Suppose further that $-\alpha_1 A_3(r) > 0$, $-\alpha_2 A_3(r) > 0$, $\delta_1 > 0$, $f_1 \geq 0$, $e_{11} := \alpha_1\beta_2^2 - \alpha_2\beta_1^2 \geq 0$, $e_{22} := \alpha_1\gamma_2^2 - \alpha_2\gamma_1^2 \geq 0$, *and* $e_{33} := \alpha_1\delta_2^2 - \delta_2 c_1^2 + 2\alpha_1\alpha_2(g_{1,3} - g_{2,3}) \geq 0$. *Then if* $h_{1,1} > 0$, $h_{1,2} > 0$ *and* $h_{2,2} > 0$ *where*

$$\begin{aligned}
h_{11} &= [2\alpha_2\beta_1\delta_1(g_{1,1} - g_{2,1}) - 2\alpha_2\beta_1^2(g_{1,3} - g_{2,3}) - 2\alpha_2\beta_1\delta_1 - \beta_1^2\delta_2^2 - \beta_2^2\delta_1^2],\\
h_{12} &= [2\alpha_2\gamma_1\delta_1(g_{1,1} - g_{2,1}) - 2\alpha_2\beta_1\delta_1(g_{2,2} - g_{1,2}) - 4\alpha_2\beta_1\gamma_1(g_{1,3} - g_{2,3})\\
&\qquad + 2(\beta_2\delta_1 - \beta_1\delta_2)(\gamma_1\delta_2 - \gamma_2\delta_1)],\\
h_{22} &= [2\alpha_2\gamma_1\delta_1(g_{1,2} - g_{2,2}) - 2\alpha_2\gamma_1^2(g_{1,3} - g_{2,3}) + 2\gamma_1\gamma_2\delta_1\delta_2 - \gamma_2^2\delta_1^2],
\end{aligned}$$

it follows that $\rho_2^{(1)} > \rho_2^{(2)}$.

Proof It is clear that $2\alpha_1\alpha_2A_3(r)(\rho_2^{(1)} - \rho_2^{(2)})$ has the same sign as $\rho_2^{(1)} - \rho_2^{(2)}$. It follows that

$$\begin{aligned}
2\alpha_1\alpha_2A_3(r)(\rho_2^{(1)} - \rho_2^{(2)}) &= e_{11}A_1^2(r) + e_{22}A_2^2(r) + e_{33}A_3^2(r)\\
&\quad + e_{12}A_1(r)A_2(r) + e_{13}A_1(r)A_3(r) + e_{23}A_2(r)A_3(r),
\end{aligned} \tag{E.4}$$

where the e_{ij} are given explicitly by

$$\begin{aligned}
e_{11} &= \alpha_1\beta_2^2 - \alpha_2\beta_1^2,\\
e_{22} &= \alpha_1\gamma_2^2 - \alpha_2\gamma_1^2,\\
e_{33} &= \alpha_1\delta_2^2 - \alpha_2\delta_1^2 + 2\alpha_1\alpha_2(g_{1,3} - g_{2,3}),\\
e_{12} &= 2\alpha_1\beta_2\gamma_2 - 2\alpha_2\beta_1\gamma_1,\\
e_{13} &= 2\alpha_1\beta_2\delta_2 - 2\alpha_2\beta_1\delta_1 + 2\alpha_1\alpha_2(g_{1,1} - g_{2,1}),\\
e_{23} &= 2\alpha_1\gamma_2\delta_2 - 2\alpha_2\gamma_1\delta_1 + 2\alpha_1\alpha_2(g_{1,2} - g_{2,2}).
\end{aligned}$$

By the explicit form of f_1 and the inequalities $f_1 \geq 0$, $-\alpha_1A_3(r) > 0$ and

$\delta_1 > 0$, it follows that

$$A_3(r) \geq -\frac{\beta_1 A_1(r) + \gamma_1 A_2(r)}{\delta_1}. \tag{E.5}$$

Suppose that e_{13}, e_{23} and e_{33} are all non-negative and hence inserting (E.5) into δ_1^2 times (E.4) (which still has the same sign) gives

$$2\delta_1^2 \alpha_1 \alpha_2 A_3(r)(\rho_2^{(1)} - \rho_2^{(2)}) =: \tilde{h}_{11} A_1^2(r) + \tilde{h}_{12} A_1(r) A_2(r) + \tilde{h}_{22} A_2^2(r),$$

where $\tilde{h}_{ij} = -\alpha_1 h_{ij}$ as given the statement of the lemma. The result follows by noting that, since $A_3(r) \geq 0$ we must also have that $-\alpha_1 \geq 0$, and so $\tilde{h_{ij}}$ and h_{ij} have the same sign. □

Lemma E.6. *Suppose that $A_i(r) \geq 0$ and let*

$$\begin{aligned}
\rho_2^{(1)} &= -\frac{(\beta_1 A_1(r) + \gamma_1 A_2(r) + \delta_1 A_3(r))^2}{2\alpha_1 A_3(r)} + g_{1,1} A_1(r) + g_{1,2} A_2(r) + g_{1,3} A_3(r), \\
\rho_2^{(2)} &= -\frac{(\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r))^2}{2\alpha_2 A_3(r)} + g_{2,1} A_1(r) + g_{2,2} A_2(r) + g_{2,3} A_3(r), \\
f_1 &= -\frac{(\beta_1 A_1(r) + \gamma_1 A_2(r) + \delta_1 A_3(r))}{\alpha_1 A_3(r)}, \text{ and} \\
f_2 &= -\frac{(\beta_2 A_1(r) + \gamma_2 A_2(r) + \delta_2 A_3(r))}{\alpha_2 A_3(r)}.
\end{aligned}$$

Suppose further that $-\alpha_1 A_3(r) > 0$, $-\alpha_2 A_3(r) > 0$, $-(\alpha_2 + \delta_2) > 0$, $f_2 \leq 1$, $e_{11} := \alpha_1 b_2^2 - \alpha_2 b_1^2 \geq 0$, $e_{22} := \alpha_1 c_2^2 - \alpha_2 c_1^2 \geq 0$, and $e_{33} := \alpha_1 d_2^2 - \delta_2 c_1^2 + 2\alpha_1 \alpha_2 (g_{1,3} - g_{2,3}) \geq 0$ Then if $h_{1,1} > 0$, $h_{1,2} > 0$ and $h_{2,2} > 0$ where

$$\begin{aligned}
h_{11} = {} & [2\alpha_1 \alpha_2 \beta_2 (g_{1,1} - g_{2,1}) - 2\alpha_1 \beta_2 \delta_2 (g_{2,1} - g_{1,1}) - 2\alpha_1 \beta_2^2 (g_{1,3} - g_{2,3}) \\
& - 2\alpha_2 \beta_1 (\beta_1 \delta_2 - \beta_2 \delta_1) - 2\beta_1 \beta_2 \delta_1 \delta_2 - \alpha_1 \alpha_2 \beta_2^2 + \alpha_2^2 \beta_1^2 + \delta_2^2 \beta_1^2 + \beta_2^2 \delta_1^2], \\
h_{12} = {} & [2\alpha_1 \gamma_2 (\alpha_2 + \delta_2)(g_{1,1} - g_{2,1}) + 2\alpha_1 \beta_2 (\alpha_2 + \delta_2)(g_{1,2} - g_{2,2}) \\
& - 4\alpha_1 \beta_2 \gamma_2 (g_{1,3} - g_{2,3}) - 2\delta_1 (\beta_1 \gamma_2 + \beta_2 \gamma_1)(\alpha_2 + \delta_2) + 2\beta_1 \gamma_1 (\alpha_2^2 + \delta_2^2) \\
& + 2\beta_2 \gamma_2 \delta_1^2 - 2\alpha_1 \alpha_2 \beta_2 \gamma_2 + 4\alpha_2 \beta_1 \gamma_1 \delta_2], \\
h_{22} = {} & [2\alpha_1 \alpha_2 \gamma_2 (g_{1,2} - g_{2,2}) - 2\alpha_1 \gamma_2^2 (g_{1,3} - g_{2,3}) + 2\alpha_1 \gamma_2 \delta_2 (g_{1,2} - g_{2,2}) \\
& + \delta_2^2 \gamma_1^2 + \gamma_2^2 \delta_1^2 - \alpha_1 \alpha_2 \gamma_2^2 + \alpha_2^2 \gamma_1^2 - 2\alpha_2 \gamma_1 \gamma_2 \delta_1 + 2\alpha_2 \gamma_1^2 \delta_2 - 2\gamma_1 \gamma_2 \delta_1 \delta_2],
\end{aligned}$$

it follows that $\rho_2^{(1)} > \rho_2^{(2)}$.

Proof The proof mirrors that of Lemma E.5 but the upper bound on $A_3(r)$ comes from the condition $f_2 \leq 1$ instead of $f_1 \geq 0$ and $-(\alpha_2 + \delta_2) > 0$ replaces $\delta_1 > 0$. □

Bibliography

[AdP01] Atkins, P. and de Paula, J. *Atkins' Physical Chemistry, 7th Ed.* Oxford University Press (2001).

[All72] Allen, L.C. The shape of molecules. *Theoretica Chimica Acta*, **24**, 117–131 (1972).

[AS72] Abramowitz, M. and Stegun, I.A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* National Bureau of Standards Applied Mathematics Series; 55. Washington, D.C.: U.S. Dept. of Commerce (1972).

[BB01] Bellert, D. and Breckenridge, W.H. A spectroscopic determination of the bond length of the LiOLi molecule: Strong ionic bonding. *Journal of Chemical Physics*, **114**(7), 2871–2874 (2001).

[BBBC92] Bunge, C.F., Barrientos, J.A., Bunge, A.V., and Cogordan, J.A. Hartree-Fock and Roothaan-Hartree-Fock energies for the ground states of He through Xe. *Physical Review A*, **46**(7), 3691–3696 (1992).

[Bec93] Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, **98**(7), 5648–5652 (1993).

[Bel78] Bell, S. The effects of basis set and configuration interaction on the predicted geometries of AH_2 molecules. *Journal of Chemical Physics*, **68**(7), 3014–3024 (1978).

[BM06] Bromley, M.W.J. and Mitroy, J. Convergence of the partial wave expansion of the He ground state. *International Journal of Quantum Chemistry* (2006). URL `http://www3.interscience.wiley.com/cgi-bin/abstract/113441513/ABSTRACT`.

[BO27] Born, M. and Oppenheimer, R. Zur Quantentheorie der Molekeln. *Annalen der Physik*, **389**(20), 457–484 (1927).

[BP72] Buenker, R.J. and Peyerimhoff, S.D. Extension of Walsh's rules to more general systems. *Theoretica Chimica Acta*, **24**, 132–146 (1972).

[BSKW63] Büchler, A., Stauffer, J.L., Klemperer, W., and Wharton, L. Determination of the geometry of Lithium Oxide, $Li_2O(g)$, by electric deflection. *The Journal of Chemical Physics*, **39**(9), 2299–2303 (1963).

[BWG05] Burke, K., Werschnik, J., and Gross, E.K.U. Time-dependent density functional theory: Past, present, and future. *The Journal of Chemical Physics*, **123**, 062206 (2005).

[CC98] Chen, E. and Chang, T.C. Walsh diagram and the linear combination of bond orbital method. *Journal of Molecular Structure*, **431**, 127–136 (1998).

[CL62] Calais, J.L. and Löwdin, P.O. A simple method of treating atomic integrals containing functions of r_{12}. *Journal of Molecular Spectroscopy*, **8**, 203–211 (1962).

[Cle65] Clementi, E. Ab initio computations in atoms and molecules. *IBM Journal of Research and Development*, **9**(1), 2–19 (1965).

[CRR67] Clementi, E., Raimondi, D.L., and Reinhardt, W.P. Atomic screening constants from SCF functions. II. Atoms with 37 to 86 electrons. *The Journal of Chemical Physics*, **47**(4), 1300–1307 (1967).

[CS35] Condon, E. and Shortley, G. *The Theory of Atomic Spectra.* Cambridge University Press, London and New York (1935).

[CV66] Clementi, E. and Veillard, A. Correlation energy in atomic systems IV. Degeneracy effects. *The Journal of Chemical Physics*, **44**(8), 3050–3053 (1966).

[DFRSP82] De Frees, D.J., Raghavachari, K., Schlegel, H.B., and Pople, J.A. Effect of electron correlation on theoretical equilibrium geometries. 2. Comparison of third-order perturbation and configuration interaction results with experiment. *Journal of the American Chemistry Society*, **104**(21), 5576–5580 (1982).

[FF69] Froese Fischer, C. A Multi-Configuration Hartree-Fock program. *Computer Physics Communication*, **1**, 151–161 (1969).

[FF72] Froese Fischer, C. A Multi-Configuration Hartree-Fock program with improved stability. *Computer Physics Communication*, **4**, 107–116 (1972).

[FF77] Froese Fischer, C. *The Hartree-Fock Method for Atoms. A Numerical Approach.* Wiley-Interscience (1977).

[FHOHOS05] Fournais, S., Hoffmann-Ostenhof, M., Hoffmann-Ostenhof, T., and Sørensen, T.Ø. Sharp regularity results for Coulombic many-electron wave functions. *Communications in Mathematical Physics*, **255**, 183–227 (2005).

[Foc30] Fock, V. Näherungsmethoden zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik*, **61**, 126–148 (1930).

[Fri02] Friesecke, G. On the infinitude of non-zero eigenvalues of the single-electron density matrix for atoms and molecules. *Proceedings of the Royal Society of London A*, **459**, 47–52 (2002).

[Fri04] Friesecke, G. *Private Communication* (2004).

[Fri05] Friesner, R.A. Chemical theory and computation special feature: Ab initio quantum chemistry: Methodology and applications. *Proceedings of the National Academt of Sciences of the United States of America*, **102**, 6648–6653 (2005).

[Fri07] Friesecke, G. *Lectures on Fourier Analysis.* submitted to Cambridge University Press (2007).

[Fri0X] Friesecke, G. *In preparation* (200X).

[Fro67] Froese, C. Some Hartree-Fock results for two- and four-electron atomic systems. *The Journal of Chemical Physics*, **47**(10), 4010–4015 (1967).

[Gil60] Gillespie, R.J. Electron correlation and molecular shape. *Canadian Journal of Chemistry*, **38**, 818–826 (1960).

[GN57] Gillespie, R.J. and Nyholm, R.S. Inorganic stereochemistry. *Quarterly Reviews of the Chemical Society*, **11**, 339–380 (1957).

[Gri95] Griffiths, D.J. *Introduction to quantum mechanics.* Englewood Cliffs, N.J.: Prentice Hall (1995).

[Har28] Hartree, D.R. The wave mechanics of an atom with a non-Coulomb central field. *Proceedings of the Cambridge Philosophical Society*, **24**, 89–132 (1928).

[HC96a] He, Z. and Cremer, D. Sixth-order many-body perturbation theory. I. Basic theory and derivation of the energy formula. *International Journal of Quantum Chemistry*, **59**, 15–29 (1996).

[HC96b] He, Z. and Cremer, D. Sixth-order many-body perturbation theory. II. Implementation and application. *International Journal of Quantum Chemistry*, **59**, 31–55 (1996).

[HDGP75] Hosteny, R.R., Dunning, T.H., Gilman, R.R., and Pipano, A. Ab initio study of the π-electron states of trans-butadene. *The Journal of Chemical Physics*, **62**, 4746–4779 (1975).

[Her66] Herzberg, G. *Molecular spectra and molecular structure. 3. Electronic spectra and electronic structure of polyatomic molecules.* Princeton, N.J.; London: Van Nostrand (1966).

[HG96] Head-Gordon, M. Quantum chemisty and molecular processes. *The Journal of Chemical Physics*, **100**, 13212–13225 (1996).

[HH36] Hartree, D. and Hartree, W. Self-Consistent Field, with exchange, for Beryllium - the $(2s)(2p)$ ^3P and ^1P excited states. *Proceedings of the Royal Society of London A*, **154**(883), 588–607 (1936).

[HH83] Harrison, R.J. and Handy, N.C. Full CI calculations on BH, H_2O, NH_3 and HF. *Chemical Physics Letters*, **95**(4,5), 386–391 (1983).

[Hil85] Hill, R.N. Rates of convergence and error estimation formulas for the Rayleigh-Ritz variational method. *Journal of Chemical Physics*, **83**(3), 1173–1196 (1985).

[HJ67] Herzberg, G. and Johns, J.W.C. The spectrum and structure of the free BH_2 radical. *Proceedings of the Royal Society of London A*, **298**(1453), 142–159 (1967).

[HK64] Hohenberg, P. and Kohn, W. Inhomogeneous electron gas. *Physical Review*, **136**(3B), 864–871 (1964).

[Huh93] Huheey, J.E. *Inorganic chemistry : principles of structure and reactivity.* Harper Collins (1993).

[IS75] Ivanova, E. and Safranova, U. Perturbation theory in calculations of atomic energy levels. *Journal of Physics B*, **8**(10), 1591–1602 (1975).

[JLH62] Jordan, P.C.H. and Longuet-Higgins, H.C. The lower electronic levels of the radicals CH, CH_2, CH_4, NH, NH_2, BH, BH_2 and BH_3. *Molecular Physics*, **5**, 121–138 (1962).

[Joh32] Johnson, M.H. The theory of complex spectra. part i. energy levels. *Phys. Rev.*, **39**(2), 197–209 (Jan 1932). doi:10.1103/PhysRev.39.197.

[Jos02] Jost, J. *Riemannian geometry and geometric analysis.* Berlin ; New York : Springer (2002).

[Kat57] Kato, T. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, **10**(2), 151–177 (1957).

[Kat67] Kato, T. *Perturbation Theory for Linear Operators.* Springer (1967).

[KB77] Klahn, B. and Bingel, W.A. The convergence of the Rayleigh-Ritz method in quantum chemistry. *Theoretica Chimica Acta*, **44**(1), 27–43 (1977).

[KM92] Kutzelnigg, W. and Morgan, J.D. Rates of convergence of the partial-wave expansions of atomic correlation energies. *Journal of Chemical Physics*, **96**(6), 4484–4508 (1992).

[KNB89] Kucharski, S.A., Noga, J., and Bartlett, R.J. Fifth-order many-body perturbation theory for molecular correlation energies. *Journal of Chemical Physics*, **90**(12), 7282–7290 (1989).

[Kni06] Knight, W. Atomic simulation most intensive computer program ever. *New Scientist Tech* (2006). URL `http://www.newscientisttech.com/channel/tech/dn9394.html`.

[KS65] Kohn, W. and Sham, L.J. Self-Consistent equations including exchange and correlation effects. *Physical Review*, **140**(4A), 1133–1138 (1965).

[KSM99] Kim, Y.H., Städele, M., and Martin, R.M. Density-functional study of small molecules within the Krieger-Li-Iafrate approximation. *Physical Review A*, **60**(5), 3633–3640 (1999).

[LASS00] Leininger, M.L., Allen, W.D., Schaefer, H.F., and Sherrill, C.D. Is Møller-Plesset perturbation theory a convergent ab initio method? *The Journal of Chemical Physics*, **122**(21), 9213–9222 (2000).

[Löw55] Löwdin, P.O. Quantum theory of many-particle systems.I. Physical interpretations by means of density matrices, natural spin-orbitals and convergence problems in the method of Configuration Interaction. *Physical Review*, **97**(6), 1474–1489 (1955).

[LS60] Linderberg, J. and Shull, H. Electron correlation energy in 3- and 4-electron atoms. *Journal of Molecular Spectroscopy*, **5**, 1–16 (1960).

[LYP88] Lee, C., Yang, W., and Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, **37**(2), 785–789 (1988).

[McW89] McWeeny, R. *Methods of molecular quantum mechanics*. London: Academic Press (1989).

[MH65] Moskowitz, J.W. and Harrison, M.C. Gaussian wavefunctions for the 10-electron systems. III. OH^-, H_2O, H_3O^+. *The Journal of Chemical Physics*, **43**(10), 3550–3555 (1965).

[Moo70] Moore, C.E. *Selected Tables of Atomic Spectra (NSRDS-NBS 3)*. Washington, DC: National Bureau of Standards (1970).

[MP34] Møller, C. and Plesset, M.S. Note on an approximation treatment for many-electraon systems. *Physical Reveiw*, **46**, 618–622 (1934).

[Mul55] Mulliken, R.S. Bond angles in water-type and ammonia-type molecules and their derivatives. *Journal of the American Chemical Society*, **77**, 887–891 (1955).

[NIS05] NIST. *NIST Computational Chemistry Comparison and Benchmark DataBase (release 12)*. National Institute of Standards and Technology (2005). URL `http://srdata.nist.gov/cccbdb/`.

[NIS06] NIST. *NIST Atomic Spectra Database (version 3.1.0)*. National Institute of Standards and Technology (2006). URL `http://physics.nist.gov/asd3`.

[OJK+96] Olsen, J., Jørgensen, P., Koch, H., Balkova, A., and Bartlett, R.J. Full configuration-interaction and state of the art correlation calculations on water in a valence double-zeta basis with polarization functions. *The Journal of Chemical Physics*, **104**(20), 8007–8015 (1996).

[Pau31] Pauling, L. The nature of the chemical bond. application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. *Journal of the American Chemical Society*, **53**, 1367–1400 (1931).

[Pau79] Pauncz, R. *Spin Eigenfunctions.* Plenum, New York (1979).

[PTSS04] Perdew, J.P., Tao, J., Staroverov, V.N., and Scuseria, G.E. Meta-generalized gradient approximation: Explanation of a realistic nonempirical density functional. *The Journal of Chemical Physics*, **120**(15), 6898–6911 (2004).

[PY95a] Parr, R.G. and Yang, W. *Density Functional Theory of Atoms and Molecules.* International Series of Monographs on Chemistry. Oxford University Press Inc, USA (1995).

[PY95b] Parr, R.G. and Yang, W. Density-functional theory of the electronic structure of molecules. *Annual Reviews of Physical Chemistry*, **46**, 701–728 (1995).

[Roo60] Roothaan, C.C.J. Self-Consistent Field theory for open shells of electronic systems. *Reviews Of Modern Physics*, **32**(2), 179–185 (1960).

[RS78] Reed, M. and Simon, B. *Methods of Modern Mathematical Physics. Volumes I-IV.* Dover Publications (1978).

[SACKH95] St.-Amant, A., Cornell, W.D., Kollman, P.A., and Halgren, T.A. Calculation of molecular geometries, relative conformational energies, dipole moments, and molecular electrostatic potential fitted charges of small organic molecules of biochemical interest by Density Functional Theory. *Journal of Computational Chemistry*, **16**(12), 1483–1506 (1995).

[SC62] Sharma, C.S. and Coulson, C.A. Hartree-Fock and correlation energies for 1s2s ^{3}S and ^{1}S states of helium-like ions. *Proceedings of the Physical Society*, **80**, 81–96 (1962).

[Sch26] Schrödinger, E. Quantisierung als Eigenwertproblem. *Annalen der Physik*, **79**, 361 (1926).

[Sch62] Schwartz, C. Importance of angular correlations betweeen atomic electrons. *Physical Review*, **126**(3), 1015–1019 (1962).

[Sch01] Schwabl, F. *Quantum Mechanics*. Springer (2001).

[SD73] Stenkamp, L.Z. and Davidson, E.R. An ICSCF investigation of Walsh's rules. *Theoretica Chimica Acta*, **30**, 283–314 (1973).

[Sha62] Sharma, C.S. A perturbation treatment of the Unrestricted Hartree-Fock equations for the ground state of Lithium-like ions. *Proceedings of the Physical Society*, **80**, 839–848 (1962).

[Sla29] Slater, J.C. The theory of complex spectra. *Physical Review*, **34**(10), 1293–1322 (Nov 1929). doi:10.1103/PhysRev.34.1293.

[Sla30] Slater, J.C. Atomic shielding constants. *Physical Review*, **36**(1), 57–64 (1930).

[Sla64] Slater, J.C. Atomic radii in crystals. *The Journal of Chemical Physics*, **41**(10), 3199–3204 (1964).

[SO96] Szabo, A. and Ostlund, N.S. *Modern Quantum Chemistry*. Dover Publications (1996).

[SP40] Sidgwick, N.V. and Powell, H.M. Bakerian lecture. Stereochemical types and valency groups. *Proceedings of the Royal Society of London A*, **176**(965), 153–180 (1940).

[Sug27] Sugiura, Y. Über die Eigenschaften des Wasserstoffmoleküls im Grundzustande. *Zeitschrift für Physik*, **45**, 484–492 (1927).

[SW66] Sharma, C.S. and Wilson, R.G. A perturbation treatment of the Hartree-Fock equations for $2p^2$, ^1D and ^3P weakly quantized states of Helium-like ions. *Physical Review*, **181**(1), 39–42 (1966).

[SW68a] Sharma, C.S. and Wilson, R.G. A perturbation treatment of the Unrestricted Hartree-Fock equations for 1s2p ^3P and ^1P states of Helium-like ions. *Journal of Physics B*, **1**(1), 1–12 (1968).

[SW68b] Sharma, C.S. and Wilson, R.G. A perturbation treatment of the Unrestricted Hartree-Fock equations for 2s2p ^{3}P and ^{1}P weakly quantized states of Helium-like ions. *Molecular Physics*, **15**(3), 271–278 (1968).

[tH75] ter Harr, D. *Problems in Quantum Mechanics.* Pion, London (1975).

[TT74] Tatewaki, H. and Tanaka, K. Correlation holes and wavefunctions for the ^1S, $^1P^0$, and $^3P^0$ states of B, Be$^+$, and C^{2+}. *The Journal of Chemical Physics*, **60**(2), 601–606 (1974).

[TTST94] Tatewaki, H., Toshikatsu, K., Sakai, Y., and Thakkar, A.J. Numerical Hartree-Fock energies of low-lying excited states of neutral atoms with $Z \leq 18$. *Journal of Chemical Physics*, **101**(6), 4945–4948 (1994).

[Uff38] Ufford, C.W. Term values in carbon. *Physical Review*, **53**(7), 568–569 (Apr 1938). doi:10.1103/PhysRev.53.568.

[VM70] Verhaegen, G. and Moser, C.M. Hartree-Fock, correlation and term energies of valence excited states of atoms and ions of the second row of the periodic table. *Journal of Physics B: Atomic and Molecular Physics*, **3**(4), 478–492 (1970).

[VVS98] Van Voorhis, T. and Scuseria, G.E. A novel form for the exchange-correlation energy functional. *The Journal of Chemical Physics*, **109**(2), 400–410 (1998).

[Wal53] Walsh, A.D. The electronic orbitals, shapes, and spectra of polyatomic molecules. Part I. AH$_2$ molecules. *Journal of the Chemical Society*, pages 2260–2266 (1953).

[Wil71] Wilson, R.G. A perturbation treatment of the unrestricted Hartree-Fock equations for the ground states of 3- to 10-electron atoms. *Journal of Physics B: Atomic and Molecular Physics*, **4**, 228–240 (1971).

[ZG29] Zener, C. and Guillemin, V. The B-state of the Hydrogen molecule. *Physical Review*, **34**(7), 999–1009 (1929).

[Zhi60] Zhislin, G.M. Discussion of the spectrum of Schrödinger operators for systems of many particles. *Trudy Moskovskogo matematiceskogo obscestva*, **9**, 81–120 (1960).