Variational Solutions in Orbital Free Density Functional Theory

Matthew Stephen Ryley



A thesis presented for the degree of Doctor of Philosophy School of Chemistry



Dedication

I dedicate this to my wife. Without her encouragement and support I would have definitely carried out my various threats of running away to the mountains and never thinking about OF-DFT again.

Declaration

The material contained in this thesis has not been previously submitted for a degree at the University of Nottingham or any other university.

I do declare that this thesis and the work contained within is my own and has been generated as a result of my own research throughout my PhD studies.

Abstract

This work focusses on testing and developing methods to find solutions to the variational problem at the heart of orbital free density functional theory (OF-DFT). OF-DFT is an inherently linear scaling quantum chemical method which can, in theory, be used to simulate systems with millions of atoms. However, there is a limited choice of OF-DFT codes which are suitable for the development of the theory for chemical systems. In this work we compare and contrast three methods: the Lopez-Acevedo scheme;¹ the Chan, Cohen and Handy Scheme^{2,3}(CCH); and the trust-region image method (TRIM) scheme.⁴ We find that the scheme developed in this work — the TRIM scheme — offers the most efficient methodology for converging the energy for a wide range of functionals in an all electron context for finite chemical systems.

In Chapter 1 some mathematical topics are introduced which are required to understand how the foundations of density functional theory (DFT) built upon convex analysis underpins the variational principles of CCH and TRIM. In Chapter 2 we introduce electronic structure theory and discuss the theoretical foundations of DFT. We also include a discussion on Kohn-Sham DFT.⁵ Following on from this in Chapter 3 we discuss OF-DFT and introduce the concept of the orbital free approximation of the noninteracting kinetic energy functional (OF-KEF). In addition, explanations of various forms of OF-KEFs found in the literature are given.

We will then shift focus to discussing some established variational schemes — Lopez-Acevedo and CCH — in the literature and discuss how we modified CCH to converge the energy for a wider range of OF-KEFs than had been previously reported. The last section of Chapter 3 will discuss the theory behind the TRIM method we have developed in this work. We will see that the TRIM scheme relies on the fact that the optimisation problem at the heart of OF-DFT is a saddle point optimisation problem.

In Chapter 4 a detailed description of the CCH scheme is given and how we implemented this scheme in order to converge the energy for a wide range of OF-KEFs. It is shown that most functionals predict very poor energies and densities. Furthermore, most OF-KEFs do not predict molecular binding. In Chapter 5 we discuss the form of the potentials generated by the OF-KEFs. It is shown that the sum of the kinetic potential and the effective potential approaches a constant in an oscillating manner, pointwise in space, when using Gaussian basis sets. This means if one wants to compute forces there is a Pulay like term in the equations. In addition, the balances between the potentials in the Euler equation are examined. This provides an explanation for the small chemical potential values one computes using these OF-KEFs.

In Chapter 6 the TRIM scheme is presented and our implementation of the scheme is discussed in detail. A demonstration of why the TRIM scheme is more efficient than both the Lopez-Acevedo and CCH schemes is provided. Furthermore, it is highlighted that the importance of the guess density for molecular systems in ensuring convergence is optimal.

In Chapter 7 we examine a variation principle at the interface of wavefunction and density functional theories. A recently proposed variation principle⁶ is examined and is shown⁷ that it can be expressed in terms of the well known Lieb functional. The equivalence between the information obtained from the two approaches is illustrated numerically by their implementation in a common framework.

In Chapter 8 a summary of the work is presented and a view on the future of research in OF-DFT is given. These views are supplemented with results from preliminary investigations on these future research directions.

Acknowledgements

I would like to thank my supervisor, Dr Andy Teale. Since we first met when I took up a summer project with him he has shown nothing but support and kindness. He is a fantastic person and I'm glad that I had the opportunity to work with him.

I also wish to extend my thanks to all members of the Teale group over the years. In particular for this project I'd like to thank Mike Withnall who tested my codes to the max and to Tom Irons who has an impressive knowledge of Python and without whose help this work would not have been possible to the extent shown in this thesis.

Acronyms

- **CC** Coupled Cluster.
- **CCH** Chan, Cohen and Handy optimisation scheme.
- **CCSD** Coupled Cluster theory with Single and Doubly excited determinants.
- **DFT** Density Functional Theory.
- **DIIS** Direct Inversion of Iterative Subspace.
- **ETOF** Even Tempered Orbital Free basis set.
- **ETOF-MOL** Even Tempered Orbital Free basis set for MOLecular systems.
- **GCE** Grand Canonical Ensemble.
- **GGA** Generalised Gradient Approximation.
- **GTO** Gaussian type Orbitals.
- **HF** Hartree-Fock.
- **HK** Hohenberg-Kohn.
- **KE** non-interacting Kinetic Energy.
- **KEF** non-interacting Kinetic Energy Functional.
- **KS-DFT** Kohn-Sham Density Functional Theory.
- **LDA** Local Density Approximation.
- **MO** Molecular Orbitals.
- **OF-DFT** Orbital Free Density Functional Theory.

OF-KEF orbital free non-interacting Kinetic Energy Functional.

- **PKSA** Post Kohn Sham Analysis.
- SC Self Consistent.
- **SCF** Self Consistent Field.
- **SCOFA** Self Consistent Orbital Free density functional theory Analysis.
- **SGA** Second order Gradient Expansion of the non-interacting kinetic energy density.
- **STO** Slater type Orbitals.
- **TF** Thomas-Fermi non-interacting kinetic energy functional.
- **TFD** λ **W** Thomas Fermi + Dirac exchange + λ -von-Weizsäcker functional.
- **TFDW** Thomas Fermi + von Weizsäcker with Dirac exchange.
- **TFVW** Thomas Fermi + Von Weizsäcker.
- **TRIM** Trust Region Image Method.
- **VW** von Weizsäcker non-interacting kinetic energy functional.
- **XC** eXchange Correlation.

Contents

1.	Mat	hemati	cal Background	1
	1.1.	An Int	roduction to Topological and Metric spaces	1
		1.1.1.	Dual Spaces	3
		1.1.2.	Convergence and Continuity in Banach Spaces	4
		1.1.3.	Gâteaux Derivatives	6
	1.2.	Lebesg	gue Spaces	7
		1.2.1.	Lebesgue vs Riemann Integration	7
		1.2.2.	The Lebesgue Measure	8
		1.2.3.	Lebesgue Space	10
		1.2.4.	The Dual Spaces of L_p	15
		1.2.5.	The Sums and Intersections of L_p	18
		1.2.6.	Functional Derivatives in L_p spaces $\ldots \ldots \ldots$	20
	1.3.	Conver	x Analysis	21
		1.3.1.	Convex Sets	21
		1.3.2.	Convex Functions	22
		1.3.3.	Lower Semi-Continuous Functions	23
		1.3.4.	Sub-Derivatives	25
	1.4.	Introd	ucing $\Gamma(V)$	26
		1.4.1.	Γ-regularisation	27
		1.4.2.	Legendre-Fenchel Transformations	27
	1.5.	Saddle	Functions	29
		1.5.1.	Sub-Derivatives on Saddle Functions	30
		1.5.2.	Constrained Minimisation	30
	1.6.	Summ	ary	31
2.	Den	sity Fu	nctional Theory	33
	2.1.	Quant	um Mechanics	33
	2.2.	Electro	onic Structure Theory	37
		2.2.1.	Hartree-Fock Theory	38
		2.2.2.	Coupled-Cluster Theory	43
		2.2.3.	Rayleigh-Ritz Variation Principle	44

	2.3.	Formal Density Functional Theory	45
		2.3.1. On the Development of DFT	45
		2.3.2. Constrained Search Formalisms of DFT	46
		2.3.3. Topology of Density and Potential Sets	48
		2.3.4. Lieb Ensemble DFT	49
		2.3.5. Lieb's Convex-Conjugate Theory	51
	2.4.	Kohn-Sham Density Functional Theory	53
		2.4.1. Solving the KS-DFT Equations	54
		2.4.2. Exchange-Correlation Functionals	56
	2.5.	Summary	58
3.	Orbi	ital Free Density Functional Theory	59
	3.1.	Orbital Free Density Functional Theory	59
	3.2.	Thomas Fermi and von Weizsäcker Models	61
	3.3.	The Non-interacting Kinetic Energy Functional	66
		3.3.1. A Review on Functional Development	67
		3.3.2. A Review on Solving the Optimisation Problem	74
	3.4.	Solving the Optimisation Problem: A Formal Approach	79
		3.4.1. Grand Canonical Ensembles in OF-DFT	79
		3.4.2. The Importance of Saddle Functions in OF-DFT $~$	81
4.	The	3.4.2. The Importance of Saddle Functions in OF-DFT Importance of Self Consistency in Orbital Free Density	81
4.	The Fund	3.4.2. The Importance of Saddle Functions in OF-DFT Importance of Self Consistency in Orbital Free Density ctional Theory	81 85
4.	The Fune 4.1.	3.4.2. The Importance of Saddle Functions in OF-DFT Importance of Self Consistency in Orbital Free Density ctional Theory	81 85 85
4.	The Fun 4.1.	3.4.2. The Importance of Saddle Functions in OF-DFT Importance of Self Consistency in Orbital Free Density ctional Theory Introduction	81 85 85 85
4.	The Fune 4.1. 4.2.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme .	 81 85 85 85 86
4.	The Fund 4.1. 4.2.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme .	 81 85 85 85 86 86
4.	The Fund 4.1. 4.2.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions .	 81 85 85 85 86 86 88
4.	The Fund 4.1. 4.2.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme .	 81 85 85 86 86 88 92
4.	The Fune 4.1. 4.2.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . The Gradient and Hessian .	 81 85 85 86 86 88 92 94
4.	The Fund 4.1. 4.2. 4.3.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm .	 81 85 85 86 86 88 92 94 94
4.	The Fund 4.1. 4.2. 4.3.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient .	 81 85 85 86 86 88 92 94 94 97
4.	The Fund 4.1. 4.2. 4.3.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian .	 81 85 85 86 86 88 92 94 94 97 98
4.	The Fund 4.1. 4.2. 4.3.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . Results .	 81 85 85 86 86 88 92 94 94 97 98 99
4.	The Fund 4.1. 4.2. 4.3. 4.4.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . 4.3.4.1. Atomic Systems .	 81 85 85 86 86 88 92 94 97 98 99 100
4.	The Fund 4.1. 4.2. 4.3. 4.4.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . 4.4.1. Atomic Systems . 4.4.2. Molecular Systems .	81 85 85 86 86 88 92 94 94 97 98 99 100 106
4.	The Fund 4.1. 4.2. 4.3. 4.4. 4.5.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . The Gradient and Hessian . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . 4.4.1. Atomic Systems . 4.4.2. Molecular Systems .	 81 85 85 86 86 88 92 94 97 98 99 100 106 110
4.	The Fund 4.1. 4.2. 4.3. 4.4. 4.5. Pote	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . 4.4.1. Atomic Systems . 4.4.2. Molecular Systems . 4.4.2. Molecular Systems .	 81 85 85 86 86 88 92 94 94 97 98 99 100 106 110 113
4.	The Fund 4.1. 4.2. 4.3. 4.4. 4.5. Pote 5.1.	3.4.2. The Importance of Saddle Functions in OF-DFT . Importance of Self Consistency in Orbital Free Density ctional Theory Introduction . 4.1.1. Recap of Variation Principles . The Chan Cohen Handy Scheme . 4.2.1. Outline of the Scheme . 4.2.2. The Choice of Basis Functions . 4.2.3. Convergence Properties of the CCH Scheme . 4.3.1. The Line Search Algorithm . 4.3.2. The Gradient . 4.3.3. Hessian . 4.4.1. Atomic Systems . 4.4.2. Molecular Systems . conclusions . conclusions .	81 85 85 86 86 88 92 94 97 98 99 100 106 110 113 113

	5.2.	The Basis Set Dependence of Potential Terms	116	
		5.2.1. The Implication for Computing Molecular Forces	120	
	5.3.	The Effect of Potentials on the Density	122	
	5.4.	Conclusions	126	
6.	. TRIM Scheme			
	6.1.	Introduction	127	
	6.2.	Theory	128	
		6.2.1. OF-DFT Variational Recap	128	
		6.2.2. A Second Order Optimization Scheme	129	
	6.3.	Results	137	
		6.3.1. Comparing Lopez-Acevedo, CCH and TRIM	137	
		6.3.2. Comparing CCH and TRIM	140	
		6.3.3. Molecular Calculations and the Initial Guess	142	
	6.4.	Conclusions	146	
_	-		-	
7.	Con	nections between variation principles	147	
	7.1.		147	
	7.2.	Variation Principles	148	
		7.2.1. Rayleigh–Ritz Variation Principle	148	
		7.2.2. Hohenberg–Kohn and Lieb Variation Principles	150	
	7.3.	Gidopoulos-Davidson Variation Principles	150	
		7.3.1. Relationship to Lieb Variation Principle	151	
		7.3.2. Objective Functions	152	
		7.3.3. Functional Derivatives of the Objective Functions	152	
		7.3.4. Generalisation to Ensembles	154	
		7.3.5. Generalization to Arbitrary Interaction Strengths	154	
7.4.		Adiabatic Connection	155	
	7.5.	Results	156	
		7.5.1. Computational Details	157	
		7.5.2. Kohn–Sham Non-Interacting System	158	
		7.5.3. General Interaction Strengths	158	
		7.5.4. Constructing the Adiabatic Connection	160	
	7.6.	Conclusions	161	
8.	Futi	ire Work and Summary	163	
	8.1.	Future Areas of Research	163	
		8.1.1. Laplacian Level Functionals	163	
		8.1.2. Non-Local Functionals	164	
		8.1.3. Machine Learning in OF-DFT	168	

	8.2.	Conclusions	169
Α.	Equa	ations for the gradient and Hessians used in this work	185
	A.1.	Preliminaries	185
	A.2.	Gradient	185
	A.3.	Hessian	186
		A.3.1. LDA	188
		A.3.2. GGA	188
		A.3.3. Laplacian Level	192
в.	Equa	ations for the spin-polarised potential used in this work	197
	В.1.	Derivation	197

1. Mathematical Background

1.1. An Introduction to Topological and Metric spaces

In this thesis we will be utilising ideas derived from the formulation of density functional theory (DFT) using convex analysis by Lieb.⁸ In this chapter we will introduce and explain in some detail the mathematics required to understand this formulation. To begin with we will introduce the concept of the topological space. As we will explain in Section 2.3.3 the sets of ground state densities and potentials are examples of topological spaces^{9,10} called Lebesgue spaces which are also introduced in this chapter.

We will also discuss some tools from convex analysis required for understanding Lieb's convex formulation of DFT discussed in Chapter 2. We will finish this chapter by introducing saddle functions which are central to the idea introduced in Chapter 3 and applied in Chapter 6.

We begin by defining the topological space which underlines the material in this chapter.

Definition 1.1.1. A topological space $T = (X, \mathcal{T})$ consists of a nonempty set X with a fixed family \mathcal{T} of subsets of X satisfying

- 1. $X, \emptyset \in \mathcal{T}$
- 2. The finite intersection of any members in \mathcal{T} is also in \mathcal{T}
- 3. The countable^{*} union of any members in \mathcal{T} is also in \mathcal{T}

Where \emptyset denotes the empty set. The family \mathcal{T} is called a topology for X and its elements are the open sets on X. Given a set $C \in X$ then it is a closed set in X if its complement $C^c = X \setminus C$ is an open set. Or in other words, if C is in X but not in \mathcal{T} then C is closed. In a topological space it is often convenient to have some subfamily \mathcal{B} of open sets such that any open set in X can be constructed from a union of sets in \mathcal{B} . This leads to the concept of the topological basis.

^{*} Either finite or countably infinite.

Definition 1.1.2. Given a topological space (X, \mathcal{T}) a basis for \mathcal{T} is a subfamily $\mathcal{B} \subset \mathcal{T}$ such that every set in \mathcal{T} is a union of sets from \mathcal{B} .

The topological basis is the smallest set of open sets which can generate the topology for X. We now need to introduce the concept of the *metric*

Definition 1.1.3. If X is a set, a metric on X is a function $d: X \times X \to \mathcal{R}$ such that $\forall x, y, z \in X$:

- 1. $d(x,y) \ge 0$;
- 2. $d(x,y) = 0 \Leftrightarrow x = y;$
- 3. d(x, y) = d(y, x);
- 4. $d(x, y) + d(y, z) \ge d(x, z)$.

A metric space consists of a set X together with a metric, d, where X is given the metric topology induced by the metric. In other words, a metric space is a special type of topological space. For example, we can generate a topology using a basis consisting of open balls, $B_{\epsilon}(x)$

$$B_{\epsilon}(x) = \{ y \in X \mid d(x, y) < \epsilon \}, \qquad (1.1)$$

with $x \in X$ and $\epsilon > 0$. This metric topology is important for discussing concepts such as convergence of a sequence. Let (X, d) be a metric space. A sequence (x_n) in X is said to be a Cauchy sequence if, given $\epsilon > 0$, there exists an integer n_0 such that $d(x_p, x_q) \leq \epsilon \quad \forall \quad p, q \geq n_0$. Alternatively if a sequence is Cauchy it cannot escape from any open balls it enters. Obviously every convergent sequence is Cauchy. If every Cauchy sequence in X converges to an element in X then we say this space is complete.

Now let V be a linear space over a field \mathcal{F} . A field is a set on which addition, subtraction, multiplication, and division are defined for every element of the field. An example of a field would be the positive real numbers excluding zero. V being a linear space means that it is closed with respect to the following operations: vector addition

$$x + y \in V \quad \forall \quad (x, y) \in V \times V \tag{1.2}$$

and scalar multiplication

$$\alpha x \in V \ \forall \ (\alpha, x) \in \mathcal{F} \times V. \tag{1.3}$$

A **norm**, ||.|| on V is a mapping from V to the non-negative real numbers, \mathcal{R}_+ , which satisfies, over the field \mathcal{F} ,

- 1. $||x|| \ge 0 \quad \forall x \in X$
- 2. $||x+y|| \le ||x|| + ||y|| \quad \forall x, y \in X,$
- 3. $||\alpha x|| = |\alpha|||x|| \quad \forall \quad (\alpha, x) \in \mathcal{F} \times V,$
- 4. $||x|| = 0 \Leftrightarrow x = 0.$

A linear space equipped with a norm is called a normed linear space or normed space. The norm satisfies all the requirements for a metric and so the normed space is a special type of metric space. This means we can discuss the continuity and differentiability of functions acting on any normed space using the norm.

We can now define the mathematical object that underpins so much of DFT, the Banach space.

Definition 1.1.4. A normed space which is complete with respect to the associated metric is called a **Banach space**.

We will explain in Section 2.3.3 that the sets of potentials and densities used in Lieb's formulation of DFT are in fact examples of Banach spaces.

1.1.1. Dual Spaces

Having introduced the concept of the normed space we now consider general linear transformations between two such spaces which will give rise to the concept of dual spaces. Dual spaces are crucial to the mathematical analysis of DFT in which the set of ground state potentials is the dual space of the set of densities. We will discuss this duality in the next chapter.

Let X and Y be linear spaces over the same field \mathcal{F} . A linear mapping $T: X \to Y$ is one that preserves the linear operations i.e.

$$T(x + x') = T(x) + T(x')$$
(1.4)

$$T(\alpha x) = \alpha T(x). \tag{1.5}$$

for $x, x' \in X$ and $\alpha \in \mathcal{F}$.

Theorem 1.1.1. Let X and Y be normed spaces over the same field \mathcal{F} and let T be a linear mapping from X to Y. Then the following statements are equivalent

- T is continuous
- $\exists M \in \mathcal{R} \text{ such that } ||T(x)|| \le M ||x|| \quad \forall x \in X$

Linear mappings into the field \mathcal{F} underlying the linear field are called linear functionals. For example, the ground state potential, v is a linear functional acting on the set of densities. If the ground state potential is linear functional of the density — which is the case in DFT where we use the Coulomb potential — we are then guaranteed that $\int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$ is finite. As we will see in the next chapter this integral is required in the evaluation of the total energy.

We now denote the set of all linear mappings from X to Y by L(X, Y)and the set of all continuous linear mappings from X to Y as $L^{C}(X, Y)$. The set $L(X, \mathcal{F})$ is known as the algebraic dual. We are more interested in the set $L^{C}(X, \mathcal{F})$ which is known as the topological dual of X. As we are only concerned with the latter we will just refer to this as the **dual space** of X which we denote by X^* . One can show that X^* is a Banach space even if X is not. A simple way of describing the dual space of X using Theorem 1.1.1 — is that this space is the set of all bounded linear functionals whose domain is a subset of X.

We can continue by defining the bidual

$$X^{**} = L^C(X^*, \mathcal{F}).$$
(1.6)

which is the set of all bounded linear functionals whose domain is a subset of the dual space of X. X is said to be **reflexive** if $X = X^{**}$. If $X \subset X^{**}$ then X is referred to as **non-reflexive**. The Banach spaces underlining DFT are non-reflexive. This is in contrast to quantum mechanics where the spaces of wavefunctions are reflexive. The reflexive Banach spaces used in quantum mechanics are called Hilbert spaces and will be introduced in Chapter 2.

1.1.2. Convergence and Continuity in Banach Spaces

Another topic which we must discuss are the various types of convergence one can have in a Banach space. This will allow us to define the various concepts of continuity of functions in Banach spaces. The concepts in this section are required in defining the set Γ — introduced in Section 1.4 which contains the Lieb functional.

We will begin by considering sequences in a Banach space X and its dual X^* . There exists a duality between X and X^* given by the bi-linear

functional

$$(x^*|x) = x^*(x); \quad x \in X, x^* \in X^*.$$
(1.7)

This duality sets up a bounded linear functional $x \to (x^*|x)$ on X and also a bounded linear functional $x^* \to (x|x^*)$ on X^* . Only in reflexive spaces is $(x^*|x) = (x|x^*)$.

For example, in quantum mechanics we often write $\langle \Psi | \Psi \rangle$ without concern whether we are acting on the ket or bra space because of the reflexive nature of the Hilbert space. However, in DFT the duality $(v|\rho) \neq (\rho|v)$ due to the non-reflexive nature of the Banach spaces. The form we use is dependent on when we are evaluating the energy functional on the density space — $(v|\rho)$ — or the potential space — $(\rho|v)$. In practice, however, we are not strict on this.

We can now define both strong and weak convergence. Let $\{x_n\}$ be a sequence in the normed space X.

Definition 1.1.5. Strong Convergence

$$||x_n - x|| \to 0 \quad as \quad n \to \infty$$

This is usually denoted by $x_n \to x$ and we say $\{x_n\}$ converges strongly to x.

Definition 1.1.6. Weak Convergence

$$\forall x^* \in X^*, ||(x^*|x - x_n)|| \to 0 \quad as \quad n \to \infty$$

This is usually denoted by $x_n \rightarrow x$ and we say $\{x_n\}$ converges weakly to x. It's important to note that, in a Banach space, strong convergence implies weak convergence. If the Banach space is reflexive the reverse is also true.

Now consider a $\{x_n^*\}$ in the dual space X^* . Strong convergence is defined in an obvious manner and weak convergence is defined using the bidual X^{**} . However, when discussing sequences in X^* we have an even weaker notion of convergence: weak star convergence.

Definition 1.1.7. Weak-star Convergence

$$\forall x \in X, ||(x^* - x_n^*|x)|| \to 0 \quad as \quad n \to \infty.$$

This is usually denoted by $x_n^* \stackrel{*}{\rightharpoonup} x^*$. Weak convergence implies weakstar convergence but the converse is true only when the Banach space is reflexive. We use these ideas of convergence in defining the continuous nature of functions acting on Banach spaces. In the Banach space X for a given $f: X \to \mathcal{R}^*$

$$x_n \to x \Rightarrow f(x_n) \to f(x)$$
 strong continuity (1.8)

$$x_n \rightarrow x \Rightarrow f(x_n) \rightarrow f(x)$$
 weak continuity. (1.9)

For a function acting on the dual space X^* we have three different forms of continuity

$$x_n^* \to x^* \Rightarrow f(x_n^*) \to f(x^*)$$
 strong continuity (1.10)

$$x_n^* \to x^* \Rightarrow f(x_n^*) \to f(x^*)$$
 weak continuity (1.11)

$$x_n^* \stackrel{*}{\rightharpoonup} x^* \Rightarrow f(x_n^*) \to f(x^*)$$
 weak-star continuity (1.12)

The relative strengths of continuity are given in the following theorem

Theorem 1.1.2. On the dual X^* of a normed space X, weak-star continuity of $g : X \to \mathcal{R}$ at x^* implies weak continuity at x^* , whereas weak continuity implies strong continuity. The converse statements do not hold. However, all $x^{**} \in X^{**}$ are both weakly and strongly continuous on X^* , whereas all $x \in X$ are weak-star, weakly, and strongly continuous on X^* .

1.1.3. Gâteaux Derivatives

This thesis is concerned with variational solutions in DFT. We will therefore need to be able to compute how functionals vary with changes in electron density. This requires knowledge of Gâteaux derivatives which is closely related to the more familiar directional derivative.

Let us consider a mapping $f: V \to \overline{\mathcal{R}}$ where $\overline{\mathcal{R}}$ is the extended real number system i.e $\mathcal{R} \cup \{\pm \infty\}$. If we change the argument x_0 in the direction of x by adding ϵx to x_0 with $\epsilon > 0$ the **directional derivative** in this direction is defined as

$$f'(x_0; x) = \lim_{\epsilon \to 0^+} \frac{f(x_0 + \epsilon x) - f(x_0)}{\epsilon}.$$
 (1.13)

If f' is linear in x then

$$\lim_{\epsilon \to 0^+} \frac{f(x_0 - \epsilon x) - f(x_0)}{-\epsilon} = \lim_{\epsilon \to 0^+} \frac{f(x_0 + \epsilon x) - f(x_0)}{\epsilon}$$
(1.14)

which shows the left and right hand derivatives are identical allowing us to

^{*} \mathcal{R} denotes the set of real numbers

write:

$$f'(x_0; x) = \lim_{\epsilon \to 0} \frac{f(x_0 + \epsilon x) - f(x_0)}{\epsilon}.$$
 (1.15)

So now let us consider a function $f : X \to \overline{\mathcal{R}}$. The function is said to be Gâteaux differentiable at x_0 if there exists a bounded linear functional $\nabla f(x_0) \in X^*$ known as the **Gâteaux derivative**

$$\nabla f(x_0)(x) = \lim_{\epsilon \to 0^+} \frac{f(x_0 + \epsilon x) - f(x_0)}{\epsilon}$$
(1.16)

It should be clear that if the Gâteaux derivative exists at x_0 then the directional derivative also exists. However, the existence of a directional derivative for all x at x_0 does not imply Gâteaux differentiability unless the directional derivatives are bounded and linear in x at x_0 .

1.2. Lebesgue Spaces

We will now introduce a special type of Banach space: *Lebesgue space*.^{11,12} As the properties of these types of spaces are used frequently in Lieb's convex formulation of DFT we will explore this subject in greater depth than we did for general Banach spaces. We will first discuss the concept of measure spaces and Lebesgue integration. Then we will introduce the Lebesgue space and discuss important properties such as derivatives, dual spaces and sum and intersections of these spaces.

1.2.1. Lebesgue vs Riemann Integration

The Riemann integral is the integration technique we are first introduced to. Whilst powerful it does require that the integrand is finite over the entire integration range. The Lebesgue theory of integration circumvents this problem by enlarging the collection of functions for which the integral is defined. This will become desirable in the formulation of DFT for example $\int_{\mathcal{R}^3} |\mathbf{r}|^{-1} d\mathbf{r}$ is not Riemann integrable but is Lebesgue integrable.

Recall that we can have the following interval types in \mathcal{R} :

$$[a,b] = \{x \in \mathcal{R} | a \le x \le b\} \text{ closed interval};$$
(1.17)

$$[a,b) = \{x \in \mathcal{R} | a \le x < b\} \text{ clopen interval};$$
(1.18)

$$(a,b) = \{x \in \mathcal{R} | a < x < b\} \text{ open interval.}$$
(1.19)

We define the length of the interval to be b - a.

Definition 1.2.1. Let E be a set. Then the characteristic function of E, χ_E , is defined by

$$\chi_E(x) = 1 \quad if \ x \in E$$
$$= 0 \quad if \ x \notin E.$$

We can now define a step function, ϕ , as a linear combination of characteristic functions, χ , of intervals, I_i ,

$$\phi = \sum_{j=1}^{n} c_j \chi_{I_j} \tag{1.20}$$

where $c_j \in \mathcal{R}$. If the end-points of the interval I_j are a_j and b_j then one can define an integral of ϕ

$$\int \phi d\mathbf{r} = \sum_{j=1}^{n} c_j (b_j - a_j). \tag{1.21}$$

If ϕ is a bounded function over the entire interval then the Riemann integral of ϕ is defined to be the limit of the integrals of step-functions which approximate ϕ . To lift the bounded requirement we generalize the notion of length to a suitable collection of subsets, X of \mathcal{R} . To do this we need to pause and consider measures.

1.2.2. The Lebesgue Measure

Let us begin by defining an important concept: the σ -algebra

Definition 1.2.2. A family, \mathcal{X} , of subsets of a set X is said to be a σ -algebra if:

- 1. $\emptyset, X \in \mathcal{X},$
- 2. If $A \in \mathcal{X}$ then A^c (the complement) belongs to \mathcal{X} ,
- 3. If (A_n) is a sequence of sets in \mathcal{X} , then the finite or countable union of these sets belongs to \mathcal{X} .

We call (X, \mathcal{X}) a measurable space. Note that although similar to topologies, σ -algebras are different. However, they can be reunited through the construction of the Borel algebra.

Definition 1.2.3. Let X be the set \mathcal{R} . The **Borel algebra** is the σ -algebra, \mathcal{B} , generated by all open intervals in \mathcal{R} . The elements of \mathcal{B} are called **Borel sets**.

The Borel sets are very important in many areas of mathematics — for example, unifying topologies and σ -algebras — but for us it is important as it helps define the Lebesgue measure. So now we have defined a measurable space what exactly is a measure?

Definition 1.2.4. Let \mathcal{X} be a σ -algebra generated by X. Let μ be an extended real valued function (its range is $\mathcal{R} \cup \{+\infty\}$), defined on \mathcal{X} . We call μ a measure if

- 1. $\mu(\emptyset) = 0$,
- 2. $\mu(E) \ge 0 \quad \forall \quad E \in \mathcal{X},$
- 3. μ is countably additive in the sense that if (E_n) is any disjoint sequence of sets in \mathcal{X} then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

We call (X, \mathcal{X}, μ) a measure space which really is a measurable space (X, \mathcal{X}) endowed with the specific measure μ . We are now in the position to define the Lebesgue measure.

Definition 1.2.5. Let $X = \mathcal{R}$ and let its σ -algebra be the Borel algebra \mathcal{B} then it can be shown that there exists a unique measure μ defined on \mathcal{B} which equals the length of open intervals. We then define μ to be the **Lebesgue measure**.

For example, if E is a singleton — a set with only one element — then $\mu(E) = 0$. We can now define the **Lebesgue integral** of a function f over a measure space $\mathcal{M} = (X, \mathcal{X}, \mu)$ as

$$\int_{\mathcal{M}} f d\mu := \lim_{n \to \infty} S_n = \lim_{n \to \infty} \sum_{i=1}^{n} \eta_i \mu(E_i)$$
(1.22)

where $\mu(E_i)$ is the Lebesgue measure of the set E_i of points on the x-axis for which f(x) approximates η_i . The limit is known as a refinement of the domain of f i.e. we split the domain into smaller and smaller intervals. So imagine we have f that is bounded except at a finite, isolated, number of points. We say that f is bounded **almost everywhere** (a.e.). What this means is that whenever $n_i = +\infty$, $\mu(E_i) = 0$ and so its contribution to the integral is zero. Whereas for the same function using Riemann integration will yield a divergent integral. To briefly summarise, without proof, **Theorem 1.2.1.** Let \mathcal{M} be a measure space endowed with the Lebesgue measure. Let $dom(f) \subset \mathcal{M}$. Then if f is bounded everywhere the Riemann integration and Lebesgue integration are equal. If f is bounded a.e. then the Riemann integral diverges but the Lebesgue remains finite.

It is clear that we would rather use the Lebesgue integral in our analysis of DFT especially as some integrands in DFT diverge at isolated points. For example, the integrand of the Coulomb energy diverges at the origin of the integration range. We are now ready to introduce the Lebesgue space.

1.2.3. Lebesgue Space

Throughout this section μ will denote the Lebesgue measure. In this section we will impose a Banach space structure on the set of all (Lebesgue) integrable functions on a measure space (X, \mathcal{X}, μ) and introduce the Lebesgue spaces, L_p , where $1 \leq p \leq \infty$. We will then discuss the properties of L_p spaces that are essential to understanding Lieb's convex formulation of DFT. Due to the importance of these spaces, proofs of some of the key properties are included.

Definition 1.2.6. The space $L(X, \mathcal{X}, \mu)$ is a linear space of Lebesgue functions whose domain is the measure space $\mathcal{M}(X, \mathcal{X}, \mu)$.

To take a step closer to imposing the Banach structure on L we need to make L a normed space. We define two functions to be equal a.e. if the Lebesgue integral of these two functions are equal. The class of functions that are equal a.e. are called μ -equivalence classes. So we are really interested in the μ -equivalence classes generated by the members of L. We will denote them by [f].

Definition 1.2.7. The Lebesgue space L_1 is a set of μ -equivalence classes generated by $L(X, \mathcal{X}, \mu)$. If $[f] \in L_1$ then it has a norm defined through

$$||[f]||_1 = \int |[f]| d\mu.$$

Theorem 1.2.2. The Lebesgue space $L_1(X, \mathcal{X}, \mu)$ is a normed space.

Proof. Let $[f], [g] \in L_1(X, \mathcal{X}, \mu)$. Firstly it is obvious that $||[f]||_1 = 0$ if and only if [f] = 0. It is also obvious that $||[f]||_1 \ge 0 \forall L_1$. Now let $\alpha \in \mathcal{R}$. Then

$$||[\alpha[f]]||_{1} = \int |\alpha[f]| d\mu = |\alpha| \int |[f]| d\mu = |\alpha|||[f]||_{1}.$$

And finally

$$|[f+g]]_1 = \int |[f+g]| d\mu \le \int |[f]| + |[g|d\mu = ||[f]||_1 + ||[g]||_1$$

For notational ease we will drop the equivalence class notation but it should be remembered that all results in this section are for equivalence classes of functions not just for single functions.

We will now consider the spaces L_p : $1 \le p < \infty$.

Definition 1.2.8. If $1 \le p < \infty$ the space L_p consists of all μ -equivalence classes of functions f for which $|f|^p$ has a finite Lebesgue integral. We then set $\left(\int_{a}^{b} \int_{a}^{\frac{1}{p}} \int_{a}^$

$$||f||_p = \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$$

as the norm of L_p .

It can be shown that L_p is a normed space, we will prove that it is also complete and so is a Banach space. However, we need some intermediate theorems the first of which is key to proving a lot of the properties of Lebesgue spaces and so has an accompanying proof for the interested reader.

Theorem 1.2.3. Holders Inequality

Let $f \in L_p$ and $g \in L_q$ where p > 1 and pq = p + q. Then $fg \in L_1$ and $||fg||_1 \leq ||f||_p ||g||_q$

Proof. Let $\alpha \in (0, 1)$ and $t \ge 0$. Consider

$$\phi(t) = \alpha t - t^{\alpha}.$$

Because $\phi'(t) < 0 \ \forall t \in (0,1)$ and $\phi'(t) > 0, \ \forall t \in (1,+\infty)$ then $\phi(t) \ge \phi(1)$ with equality only if t = 1. Therefore

$$t^{\alpha} \le \alpha t + (1 - \alpha), \ t \ge 0.$$

Assume $ab \ge 0$ and let t = a/b, we get

$$a^{\alpha}b^{1-\alpha} \le \alpha a + (1-\alpha)b.$$

Equality is reached if and only if a = b. Now let $p \in (1, \infty)$ and $q^{-1} = 1 - p^{-1}$. Set $\alpha = p^{-1}$. It follows that

$$AB \le \frac{A^p}{p} + \frac{B^q}{q}$$

if A and B are non-negative real numbers. Suppose now that $f \in L_p$ and $g \in L_q$ and $||f||_p ||g||_q \neq 0$. If we allow $A = |f(x)|/||f||_p$ and $B = |g(x)|/||g||_q$ then we have

$$\frac{|f(x)g(x)|}{||f||_p||g||_q} \le \frac{|f(x)|^p}{p||f||_p^p} + \frac{|g(x)|^q}{q||g||_q^q}.$$

As the RHS is integrable so is fg and so

$$\frac{||fg||_1}{||f||_p||g||_q} \le p^{-1} + q^{-1} = 1.$$

 $\{p,q\}$ are known as **Holder conjugates**. The next theorem can be proved using the Holder Inequality

Theorem 1.2.4. Minkowski's Inequality

If $f, h \in L_p$, if $p \ge 1$, then $f + h \in L_p$ and

$$||f + h||_p \le ||f||_p + ||h||_p.$$

The following two theorems arise from the study of integration.

Theorem 1.2.5. Fatou's Lemma

If (f_n) is a sequence of functions bounded from below then

$$\int (\liminf_{n \to \infty} f_n) d\mu \ge \liminf_{n \to \infty} \int f_n d\mu.$$

Theorem 1.2.6. Dominated Convergence Theorem

Let (f_n) be a sequence of integrable functions which converge a.e. to a realvalue bounded function f. If there exists an integrable function g such that $|f_n| \leq g \forall n$ then f is integrable and

$$\int f d\mu = \lim_{n \to \infty} \int f_n d\mu.$$

We are now in a position to prove that L_p is a complete and so is a Banach space for $1 \le p < \infty$.

Theorem 1.2.7. Riesz-Fischer Theorem

If $p \in [1,\infty)$ then L_p is a Banach space under the norm

$$||f||_{p} = \left[\int |f|^{p} d\mu\right]^{\frac{1}{p}}.$$
(1.23)

Proof. Let (f_n) be a Cauchy sequence with respect to the norm $||f||_p$. Hence if ϵ there exists a N_{ϵ} such that if $n, m \geq N_{\epsilon}$ then

$$||f_m - f_n||_p^p < \epsilon^p.$$

There exists a subsequence (g_k) of (f_n) such that $||g_{k+1} - g_k|| < 2^{-k}$ for some integer k. Now we define

$$g(x) = |g_1(x)| + \sum_{k=1}^{\infty} |g_{k+1}(x) - g_k(x)|$$

so we can apply Fatou's Lemma in the following manner

$$\int |g|^p \mathrm{d}\mu \le \liminf_{n \to \infty} \int \left[|g_1| + \sum_{k=1}^n |g_{k+1} - g_k| \right]^p \mathrm{d}\mu.$$

We can then write, using the Minowski Inequality,

$$\left[\int |g|^{p} \mathrm{d}\mu\right]^{\frac{1}{p}} \leq \liminf_{n \to \infty} \left(||g_{1}||_{p}^{p} + \sum_{k=1}^{n} ||g_{k+1} - g_{k}||_{p} \right)$$
$$\leq ||g_{1}||_{p} + 1$$

Hence if $E = \{x \in X : y(x) < +\infty\}$ then $E \in X$ and $\mu(X/E) = 0$. So (g_k) converges a.e. Now define

$$f(x) = g_1(x) + \sum_{k=1}^{\infty} (g_{k+1}(x) - g_k(x)), x \in E$$

= 0, x \notice E

Since $||g_k|| \leq g$ and (g_k) converges a.e. to f then by the dominated convergence theorem (DCT), $f \in L_p$. Since $|f - g_k|^p \leq 2^p g^p$ we infer using the DCT that $0 = \lim ||f - g_k||_p$ such that (g_k) converges in L_p to f. Then

$$\int |f_m - g_k|^p \mathrm{d}\mu < \epsilon^p.$$

Applying Fatou's lemma again we conclude

$$\int |f_m - f|^p \mathrm{d}\mu \le \liminf \int |f_m - g_k|^p \mathrm{d}\mu \le \epsilon^p \tag{1.24}$$

 \square

Hence (f_n) converges to f in the norm of L^p .

We have now discussed the Banach space of L_p where $p < +\infty$ but we also need to discuss L_{∞} as this Banach space is important for DFT as we will show that the set of potentials is constructed using this space.

Definition 1.2.9. The space $L_{\infty} = L_{\infty}(X, \mathcal{X}, \mu)$ consists of all the equivalence classes of \mathcal{X} -measurable real-valued functions which are bounded a.e. If $f \in L_{\infty}$ and $N \in \mathcal{X}$ with $\mu(N) = 0$, we define

$$S(N) = \sup_{x} \left\{ |f(x)| : x \notin N \right\}$$

and

$$||f||_{\infty} = \inf \left\{ S(N) | N \in \mathcal{X}, \mu(N) = 0 \right\}$$

If $f \in L_{\infty}$ then we say f is an essentially bounded function

What we mean when we say that f is a \mathcal{X} -measurable function is that $\{x \in X | f(x) < \alpha\} \in \mathcal{X} \quad \forall \ \alpha \in \mathcal{R}$. The L_{∞} space is a normed space which is easily proved. To complete our analysis of the types of Lebesgue space we will prove the following theorem.

Theorem 1.2.8. The space L_{∞} is a Banach space with respect to the norm, $||.||_{\infty}$

Proof. Let (f_n) be a Cauchy sequence in L_{∞} . Let $M \subset \mathcal{X}$ with $\mu(M) = 0$, such that $|f_n(x)| \leq ||f_n||_{\infty}$ for $x \notin M, n = 1, 2, \cdots$ and also such that $|f_n(x) - f_m(x)| \leq ||f_n - f_m||_{\infty} \quad \forall x \notin M, n, m = 1, 2, \cdots$. Then the sequence (f_n) uniformly convergent on $X \setminus M$ and we let

$$f(x) = \lim f_n(x), x \notin M,$$

= 0, x \in M.

It follows that f is measurable and that $||f_n - f_{\infty}|| \to 0$. Hence L_{∞} is complete.

Now that we have discussed the Lebesgue space we need to discuss how one computes the dual of such spaces and the sum and intersections of such spaces. Furthermore, we need to know how functional derivatives are computed in Lebesgue spaces. These concepts are important for understanding formal DFT where we will find that L_p spaces contain densities and potentials used in DFT. By utilising the fact these spaces are Banach we will be able to derive some important properties of these spaces.

Consider, for example, the function

$$f(\mathbf{r}) = \frac{1 - e^{-|\mathbf{r}|}}{|\mathbf{r}|}.$$
 (1.25)

This function is a monotonically decreasing function as $|r| \to \infty$ and hence is bounded and so $f \in L_{\infty}$. For another example consider the function

$$g(\mathbf{r}) = \frac{e^{-|\mathbf{r}|}}{|\mathbf{r}|}.$$
(1.26)

As $\int g(\mathbf{r})^{5/2} d\mathbf{r} \sim \int r^2 g(r) dr$ as one goes from 3-dimensional to 1-dimensional integration. Then $\int r^2 g(r) dr = \int r^{-1/2} e^{-5/2} dr < \infty$ and so $f \in L_{5/2}$. As the Coulomb potential energy can be expressed as the sum of g and f we can expect that the Lebesgue space of the Coulomb potential to be a sum of L_{∞} and $L_{5/2}$. We will demonstrate this rigorously below.

1.2.4. The Dual Spaces of L_p

In this section we will compute the dual space of a given Lebesgue space, L_p and we will let $\{p, q\}$ be Holder conjugates. Dual spaces of Lebesgue spaces are very important in the development of DFT. We will find that the dual space of densities are the potentials for example. Therefore, we must be able to compute the dual of a Lebesgue space. Recall that the dual space of any Banach space (e.g. Lebesgue spaces) consists of all bounded linear functionals on the space. If X is a real Banach space, the dual space of X^* consists of all bounded linear functionals $F: X \to \mathcal{R}$, with norm

$$||F||_{X^*} = \sup_{x \in X \setminus 0} \left(\frac{|F(x)|}{||x||_X} \right)$$
(1.27)

Here we have introduced a new notational device $||.||_S$ which signifies that the we are using the norm associated with the normed space S. Holders inequality implies that functions in L_q define bounded linear functionals on L_p with the same norm.

Theorem 1.2.9. Let (X, \mathcal{X}, μ) be a measure space and $1 \leq p \leq \infty$. If

 $f \in L_q(X)$ then

$$F(g) = \int fg d\mu$$

defines a bounded linear functional $F: L_p(X) \to \mathcal{R}$ and

$$||F||_{(L_p)^*} = ||f||_{L_q}$$

Furthermore, if X is σ -finite, then the same result holds for p = 1.

A point to note here is that the Lebesgue measure is σ -finite on the real numbers. What this theorem says is that the map F = J(f)

$$J: L_q(X) \to (L_p(X))^* \quad J(f): g \mapsto \int fg \mathrm{d}\mu, \qquad (1.28)$$

is a (isometric) map from L_q to the dual space of L_p . The main part of the following theorem, which we shall sketch a proof, is that for $1 every bounded linear functional on <math>L_p$ arises from an L_q function using J as a map.

Theorem 1.2.10. Let (X, \mathcal{X}, μ) be a measure space. Let p and q be Holder conjugates. If 1 then <math>J defines an isometric isomorphism * of $L_q(X)$ onto the dual space of $L_p(X)$

Proof. The objective here is to show that every $F \in (L_p(X))^*$ is given by J(f) for some $f \in L_q(X)$.

Suppose that $\mu(X) < \infty$ and let

$$F: L_p \to \mathcal{R}$$

be a bounded linear functional on L_p . If $B \in \mathcal{X}$, then $\chi_B \in L_p(X)$ (the characteristic function of B) and one may define $m : B \to \mathcal{R}$ by

$$m(B) = F(\chi_B)$$

If $B = \bigcup_{i=1}^{\infty} B_i$ is a disjoint union of measurable sets, then

$$\chi_B = \sum_{i=1}^{\infty} \chi_{B_i},$$

^{*} An isometric isomorphism is a bijective distance preserving map. To make this more explicit, let $f: X \to Y$ be a bijective map. Then if d(f(a), f(b)) = d(a, b) then f is a isometric isomorphism.

and using the DCT it's clear that

$$\lim_{n \to \infty} ||\chi_B - \sum_{i=1}^{\infty} \chi_{B_i}||_{L_p} \to 0$$

Hence, since F is a continuous linear functional on L_p , m is a signed measure on (X, \mathcal{X}) . One can show that there must exist a function $f: X \to \mathcal{R}$ such that

$$F(\chi_B) = \int f \chi_B \mathrm{d}\mu \ \forall \ B \in \mathcal{X}.$$

Hence for all functions ϕ ,

$$F(\phi) = \int f \phi \mathrm{d}\mu,$$

and

$$\left|\int f\phi \mathrm{d}\mu\right| \le M ||\phi||_{L_p}$$

where $M = ||F||_{(L_p)^*}$. Taking $\phi = \operatorname{sgn}(f)$ we see that $f \in L_1(X)$. We may then extend the integral of f against bounded functions by continuity. Explicitly, if $g \in L_\infty$, then there is a sequence of simple functions $\{\phi_n\}$ with $|\phi_n| \leq |g|$ such that $\phi_n \to g$, and, therefore, also in L_p . Since

$$|f\phi_n| \le ||g||_{L_{\infty}}|f| \in L_1(X),$$

the DCT and the continuity of F imply that

$$F(g) = \lim_{n \to \infty} F(\phi_n) = \int f g \mathrm{d}\mu$$

such that

$$\left|\int fg\mathrm{d}\mu\right| \le M||g||_{L_p} \ \forall \ g \in L_{\infty}(X)$$

We will now prove that $f \in L_q(X)$, where q is the Holder conjugate to p. Let $\{\phi_n\}$ be a sequence of functions such that

$$\phi_n \to f$$
 pointwise a.e. as $n \to \infty$

and $|\phi_n| \leq f$. Now we define

$$g_n = \operatorname{sgn}(f) \left(\frac{|\phi_n|}{||\phi_n||_{L_q}} \right)^{q/p}.$$

Then $g_n \in L_{\infty}(X)$ and $||g_n||_{L_p} = 1$. In addition, $fg_n = |fg_n|$ and

$$\int |\phi_n g_n| \mathrm{d}\mu = ||\phi_n||_{L_q}.$$

Now using these equalities, Fatous Lemma, $|\phi_n| \leq |f|$ and Section 1.2.4 one finds

$$||f||_{L_q} \leq \liminf_{n \to \infty} ||\phi_n||_{L_q}$$
$$\leq \liminf_{n \to \infty} \int |\phi_n g_n| d\mu$$
$$\leq \liminf_{n \to \infty} \int |fg_n| d\mu$$
$$\leq M$$

and so $f \in L_q$.

The extension to non-finite measure spaces is straightforward and unenlightening so we will not include it here. To summarise what we have just proved:

Theorem 1.2.11. If 1 and <math>p, q are holder conjugates, then the dual space of L_p is L_q .

We may inquire into the nature of the bi-duals of Lebesgue spaces. We say that a Banach space X is reflexive if and only if its bi-dual X^{**} is equal to X. So for 1 is reflexive.

So far we have not discussed the duals of the spaces L_1 and L_{∞} . Without proof we state the following theorem

Theorem 1.2.12. The dual space of L_1 is L_{∞} . The dual of L_{∞} contains L_1 . Both L_1 and L_{∞} are not reflexive.

1.2.5. The Sums and Intersections of L_p

Having discussed the dual spaces of L_p we now discuss another important property of L_p spaces, which is how do we combine two different Lebesgue spaces? This is important for DFT as the space of densities and potentials are combinations of Lebesgue spaces as we indicated previously.

Definition 1.2.10. The sums and intersections of Lebesgue spaces are de-

fined as

$$L_{(p,q)} = L_p \cap L_q, \quad p \ge q,$$
$$= L_p + L_q, \quad p < q,$$

respectively. It follows that $L_{(p,p)} := L_p$.

The intersection is constructed through

$$L_p \cap L_q = \{ x | x \in L_p, x \in L_q \}$$
(1.29)

and the sum through

$$L_p + L_q = \{x + y | x \in L_p, y \in L_q\}.$$
(1.30)

For both the sum and intersection a new norm is induced by the combination of spaces. By repeated use of the Holder inequality it is not too difficult to prove the following theorem

Theorem 1.2.13. Let p and q be Holder conjugates. Then

$$||f||_{L(p,q)} = \max(||f||_p, ||f||_q), \quad p \ge q,$$

=
$$\inf_{f_q + f_p = f}(||f_p||_p + ||f_q||_q), \quad p < q.$$

So now we know how to construct the sum and intersection of L_p spaces it is natural to inquire about the dual spaces of such constructions. We state without proof the following theorem

Theorem 1.2.14. If $1 \le p; q < 1$ with Holder conjugates $1 < p'; q' \le \infty$, then

$$(L_{(p,q)})^* = L_{(p',q')} \tag{1.31}$$

For example, the Banach space of densities which yield finite kinetic energies is given by $L_3 \cap L_1$. Its dual, the space of ground state potentials, is given by $L_{3/2} + L_{\infty}$. Based on the discussion we had at the end of the previous section we expected that the Coulomb potential to be in the space $L_{5/2} + L_{\infty}$. From the Hölder's inequality one can show $L_{5/2} \subset L_{3/2}$ which means that the Coulomb potential is indeed a linear functional acting on any $\rho \in L_3 \cap L_1$. The dual analysis also shows there can be more different functional forms of the ground state potential other than the traditional Coulomb form.

1.2.6. Functional Derivatives in L_p spaces

We have introduced Gâteaux derivatives on a general normed space in Section 1.1.3. It will be useful to us to refine this definition to a Lebesgue space L_p where $1 \leq p \leq \infty$. If the Gâteaux derivative of F exists at ρ_0 , it is by definition bounded and linear in its argument. Therefore, it can be represented by a function in the dual space, $(L_p)^*$ and so in DFT the ground state potentials can be related to the functional derivatives of the energy functional with respect to the ground state density.

Theorem 1.2.15. The functional derivative of F at ρ_0 is denoted by

$$\left. \frac{\delta F(\rho)}{\delta \rho(\mathbf{r})} \right|_{\rho = \rho_0}$$

and satisfies

$$\int \left. \frac{\delta F(\rho)}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_0} \rho(\mathbf{r}) d\mathbf{r} = \boldsymbol{\nabla} F(\rho_0)(\rho).$$
(1.32)

It will be useful for us to consider a functional of the form

$$F(\rho) = \int f(\rho(\mathbf{r}), \nabla \rho(\mathbf{r})) d\mathbf{r}.$$
 (1.33)

Using the definition of the Gâteaux derivative

$$\boldsymbol{\nabla}F(\rho_0)(\rho) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} \int f(\rho_0 + \epsilon\rho, \boldsymbol{\nabla}\rho_0 + \epsilon\boldsymbol{\nabla}\rho) \mathrm{d}\mathbf{r} \bigg|_{\epsilon=0}$$
(1.34)

$$= \int \left(\frac{\partial g}{\partial \rho} \Big|_{\rho = \rho_0} \rho + \frac{\partial g}{\partial \nabla \rho} \Big|_{\rho = \rho_0} \cdot \rho \right) d\mathbf{r}$$
(1.35)

$$= \int \left(\frac{\partial g}{\partial \rho} - \boldsymbol{\nabla} \cdot \frac{\partial g}{\partial \boldsymbol{\nabla} \rho} \right) \bigg|_{\rho = \rho_0} \rho \mathrm{d} \mathbf{r}$$
(1.36)

and, therefore,

$$\frac{\delta F(\rho)}{\delta \rho(\mathbf{r})} = \frac{\partial g}{\partial \rho} - \boldsymbol{\nabla} \cdot \frac{\partial g}{\partial \boldsymbol{\nabla} \rho}.$$
(1.37)

This concludes our discussion of Lebesgue spaces. We now move onto discuss an important topic for the development of DFT: *convex analysis*. Convex analysis was used by Lieb^{8,13,14} to put DFT on a solid mathematical footing. This will allow for us to define a new way of approaching the variational problem in orbital free DFT.


Figure 1.1.: The closed convex regular octagon, left panel, the open nonconvex octagon, middle panel, and the closed non-convex nonregular pentagon, right panel.

1.3. Convex Analysis

Convex analysis^{15,16} is important for studying formal DFT. In particular it will allow us to analyse the various universal functionals we will meet in Chapter 2.

1.3.1. Convex Sets

Let V be a linear space over \mathcal{R} . If $u, v \in V$ we can define a *line segment* denoted by [u, v] through

$$[u, v] := \{ \lambda u + (1 - \lambda)v | 0 \le \lambda \le 1 \}.$$
(1.38)

Definition 1.3.1. Convex Sets

A set $\mathcal{A} \subset V$ is said to be **convex** if and only if

$$[u,v] \in V \ \forall \ (u,v) \in \mathcal{A} \times \mathcal{A}$$

In Fig. 1.1 we have displayed: a convex closed set in \mathcal{R}^2 , left panel; a non-convex open set in \mathcal{R}^2 , middle panel; and a non-convex closed set in \mathcal{R}^2 .

If \mathcal{A} is any subset of V, the intersection of all sets containing \mathcal{A} is the smallest convex set containing \mathcal{A} . This set is called the **convex hull**, denoted by $co(\mathcal{A})$ and is formally defined through

$$\operatorname{co}(\mathcal{A}) = \left\{ \sum_{i=1}^{n} \lambda_{i} a_{i} | \sum_{i=1}^{n} \lambda_{i} = 1, \lambda_{i} \ge 0, a_{i} \in \mathcal{A}, 1 \le i \le n \right\}.$$
(1.39)

In other words the convex hull of \mathcal{A} is the set of all convex combinations of the elements of \mathcal{A} . In Fig. 1.2 we have drawn the convex hull of a highly



Figure 1.2.: The convex hull of a non-convex set in \mathcal{R}^3



Figure 1.3.: Plots of: the convex function x^2 , left panel; the concave function $-x^2$, middle panel; and the non-convex function $x^2 + \exp\left\{\frac{1}{100(x-1)^2}\right\} + 1$, right panel.

irregular polyhedron.*

1.3.2. Convex Functions

Let V be a normed space and consider mappings from $\mathcal{A} \subset V$ into $cl(\mathcal{R})$. Here the $cl(\mathcal{R})$ denotes the closure of \mathcal{R} . The closure of \mathcal{R} is just the union of the real numbers with $+\infty$.

Definition 1.3.2. Convex Functions

Let \mathcal{A} be a convex subspace of V and $F : \mathcal{A} \to cl(\mathcal{R})$. Then F is said to be a convex function on \mathcal{A} if and only if

$$F(\lambda u + (1 - \lambda)v) \le \lambda F(u) + (1 - \lambda)F(v), \quad \forall \ u, v \in \mathcal{A}, \ \forall \ \lambda \in [0, 1].$$

Additionally if -F is convex then F is **concave**. In Fig. 1.3 we have plotted the following maps $\mathcal{R}^2 \to \mathcal{R}$: convex function x^2 ; the concave function $-x^2$; and the non-convex function $x^2 + \exp\left\{\frac{1}{100(x-1)^2}\right\} + 1$.

Let $F: V \to \operatorname{cl}(\mathcal{R})$ be a convex function, then

$$dom(F) = \{ u | F(u) < +\infty \}$$
(1.40)

is called the *effective domain* and is a convex set. We include the value

^{*} Also known as a Triceratops



Figure 1.4.: The epigraph of the function x^2 in \mathcal{R}^2 .

 $+\infty$ so we can define the *extended function*, $\tilde{F} : \mathcal{A} \to \mathcal{R}$ by

$$\tilde{F}(u) = F(u) \text{ if } u \in \mathcal{A}$$

 $\tilde{F}(u) = +\infty \text{ if } u \notin \mathcal{A}$

So \tilde{F} is convex if and only if F is convex. Due to this extension, in convex analysis we need only consider functions defined everywhere. We say a convex function is **proper** if it is greater than $-\infty$ over its entire domain and there exits a least one point on the domain such that the function value is finite. If a convex function is not proper then we say it is improper.

Definition 1.3.3. The epigraph of a function $F: V \to cl(\mathcal{R})$ is the set

$$epi(F) = \{(u, a) \in V \times \mathcal{R} | f(u) \le a\}.$$

In Fig. 1.4 the epigraph of the x^2 map in \mathcal{R}^2 is shown. The epigraph is the set of points which lie above the graph of a function. One can prove straightforwardly that a function $F: V \to \mathcal{R}$ is convex if and only if its epigraph is convex.

1.3.3. Lower Semi-Continuous Functions

Lower semi-continuous functions appear in DFT. We will find that the universal functional in the convex formulation DFT is weak-star lower semicontinuous function.

Definition 1.3.4. Let V be a convex set. Then the function $F: V \to cl(\mathcal{R})$ is lower semi-continuous on V if

$$epi(F)$$
 is closed;
 $\forall v \in V, \lim_{u \to v} F(u) \ge F(v).$



Figure 1.5.: The graph of a lower semi-continuous function, left panel, and of an upper semi-continuous function, right panel. The red point means the function takes this value evaluated at 0.

In addition, if -F is lower semi-continuous then F is upper semi-continuous. Weak versions of both upper semi-continuous and lower semi continuous are defined as above but that $u \rightarrow v$. The weak-star versions are defined using the dual space and weak-star convergence. For example, weak-star lower semi-continuity is defined through

Definition 1.3.5. Let V be a convex set. Then the function $G : V^* \rightarrow cl(\mathcal{R})$ is weak-star lower semi-continuous on V^* if

$$\begin{array}{l} epi(G) \ is \ closed \\ \forall \ v^* \in V^*, \quad \liminf_{u_n^* \stackrel{*}{\rightharpoonup} v^*} G(u^*) \geq G(v^*). \end{array}$$

In Fig. 1.5 we have plotted examples of an lower semi-continuous and upper semi-continuous function on \mathcal{R}^2 . It should be clear from these plots that the epigraph of the lower semi-continuous function, left panel of Fig. 1.5, is closed whereas it is open for the the upper semi-continuous function, right panel for Fig. 1.5.

It can be shown that the point-wise supremum of lower semi-continuous functions is lower semi-continuous. Let G be a function that is not lower semi-continuous. Set \overline{G} to be the point-wise supremum of lower semi-continuous functions everywhere less than G. Then \overline{G} is the 'closet' lower semi-continuous approximation to G. This is called a lower semi-continuous regularisation.

Now consider some lower semi-continuous convex functions which assume a value $-\infty$. These are known as *improper* functions. Then the following theorem applies

Theorem 1.3.1. If $F : V \to cl(\mathcal{R})$ is a lower semi-continuous convex improper function, then it cannot take a finite value anywhere on its domain.

Therefore, for example, the universal functional in the rigorous formulation of DFT is a proper lower semi-continuous. This is discussed in the next chapter.

1.3.4. Sub-Derivatives

When formulating DFT in the next chapter we will be primarily concerned with convex (concave) functions. We also will want to find the global minima (maxima) of these functions which will require us to take derivatives. However, not all convex functions are differentiable, as easily seen when one considers |x|. However, the weak concept of sub-derivatives will allow us to identity the global minima of all convex functions.

Let $f: X \mapsto cl(\mathcal{R})$ where X is a linear space and let X^* denote the dual.

Definition 1.3.6. Let $x_0^* \in X^*$ and suppose

$$f(x) \ge f(x_0) + x_0^*(x - x_0) \quad \forall x \in X,$$

then x_0^* is the sub-gradient of f at x_0 . The sub-differential, $\partial f(x_0)$, is the set of all these sub-gradients.

The sub-differential is a closed convex set and can be linked to the Gâteaux derivative through the following theorem

Theorem 1.3.2. If the proper convex function $f : X \mapsto \mathcal{R}$ is continuous at $x_0 \in X$ with a unique sub-gradient $x_0^* \in X^*$, then its Gâtueax derivative is given by

$$\nabla f(x_0) = x_0^*$$

The following theorem gives the converse

Theorem 1.3.3. If the convex function $f : X \mapsto \mathcal{R}$ has Gâteaux derivative $\nabla f(x_0)$ at x_0 , then it is sub-differentiable at x_0 with a unique sub-gradient:

$$\partial f(x_0) = \{ \nabla f(x_0) \}$$

For concave functions we can define the super-differential, $\partial f(x_0)$, through the following theorem

Theorem 1.3.4. For $f : X \mapsto \mathcal{R}$, the super- and sub-differentials are related as

$$\bar{\partial}f(x_0) = -\partial(-f)(x_0)$$

To show how one can compute the sub-differential let us compute the sub-differential of $|x| : \mathcal{R} \to \mathcal{R}$ at x = 0. The sub-gradients at this point,

 $\{x_0^*\}$, are given through

$$\begin{aligned} |x| &\ge |0| + x_0^*(x-0) \\ |x| &\ge x_0^* x \ \forall x \in \mathcal{R}. \end{aligned}$$
(1.41)

Therefore, $x_0^* \in [-1, 1]$ and so the sub-differential $\partial |0| = [-1, 1]$. We know that |x| is minimised at x = 0. The fact that $0 \in \partial |0|$ supports this can be explained through the theorem,

Theorem 1.3.5. Let $f: X \to (-\infty, +\infty]$ and let $x_0 \in dom(f)$. Then

 $x_0^* \in \partial f(x_0) \Leftrightarrow x \mapsto f(x) - x_0^* x$ has a global minimum at x_0 $0 \in \partial f(x_0) \Leftrightarrow f$ has global minimum at x_0 .

Proof. By Definition 1.3.5 we have, $\forall x \in X, x_0 * \in \partial f(x_0) \Leftrightarrow f(x) \ge f(x_0) + x^*(x - x_0) \Leftrightarrow f(x) - x_0^* x \ge f(x_0) - x_0^* x_0$. This shows that x_0 minimises $f(x) - x_0^* x$. Let $x_0^* = 0$. Then $f(x) \ge f(x_0)$ so x_0 is a global minima of f.

1.4. Introducing $\Gamma(V)$

In this section we will introduce Γ^{13} which is the set of all lower semicontinuous convex functions which is an important space in DFT because the universal functional is an element of this space. Therefore, some of the variational principles are defined over this set in Lieb's convex formulation of DFT. This is discussed in Section 2.3.5. Let V be a convex set. The affine continuous functions $f: V \mapsto \mathcal{R}$ are functions of the type $f(v) = T(v) + \alpha$ where T is a linear, continuous functional over V, $v \in V$ and $\alpha \in \mathcal{R}$.

Definition 1.4.1. The set of functions $F : V \to cl(\mathcal{R})$ which are pointwise supremeum of a set of continuous affine functions is denoted by $\Gamma(V)$. $\Gamma_0(V)$ denotes the set of continuous proper affine functions.

By definition all functions that are members of Γ are convex and lower semi-continuous Conversely

Theorem 1.4.1. The following properties are equivalent

- 1. $F \in \Gamma(V)$
- 2. F is convex and lower semi-continuous function from V into $cl(\mathcal{R})$ and if F takes the value $-\infty$ then F is identically equal to ∞ .

Additionally if $F \in \Gamma_0(V)$ then it is also closed. Equivalently one can define $\hat{\Gamma}$ and $\hat{\Gamma}_0$ as the set of upper semi-continuous and proper semi-continuous functions. As an example if $G \in \hat{\Gamma}_0$ we know that G is a closed concave function.

We are most interested in the Γ -regularisation of functions. The Γ regularisation of any function gives us the closest approximation to it which is convex lower semi-continuous function.

1.4.1. Γ -regularisation

Definition 1.4.2. Let F and G be two functions of V into $cl(\mathcal{R})$. The following are equivalent to each other:

- 1. G is the point-wise supremum of the continuous affine functions everywhere less than F
- 2. G is the largest minorant of F in $\Gamma(V)$. G is then called the Γ -regularization of F.

In general, we can construct the epigraph of the Γ -regularization as the closed convex hull of the epigraph of the function. To finish we will note the relationship between F, its Γ -regularization, G and its lower semicontinuous regularization \overline{F} .

Theorem 1.4.2. Let $F: V \to cl$ and G be its Γ -regularization. Then

- 1. $G \leq \overline{F} \leq F$
- 2. if F is convex and admits a continuous affine minorant then $\overline{F} = G$

1.4.2. Legendre-Fenchel Transformations

In this last subsection we will now combine our knowledge of dual spaces and convex analysis to study Legendre-Fenchel Transforms - which lie at the heart of DFT optimisation theory.

In this section let V be a linear space and V^* be its dual, and let (.|.) denote the bi-linear map between them. Furthermore, we supply topological structures to each space so that concepts such as continuity can be applied to the spaces.

Let F be a function of V into $cl(\mathcal{R})$. If $u^* \in V^*$ and $\alpha \in \mathcal{R}$, the continuous affine function $u \mapsto (u|u^*) - \alpha$ is everywhere less than F if and only if

$$\forall \ u \in V, \ \alpha \ge (u|u^*) - F(u),$$

or

$$\alpha \ge F^*(u^*).$$

if we define

$$F^*(u^*) = \sup_{u \in V} \{ (u|u^*) - F(u) \}.$$
(1.42)

Definition 1.4.3. If $F : V \to cl(\mathcal{R})$, then Eq. (1.42) defines a function $F^* : V^* \to cl(\mathcal{R})$ and is called the **Legendre-Fenchel (LF) transform** of F.

The LF transform is a weak-star continuous functional. It is very clear that actually we can refine our supremum to the dom(F). This means that F^* is just the point-wise supremum of the family of continuous affine functions (u|.) - F(u), for $u \in \text{dom}(F)$. This means that $F^* \in \Gamma(V^*)$ and F is a convex weak-star lower semi-continuous function. We have the following properties all of which follow from the definition of the LF transform

Theorem 1.4.3.

1. $F^*(0) = -\inf_{u \in V} F(u)$ 2. if $F \le G$, we have $F^* \ge G^*$

3.

$$(\inf_{i \in I} F_i)^* = \sup_{i \in I} F_i^*$$
$$(\sup_{i \in I} F_i)^* \le \inf_{i \in I} F_i^*$$

for every family $(F_i)_{i \in I}$ of functions over V.

- 4. $(\lambda F)^*(u^*) = \lambda F^*(\frac{u^*}{\lambda}) \quad \forall \quad \lambda > 0$
- 5. $(F+\alpha)^* = F^* \alpha \quad \forall \quad \alpha \in \mathcal{R}$
- 6. $\forall a \in V$, we denote by F_a the translated function $F_a(v) = F(v-a)$. Then

$$(F_a)^*(u^*) = F^*(u^*) + (a|u^*)$$

We can repeat the process and LF transform F^* , the polar function, to the F^{**} , the bi-polar function. The bi-polar function is a function on V into $cl(\mathcal{R})$ and defined through

$$F^{**}(u) = \sup_{u^* \in V^*} \{ (u|u^*) - F^*(u^*) \}.$$
 (1.43)

 F^{**} is weakly lower-semi continuous. It is clear that $F^{**} \in \Gamma(V)$, which means that F^{**} is the Γ -regularisation of F. $F \in \Gamma(V)$ if and only if $F^{**} = F$ which leads to the following definition

Definition 1.4.4. The LF transform establishes a bijection between $\Gamma(V)$ and $\Gamma(V^*)$. $F \in \Gamma(V)$ and $G \in \Gamma(V^*)$ are said to be in a duality if they correspond to the bijection

$$F = G^*$$
 and $G = F^*$

The constants $\pm \infty$ on V and $\pm \infty$ are in duality. Thus $F \in \Gamma(V)$ if and only if $F^* \in \Gamma(V^*)$. Furthermore, the LF transforms provide a duality between $\Gamma(V)$ and $\Gamma(V^*)$.

In DFT we use the Lieb functional, $F_{\rm L}$, which is introduced in the next chapter, as a universal functional and is an element of $\Gamma(\mathcal{X})$. \mathcal{X} is an important Banach space which will be introduced in Section 2.3.3. The ground state energy corresponding to $F_{\rm L}$, E_0 is found through a LF transformation on $F_{\rm L}$ and $E_0 \in \tilde{\Gamma}_0(\mathcal{X}^*)$ which is the set of all proper upper semi-continuous concave functions on \mathcal{X}^* .

1.5. Saddle Functions

We will see that the Hohenberg-Kohn variation principle becomes a minimax problem when one allows for variations in the density that allow for changes in the electron number.¹³ As we will show later this is essential for practical orbital free DFT. For a general minimax problem the optimisation is described through both

$$\sup_{x \to x \in B} \inf_{x \in B} K(x, y) \tag{1.44}$$

$$\sup_{x \in A} \inf_{y \in B} K(x, y)$$
(1.44)
$$\inf_{y \in B} \sup_{x \in A} K(x, y)$$
(1.45)

where $\sup_{x \in A} \inf_{y \in B} K(x, y) \leq \inf_{y \in B} \sup_{x \in A} K(x, y)$. If the inequality is saturated by $(a_0, b_0) \in A \times B$ then we say (a_0, b_0) is the **saddle point** of $K: A \times B \to \mathcal{R}$ and $K(a_0, b_0)$ is the saddle value. Now let $K: A \times B \to \mathcal{R}$ $cl(\mathcal{R})$ have a convex mapping $x \mapsto K(x,y)$ and have a concave mapping

 $y \mapsto K(x, y)$. This is known as a convex-concave function and -K is a concave-convex function. We refer to both by the general name of saddle functions.

1.5.1. Sub-Derivatives on Saddle Functions

It will be helpful to know how one takes the sub-derivatives of a convexconcave saddle function K and a concave-convex function L.

Definition 1.5.1. The subdifferential of a concave-convex saddle function $K: X \times Y \rightarrow \mathcal{R}$ is

$$\partial K(x,y) = \bar{\partial}_1 K(x,y) \times \partial_2 K(x,y) \subset X^* \times Y^*$$

where $\bar{\partial}_1 K(x,y) = \bar{\partial} K(\cdot,y) \Big|_x \subset X^*$ and $\partial_2 K(x,y) = \partial K(x,\cdot) \Big|_y \subset X^*$. The subdifferential of a convex-concave saddle function $L: X \times Y \to \mathcal{R}$ is

$$\partial L(x,y) = \partial_1 L(x,y) \times \overline{\partial}_2 L(x,y) \subset X^* \times Y^*$$

where $\partial_1 L(x,y) = \partial L(\cdot,y) \big|_x \subset X^*$ and $\bar{\partial}_2 L(x,y) = \bar{\partial} L(x,\cdot) \big|_y \subset Y^*.$

The following theorem links sub-derivatives to the saddle-points of saddle functions

Theorem 1.5.1. Let $K : X \times Y \to \mathcal{R}$ be a saddle function. Then

$$\begin{aligned} (x^*, y^*) &\in \partial K(x, y) \Longleftrightarrow (u, v) \mapsto K(u, v) - (x^*|u) - (y^*|v) \\ has \ a \ saddle \ point \ at \quad (x, y), \\ (0, 0) &\in \partial K(x, y) \Longleftrightarrow K \ has \ a \ saddle \ point \ at \ (x, y) \end{aligned}$$

1.5.2. Constrained Minimisation

For each constrained optimisation problem discussed in this section there is a corresponding unconstrained optimisation problem. In order to show this we have to introduce the *Lagrange function* $L: X \times \mathcal{R} \mapsto \mathcal{R}$ through

$$L(x,\lambda) = f(x) + \lambda h(x) \tag{1.46}$$

where $\lambda \in \mathcal{R}$ is known as the *Lagrange multiplier*. The Lagrange function $L(x; \lambda)$ is a convex–concave saddle function, inheriting from f convexity in x and being affine and, therefore, concave in λ . The following theorem establishes the usefulness of the Lagrange function and underlies much of the later material,

Theorem 1.5.2. Let $f : X \to (-\infty, +\infty]$ be a proper convex-concave function $h : X \to \mathcal{R}$ be an affine function. Then $x_0 \in X$ is a minimiser of the constrained problem

$$f(x_0) = \inf_{x \in X} \{ f(x) | h(x) = 0 \}$$

if and only if there exists $\lambda_0 \in \mathcal{R}$ such that (x_0, λ_0) is the saddle point of the Lagrange function $L(x, \lambda) = f(x) + \lambda h(x)$. The saddle value of the Lagrange function is equal to the constrained minimum $f(x_0)$

Proof. Let f and h take the properties as ascribed in the theorem. If $x_0 \in X$ is the minimiser of the constrained problem then

$$f(x_0) \le f(x) + \lambda_0 h(x).$$

Because $h(x_0) = 0$ the we obtain

$$f(x_0) + \lambda h(x_0) \le f(x_0) + \lambda_0 h(x_0) \le f(x) + \lambda h(x) \quad \forall \ \lambda \in \mathcal{R}.$$

or equivalently

$$L(x_0, \lambda) \le L(x_0, \lambda_0) \le L(x, \lambda_0).$$

and hence (x_0, λ_0) is a saddle point of L.

Conversely if (x_0, λ_0) is a saddle point of L, then the above inequalities hold. If we assume that $h(x_0) \neq 0$ then contradictions arise when one considers $\lambda \neq \lambda_0$. Hence $h(x_0) = 0$ and so $f(x_0) \leq f(x) \quad \forall x \in X$. And so $L(x_0, \lambda_0) = f(x_0)$.

1.6. Summary

In this chapter we have introduced the mathematical concepts upon which the rest of the thesis is built. The main concepts one needs from this chapter are:

- Lebesgue spaces extend the range of integrable functions. The optimisation problems one comes across in DFT are defined over Lebesgue spaces.
- The dual space consists of functionals which are bounded from above. This guarantees that the energy functional in DFT is finite.

- The dual space of a Lebesgue space is easily found using Hölder conjugates. We will do this to construct the space of densities and potentials.
- Any Legendre-Fenchel transform is a duality which creates lower and upper semi-continuous functions. These resulting functionals are also convex and concave respectively. We will use this in the next chapter when we show the construction of the universal functional in Lieb's formulation of DFT which is automatically convex. We will utilise this in understanding Lieb's convex formulation of DFT in Chapter 3.
- The saddle value of a Lagrange function is equal to constrained minimum of the target function. This is central to this thesis as this shows that the two methods presented in Chapter 4 and Chapter 6 are different ways of getting to the same solution.

2. Density Functional Theory

2.1. Quantum Mechanics

In the early 20th century the scientific world had a problem. We had discovered the existence of the atom, and its constituents, yet the contemporary theoretical models, such as classical mechanics, predicted atoms cannot exist. If one computed the trajectories of the electrons using classical mechanics one would discover that the electrons orbit would collapse to an infinitesimal size. Thus everything around us, including ourselves, would cease to exist! So in the early part of the 20th century a revolutionary theory which adequately describe systems of electrons, and other sub-atomic and atomic systems, was created — quantum mechanics.

The fundamental axiom of quantum mechanics is:

Definition 2.1.1. The state of the system contains the probability distribution of all experimentally observable quantities. We call this the (pure) quantum state and is denoted by $|\Psi\rangle$.

A mixed quantum state is a convex combination of pure quantum states and are best represented using density matrices, as discussed later. We can characterise the set of all possible states of a system, Q_W , through

$$\mathcal{Q}_{\mathcal{W}} = \left\{ |\Psi\rangle : |\Psi\rangle \in L_2(\mathcal{R}^3; \mathcal{C}^2), \langle\Psi|\Psi\rangle = 1 \right\}.$$
(2.1)

The first restriction on the set comes from $|\Psi\rangle \in L_2(\mathcal{R}^3; \mathcal{C}^2)$ which states that $\mathcal{Q}_{\mathcal{W}}$ is a type of Lebesgue space known as the Hilbert space. The dependence on $\mathcal{C}^{2\ddagger}$ is a result of the spin property of the electron which can take two values, $\pm 1/2$. As all the work presented in this thesis is not concerned with relativity nor external electro-magnetic fields, the distinction between $|\Psi\rangle \in L_2(\mathcal{R}^3; \mathcal{C}^2)$ and $|\Psi\rangle \in L_2(\mathcal{R}^3)$ is not so important. Therefore, we will use the latter from now on for notational ease. The second restriction, $\langle \Psi | \Psi \rangle = 1$ i.e. normalisation to unity, means that $\mathcal{Q}_{\mathcal{W}}$ is not

 $[\]overline{{}^{\ddagger} \mathcal{C}^2}$ is the space of two-dimensional complex numbers

a genuine Hilbert space[†] but we will refer to it as a Hilbert space where $\langle \Psi | \Psi \rangle = 1$ is understood. This second restriction is known as an inner product and is a duality between $|\Psi\rangle$ and $\langle \Psi | \in (L_2(\mathcal{R}^3))^* = L_2(\mathcal{R}^3)$. The Hilbert space is a reflexive Banach space where the inner product defines the norm.

So far we have discussed the quantum state in an abstract manner. To proceed we need to discuss the representations of quantum mechanics. All experimentally observable properties of the quantum state have an associated Hermitian operator. A Hermitian operator \hat{A} has an eigenbasis $\{|a\rangle\}$ if

$$\hat{A}|a\rangle = a|a\rangle \quad a \in \mathcal{R},$$
(2.2)

where a is the eigenvalue. The position operator, \hat{X} , and the momentum operator, \hat{P} , are Hermitian operators and can be defined through the eigenbasis of \hat{X} , $\{|x\rangle\}$,

$$\hat{X}|x\rangle = x|x\rangle,$$
 (2.3)

$$\hat{P}|x\rangle = -i\hbar\frac{\partial}{\partial x}|x\rangle.$$
 (2.4)

This is called the co-ordinate representation of quantum mechanics. Note that this is only one of the ways that we can represent these operators. One could define them using the eigenbasis of \hat{P} but that leads to an increase in complexity in the equations central to quantum chemistry. $|x\rangle \in L_2$ and so one can define a duality between the eigenbasis $\{|x\rangle\}$ and a quantum state $|\Psi\rangle$ as

$$\langle x|\Psi\rangle := \Psi(x) \tag{2.5}$$

$$\langle \Psi | x \rangle = \langle x | \Psi \rangle^* = \Psi^*(x). \tag{2.6}$$

 $\Psi(x)$ is known as the **wavefunction** of the system. It has a topological interpretation that $\Psi(x)$ is the component of the abstract vector $\{|\Psi\rangle\}$ along the co-ordinate representation. It has the probabilistic interpretation of being the probability amplitude for the existence of the system with co-ordinate x. The eigenbasis of the co-ordinate representation is complete, continuous and uncountably infinite which means that

$$\int |x'\rangle\!\langle x'|\,\mathrm{d}x' = 1 \tag{2.7}$$

 $^{^\}dagger$ It is subset of the Hilbert space commonly called the physical Hilbert space which retains the important topological features of the full Hilbert space

where the integral must be conducted over the entire basis. The normalisation of this basis is defined by

$$\langle x|x'\rangle = \delta(x-x') \tag{2.8}$$

where δ is the Dirac distribution. This allows us to define the normalised inner-product through

$$\langle \Psi | \Psi \rangle = \iint \langle \Psi | x' \rangle \langle x' | x \rangle \langle x | \Psi \rangle \, \mathrm{d}x \mathrm{d}x'$$

$$= \iint \Psi^*(x') \delta(x - x') \Psi(x) \mathrm{d}x \mathrm{d}x'$$

$$= \int \Psi^*(x) \Psi(x) \mathrm{d}x$$

$$= \int |\Psi(x)|^2 \mathrm{d}x = 1,$$

$$(2.9)$$

where $|\Psi(x)|^2$ is known as the probability density. The wavefunction provides an intuitive interpretation of the normalisation of the quantum state. From Equation (2.9) we can interpret $\langle \Psi | \Psi \rangle$ as the probability of measuring the particle somewhere in space which must be 1 if it exists. Hence $\langle \Psi | \Psi \rangle = 1$ must be a restriction on the Hilbert space.

In quantum chemistry we are interested in N-particle systems where N is an integer. The N-electron wavefunction is an element of $L_2(\mathcal{R}^{3N})^{\dagger}$ with norm

$$\langle \Psi | \Psi \rangle = \int |\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)|^2 \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \cdots \mathrm{d}\mathbf{r}_N = 1.$$
 (2.10)

Here $\mathbf{r}_i \in \mathcal{R}^3$ is the spatial co-ordinate of the *i*th particle. For notational ease we define $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ and $d\mathbf{r} = d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N$. For the *N*-particle system the operators $\hat{\mathbf{X}}$ and $\hat{\mathbf{P}}$ are defined, using the eigenbasis of $\hat{\mathbf{X}}, \{|\mathbf{X}\rangle\}$

$$\hat{\mathbf{X}} | \mathbf{X} \rangle = \mathbf{X} | \mathbf{X} \rangle$$
 (2.11)

$$\hat{\mathbf{P}} |\mathbf{P}\rangle = -i\hbar \sum_{i=1}^{N} \boldsymbol{\nabla}_{i} |\mathbf{X}\rangle$$
(2.12)

where ∇_i is the gradient with respect to the co-ordinates of particle *i*. Using these one can define the kinetic energy operator \hat{T} and the potential

[†] Technically an N-electron quantum state is an element of $\bigotimes_{i=1}^{N} L_2^{(i)}(\mathcal{R}^3)$ which is subtly different from $L_2(\mathcal{R}^{3N})$ but too subtle to have any impact on the discussions here

energy operator, \hat{V} as

$$\hat{T}\left(\hat{\mathbf{P}}\right) = \sum_{i=1}^{N} -\frac{\hbar^2}{2m_i} \boldsymbol{\nabla}_i^2 \qquad (2.13)$$

$$\hat{V}\left(\hat{\mathbf{X}}\right) = V(\mathbf{X}) \tag{2.14}$$

for an N-electron system. The explicit form of $V(\mathbf{X})$ depends on the physical model.

We stated that the quantum state contains everything we can experimentally verify about a system. The theoretical values that one compares with experimental values are the expectation values. For a given operator \hat{O} we can define the expectation value of the operator, for a given normalised wavefunction, through

$$\left\langle \hat{O} \right\rangle_{\Psi} = \left\langle \Psi | \hat{O} | \Psi \right\rangle = \int \Psi^*(\mathbf{r}) \hat{O} \Psi(\mathbf{r}) \mathrm{d}\mathbf{r}.$$
 (2.15)

If $\Psi(\mathbf{r})$ is an eigenfunction of \hat{O} then $\langle \hat{O} \rangle_{\Psi}$ is just the corresponding eigenvalue.

In quantum chemistry we are usually interested in computing the energy of the system. This is yielded through the Hamiltonian operator, \hat{H} , defined by

$$\hat{H} = \hat{T} + \hat{V} \tag{2.16}$$

with the energy given by

$$E_{\Psi} = \left\langle \hat{H} \right\rangle_{\Psi} = \left\langle \hat{T} \right\rangle_{\Psi} + \left\langle \hat{V} \right\rangle_{\Psi}.$$
 (2.17)

From the spin-statistics theorem we have the following property for fermions

Theorem 2.1.1. If Ψ is a wavefunction describing a fermionic system then swapping the space and spin coordinates of two identical fermions leads to a change in sign of Ψ .

This antisymmetry requirement is known as the **Pauli principle**. The simplest example of an antisymmetric function is obtained by taking any N one-electron spin orbitals, $\phi_i(\mathbf{x}) \in L_2(\mathcal{R}^3; \mathcal{C}^2), i = 1, ..., N$, that are orthonormal i.e., $\int \phi_i(\mathbf{x})\phi_j(\mathbf{x}) = \delta_{ij}$. Here \mathbf{x} denotes spin-space coordinates. These spin orbitals are simply a product of a spatial orbital times a spin orbital i.e. $\phi_i(\mathbf{x}) = \phi_i(\mathbf{r}) |\sigma\rangle$, where $|\sigma\rangle$ is the spin orbital. The N-particle wave function in this case is given by

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{(N!)}} \det\{\phi_i(\mathbf{x}_j)\}_{i,j=1}^N,$$
(2.18)

and is known as the **Slater determinant**. As the Hamiltonian used in this work does not act upon the spin-orbitals we can effectively drop the explicit dependence on the spin orbitals and just treat the quantum states as antisymmetric elements of $L_2(\mathcal{R}^{3N})$. This subspace is denoted by $\bigwedge^N L_2(\mathcal{R}^3)$ where \bigwedge denotes the antisymmetric wedge product.

Throughout this thesis we are interested in finding the ground state energy. Formally we can do this using the *Variation Principle*

Theorem 2.1.2.

$$E_0 = \inf_{\Psi} \left\{ \left\langle \hat{H} \right\rangle_{\Psi} | \Psi \in \bigwedge^N L_2(\mathcal{R}^3), \int_{\mathcal{R}^{3N}} |\Psi|^2 d\mathbf{r} = 1 \right\},$$
(2.19)

$$\Psi_0 = \underset{\Psi}{\operatorname{arg inf}} \left\{ \left\langle \hat{H} \right\rangle_{\Psi} | \Psi \in \bigwedge^N L_2(\mathcal{R}^3), \int_{\mathcal{R}^{3N}} |\Psi|^2 d\mathbf{r} = 1 \right\}.$$
(2.20)

Where E_0 is the ground state energy and Ψ_0 is the ground state wavefunction of the system.

2.2. Electronic Structure Theory

The Born-Oppenheimer approximation (BOA) is an important approximation which simplifies the description of quantum states. The BOA is based on the observation that the nucleus of an atom is thousands of times heavier than its electrons. This allows one to treat the nucleus' position as fixed in space with respect to all electrons. This means the Hamiltonian is only parametrically dependent on the position of the nuclei, and so the wavefunction of the system can be written as a product of the nuclear and electronic wavefunctions. Therefore, in quantum chemistry we fix the nuclei positions and solve for the electronic wavefunction. The N-electron Hamiltonian required to yield electronic wavefunctions is, in atomic units,

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^{N} \nabla_{i}^{2} + \sum_{i=1}^{N} \upsilon(\mathbf{r}_{i}) + \sum_{1 \le i \le j \le N} \frac{1}{|\mathbf{r}_{i} - \mathbf{r}_{j}|}$$
(2.21)

where we have neglected the kinetic energy due to the nuclei since in the BOA nuclei are stationary. Therefore, to get the total energy of the system one needs to find E_0 using Eq. (2.21) and Theorem 2.1.2 and then add the constant nuclear-nuclear repulsion energy.

The external potential $v : \mathcal{R}^3 \mapsto \mathcal{R}$ corresponds to the Coulomb potential generated by the nuclei. Its explicit form is

$$\upsilon(\mathbf{r}_i) := -\sum_{A=1}^{M} \frac{z_A}{|\mathbf{r}_i - \mathbf{r}_A|}$$
(2.22)

for a molecular system consisting of M nuclei, where nucleus A has a charge z_A and is located at \mathbf{r}_A . For all but single electron systems we cannot solve the resulting variational problem, defined in Theorem 2.1.2, analytically using this choice of Hamiltonian.

2.2.1. Hartree-Fock Theory

The first approximate variational model is known as the **Hartree-Fock** (HF) model.^{17,18} It begins with the ansatz that we should constrain the variational space from $\bigwedge^{N} L_2(\mathcal{R}^3)$ to the space of all Slater determinants $\{\Phi\}$, defined in Eq. (2.18).

Let Φ belong to the set of all molecular orbital (MO) configurations , $\mathcal{M}_{\mathcal{O}}$

$$\mathcal{M}_{\mathcal{O}} := \left\{ \Phi = \{\phi_i\}_{1 \le i \le N}, \phi_i \in L_2(\mathcal{R}^3), \int_{\mathcal{R}^3} \phi_i(\mathbf{r})\phi_j(\mathbf{r}) \mathrm{d}\mathbf{r} = \delta_{ij}, 1 \le i, j \le N \right\},$$
(2.23)

where ϕ_i is once again a one-electron orbital. In the HF model we find the ground state energy and wavefunctions of a system through

$$\inf_{\Phi \in \mathcal{M}_{\mathcal{O}}} E_{\mathrm{HF}}(\Phi), \qquad (2.24)$$

where the HF energy functional is given by

$$E_{\rm HF}(\Phi) = \sum_{i=1}^{N} \int |\nabla \phi_i|^2 d\mathbf{r}_i + \sum_{i=1}^{N} \int v |\phi_i|^2 d\mathbf{r}_i + \frac{1}{2} \left(\sum_{i,j=1}^{N} D(|\phi_i|^2, |\phi_j|^2) - D(\phi_i \phi_j^*, \phi_j \phi_i^*) \right)$$
(2.25)

where

$$D(\rho_1, \rho_2) = \int \frac{\rho_1(\mathbf{r})\rho_2(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}'.$$
(2.26)

The energy terms in Eq. (2.25) are the kinetic energy, electron-nuclei interaction energy, the Coulomb and exchange energy respectively. The mathematical properties of this optimisation problem have been studied^{19–21} and it has been shown that a ground state in the HF optimization problem exists for both neutral systems and positive ions.

An issue here is that the equations of the HF model are only defined up to a unitary matrix transformation on Φ . This is a type of gauge-invariance and can be removed by formalising HF theory in terms of density matrices, $\tau_{\mathcal{D}}$

$$\tau_{\mathcal{D}}(x,y) = N \int_{\mathcal{R}^{3N-1}} \Psi(x,(\mathbf{r}-\mathbf{1})) \Psi^*(y,(\mathbf{r}-\mathbf{1})) \mathrm{d}(\mathbf{r}-\mathbf{1}), \qquad (2.27)$$

where $(\mathbf{r} - \mathbf{1}) = \mathbf{r}/r_1$. The corresponding one-electron density is denoted by $\rho_{\mathcal{D}}(x) = \tau_{\mathcal{D}}(x, x)$. The density matrix of an *N*-particle $\Psi \in L_2(\mathcal{R}^{3N})$ can be associated with a density operator, \mathcal{D} , defined through

$$\operatorname{Tr}(\mathcal{D}) = N \quad \text{and} \quad 0 \le \mathcal{D} \le 1,$$
 (2.28)

such that the electron density, $\rho_{\mathcal{D}}$, is positive and is normalised to the number of electrons in the system. It can be shown there is a bijective map between the set of finite energy Slater type density operators^{†‡}

$$\mathcal{P} = \left\{ \mathcal{D} \in \mathcal{L}_1 | \operatorname{Ran}(\mathcal{D}) \subset H^1(\mathcal{R}^3), \ \mathcal{D}^2 = \mathcal{D}, \ \operatorname{Tr}(\mathcal{D}) = N \right\}$$
(2.29)

and the set of all molecular orbital configurations up to a unitary matrix. We can now rewrite Eq. (2.25) in the gauge-variant — but energetically equivalent — density matrix form

$$\mathcal{E}_{\rm HF}(\mathcal{D}) = {\rm Tr}(h\mathcal{D}) + \frac{1}{2} {\rm Tr}(G(\mathcal{D}) \cdot \mathcal{D})$$
 (2.30)

where $h = -\nabla^2 + v$ and for all $\Psi \in L_2$ and for all $x \in \mathcal{R}^3$

$$(G(\mathcal{D}) \cdot \Psi) := \left(\rho_{\mathcal{D}} * \frac{1}{|y|}\right)(x)\Psi(x) - \int \frac{\tau_{\mathcal{D}}(x,y)}{|x-y|}\Psi(y)\mathrm{d}y, \qquad (2.31)$$

where * denotes the convolution operation. The minimization problem of HF becomes

$$\inf_{\mathcal{D}\in\mathcal{P}} \{\mathcal{E}_{\mathrm{HF}}(\mathcal{D})\}.$$
 (2.32)

[†] $H^1(\mathcal{R}^3)$ is the first-order Sobolov space and \mathcal{L}_1 denotes the linear space of trace-class operators on $L_2(\mathcal{R}^3)$ [‡] Let X be a topological space. Then $\operatorname{Ran}(X)$ is denotes a topological space consisting of all nonempty finite sub-sets of X.

One can also define the Fock operator, \mathcal{F} associated with \mathcal{D} by

$$\mathcal{F}(\mathcal{D}) = h + \mathcal{G}(D), \qquad (2.33)$$

so that

$$(\mathcal{F}(\mathcal{D})\dot{\phi})(x) = -\boldsymbol{\nabla}^2 \phi(x) + \upsilon(x)\phi(x) + \left(\rho_{\mathcal{D}} * \frac{1}{|y|}\right)(x)\phi(x) - \int \frac{\tau_{\mathcal{D}}(x,y)}{|x-y|}\phi(y)\mathrm{d}y.$$
(2.34)

which is the potential of the HF energy. Cancés and Le ${\rm Bris}^{\,21}$ proved the following theorem

Theorem 2.2.1. The following statements are equivalent:

- 1. Φ is a stationary point of E_{HF} on $\mathcal{M}_{\mathcal{O}}$
- 2. There exists a hermitian matrix λ_{ij} such that $\forall i \in [1, N]$

$$\mathcal{F}(\mathcal{D}) \cdot \phi_i = \sum_{j=1}^N \lambda_{ij} \phi_j$$

3. There exists a unitary matrix, U such that $\Psi = U\Phi$ and

$$\mathcal{F}(\mathcal{D}) \cdot \psi_i = -\epsilon_i \psi_i$$

4. $\mathcal{F}(\mathcal{D})$ and \mathcal{D} commute.

Both 2 and 3 are called the HF equations in the molecular orbital basis. The latter is known as the canonical HF equations. The set of eigenvectors $\{\psi_i\}$ form the ground state determinant — known as the HF ground state — and ϵ_i are known as the HF orbital energies. In practice we solve these equations by expanding the ψ_i , the MOs, in a finite basis of one-electron functions — known as atomic orbitals.

In the basis set approximation we assert that each MO can be expressed as a linear combination over a set of one-electron functions, $\{\eta_j\}$, called **basis functions** through

$$\phi_i(\mathbf{r}) = \sum_{j=1}^{N_{\text{bas}}} a_{ij} \eta_j(\mathbf{r})$$
(2.35)

where N_{bas} is the number of basis functions in the **basis set**. a_{ij} is known as the **basis coefficient** which tells us the fraction of η_j which will contribute to ϕ_i . This means instead of varying the MOs one can vary these basis coefficients and keep the basis functions fixed which is a lot easier for computational implementations of the HF method. There are a vast range of possible choices of the basis set. One might be tempted say that we should just use the eigenbasis of the one electron Hamiltonian as the choice for one-electron functions. These are given through

$$\eta_j^{(L)}(\mathbf{r}) = \mathcal{K} r_x^a r_y^b r_z^c \exp\{-\zeta \mathbf{r}\}$$
(2.36)

where \mathcal{K} is the normalisation factor and a + b + c = L which is the angular momentum of the orbital. ζ controls the width of the orbital (large ζ gives tight functions, small ζ gives diffuse functions). These are known as Slater type orbitals (STOs). One advantage for using STOs is that the resulting MOs have a correct radial behaviour. The problem with them is that computing the integrals in the HF equations using higher angular momentum STOs (which is required for accurate energies, for example) is expensive compared to a more popular type of basis function - the **Gaussian type orbital** (GTO). These have the form

$$\eta_j^{(L)}(\mathbf{r}) = \mathcal{K} r_x^a r_y^b r_z^c \exp\left\{-\zeta \mathbf{r}^2\right\}$$
(2.37)

and greatly simplifies the evaluation of integrals involving higher angular momentum basis functions. However, by changing the radial dependency of the basis functions the resulting MOs now have physically incorrect radial behaviour of the probability density. To fix this instead of a one-to-one replacement of STOs with GTOs in the basis expansion we replace each STO with a linear combination of GTOs. This is known as a contraction scheme and does improve the radial dependency of the resulting MOs.

The HF equations are non-linear in the sense that the operators are functions of our solutions, the set of MOs or expansion coefficients in the basis set approximation. In quantum chemistry we tackle this non-linear optimisation problem using a method called the **self-consistent field** (SCF). This is an iterative fixed point procedure where we build a pseudo-Fock operator \mathcal{F}_n at the current iteration using density operators computed in previous iterations. Using this \mathcal{F}_n we define a new density operator \mathcal{D}_{n+1} . We continue in this manner until there is a negligible change in the density matrix and/or energy. To make this more explicit we will consider the SCF algorithm called the direct inversion of iterative sub-space^{22,23} (DIIS). To solve the HF equations, using the DIIS method, we compute \mathcal{F}_n using a sample of density operators from previous iterations. To make this even more explicit we say

$$\mathcal{F}_n = \sum_{k=0}^m c_k^{(n)} \mathcal{F}(\mathcal{D}_k), \qquad (2.38)$$

where m is the number of previous iterations we are using. The coefficients $c_k^{(n)}$ are computed by solving the quadratic minimization problem

$$\inf\left\{ \left| \left| \sum_{k=0}^{m} c_{k}^{(n)} e_{k} \right| \right|_{L_{2}}^{2} \left| \sum_{k=0}^{m} c_{k}^{(n)} = 1 \right. \right\}$$
(2.39)

where the error vector is defined through

$$e_k = \mathcal{F}(\mathcal{D}_k)\mathcal{D}_k - \mathcal{D}_k\mathcal{F}(\mathcal{D}_k).$$
(2.40)

One can see from the fourth item in Theorem 2.2.1 the error vector goes to zero as we close in on the stationary point on the HF energy surface.

The HF minimization problem takes place on a subset of the set of all possible ground state wavefunctions. This means the HF energy is an upper bound to the true ground state energy due to the variational principle. We can define the energy difference $E - E_{\rm HF}$ as the **correlation energy**. The correlation energy is actually due to two effects - **static** and **dynamic** correlation. The latter can be understood using the physical interpretation of the HF equations. The HF model can be interpreted as stating that each electron interacts with the average electron density cloud of all other electrons in the system. This clearly means that if this electron moves in space the other electrons cannot individually respond instantaneously and so the resulting energy is too high in the HF model.

Static correlation is related to the fact that in certain circumstances the ground state Slater determinant is not a good approximation to the true ground state. A typical example is provided by one of the famous laboratories in quantum chemistry, the H₂ molecule. As we stretch the bond the correlation energy in the limit of very large distances is around 0.25 E_h . This cannot be all due to dynamical correlation as the electrons on each hydrogen atom very weakly interact with each other. If one expands the determinant at these large distances we find that the HF model predicts with equal probability that we end up with a proton and a H⁻ atom or two neutral H atoms. The fact that the HF wave function even at large internuclear distances consists of 50% of ionic terms, even though H₂ dissociates into two neutral hydrogen atoms, leads to an overestimation of the interaction energy and finally to the large error in the dissociation energy.

2.2.2. Coupled-Cluster Theory

One of the main tenets of electronic structure theory is how does one best incorporate this small but important electron correlation energy into our models. One such way is known as **Coupled-Cluster** (CC) theory²⁴ which was developed by Coester and Kummel^{25–28} and about ten years later the CC approach was transferred to the field of electronic structure by Sinanoğlu, Čižek, Paldus and Shavitt.^{29–31} CC theory lifts the restriction of writing the state as a single determinant. We begin by assuming we have been given a one-electron basis and we take the HF state, $|\text{HF}\rangle$, as our reference state. Within a Hilbert space spanned by the one-electron basis, any state $|\Psi\rangle : \langle \text{HF} |\Psi\rangle = 1$, can be expanded by

$$|\Psi\rangle = (1 + U_1 + U_2 + \dots + U_n) |\mathrm{HF}\rangle \tag{2.41}$$

where U_n is an excitation operator which creates n particle-hole pairs relative to the HF ground state. The above expression is equivalent to the exponential form

$$|\Psi\rangle = \exp(T) |\text{HF}\rangle \quad T = \sum_{n} T_{n}$$
 (2.42)

where T_n again is an excitation operator. All excitation operators commute and $\exp(T)$ describes all possible independent excitations from the HF ground state. $|\Psi\rangle$ is known as the coupled-cluster state, which we will rename $|\text{CC}\rangle$. The coupled-cluster wavefunction is given through $\Psi_{\text{CC}}(x) = \langle x | \text{CC} \rangle$. To compute the energy, E_{CC} one considers the following equation

$$\exp(-T)\hat{H}\exp(T)|\mathrm{HF}\rangle = E_{\mathrm{CC}}|\mathrm{HF}\rangle. \qquad (2.43)$$

Projecting against the excited and HF reference states one obtains the coupled-cluster equations

$$E_{\rm CC} = \langle \mathrm{HF} | \exp(-T) \hat{H} \exp(T) | \mathrm{HF} \rangle \qquad (2.44)$$

$$0 = \langle \mu | \exp(-T) \hat{H} \exp(T) | \mathrm{HF} \rangle \tag{2.45}$$

where μ is a generic index which counts the excitations. Some calculations in the literature are done using the singles-and-doubles approximation (CCSD) where $T = T_1 + T_2$. However, the 'gold-standard' in quantum chemistry is CCSD plus an inclusion of T_3 using perturbation theory -CCSD(T).³² This thesis will be focussed on density functional methods which incorporates the correlation energy using approximate functionals of the density. CCSD will be used in Chapter 7 to generate accurate reference densities to compare with DFT results.

2.2.3. Rayleigh-Ritz Variation Principle

In this and following sections we will consider an atomic or molecular system with N-electrons with a Hamiltonian H_N given by

$$H_N(\upsilon) = T_N + W_N + \sum_{i=1}^N \upsilon(\mathbf{r}_i)$$
(2.46)

where: T_N is the first term in Eq. (2.21) and measures the total kinetic energy of an N-electron system; W_N is the third term in Eq. (2.21) and measures the total electron-electron repulsion energy of an N-electron system. The only thing that distinguishes the Hamiltonian of two different N-electron systems is the analytic form of v - the external potential - which describes how the electronic system interacts with the nuclei structure. For the Hohenberg-Kohn (HK) theorems we require that v be multiplicative and support a ground state. The set of all possible v which support a ground state for a N-electron system is denoted by \mathcal{V}_N . For a particular element of this set we denote the corresponding ground state wave-function by Ψ_v and the ground state energy by $E_0: \mathcal{V}_N \to \mathcal{R}$. Let \mathcal{W}_N be defined in the following manner

$$\mathcal{W}_{N} = \left\{ \Psi | \Psi \in \bigwedge^{N} L_{2}(\mathcal{R}^{3}), \int_{\mathcal{R}^{3N}} |\Psi(\mathbf{r})|^{2} \mathrm{d}\mathbf{r} = 1, \langle H_{N} \rangle_{\Psi} \in \mathcal{R} \right\}.$$
(2.47)

Then we can then state the Rayleigh-Ritz Variation Principle

Theorem 2.2.2. $\forall (v, \Psi) \in \mathcal{V}_N \times \mathcal{W}_N$ then

$$E_0(v) \leq \langle \Psi | H_N(v) | \Psi \rangle$$

 $\forall v \in \mathcal{V}_N \exists \Psi_v \in \mathcal{W}_N$ as the global minimisers of the Rayleigh-Ritz variation principle, with ground state energy $E_0(v)$

$$E_0(v) = \langle \Psi_v | H_N(v) | \Psi_v \rangle, \quad \Psi_v \in \underset{\Psi \in \mathcal{W}_N}{\operatorname{arg inf}} \langle \Psi | H_N(v) | \Psi \rangle.$$

2.3. Formal Density Functional Theory

In density functional theory (DFT) we calculate the ground state energy of the system using the ground-state electronic density ρ instead of using the wavefunction Ψ through expectation values. To begin with we will focus on the formal development of DFT. By this we mean the path from the initial HK formulation of DFT to a more mathematically sound theory based on convex analysis. This will give us the theoretical foundations for the optimisers discussed later on in this work. We will also introduce Kohn-Sham DFT (KS-DFT) as a model to approximate the functionals introduced in the formal theory. We will discuss orbital-free DFT (OF-DFT) — which is more akin to the ideas of the formal theory — in the next chapter.

2.3.1. On the Development of DFT

To proceed we define the set of v-representable densities through

$$\mathcal{A}_N = \{\rho | \Psi_v \mapsto \rho\} \tag{2.48}$$

where Ψ_{υ} is the ground state of $H_N(\upsilon)$ with $\upsilon \in \mathcal{V}_N$. When we are first introduced to DFT we are taught that it rests on the following two theorems³³

Theorem 2.3.1. First Hohenberg-Kohn Theorem

There is a bijective map between \mathcal{A}_N and \mathcal{V}_N modulo a constant.

Theorem 2.3.2. Second Hohenberg-Kohn Theorem

There is a ground state energy functional $E_0: \mathcal{V}_N \to \mathcal{R}$ given by

$$E_0(\upsilon) = \min_{\rho \in \mathcal{A}_N} \left(F_{HK}(\rho) + (\upsilon|\rho) \right)$$

where $(\upsilon|\rho) := \int \upsilon(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$.

In the second theorem we have $F_{\text{HK}} : \mathcal{A}_N \to \mathcal{R}$ which is known as the **HK universal functional** and can be decomposed as as sum of $T(\rho)$ and $W(\rho)$; the kinetic energy and potential energy functional respectively. It is called universal as the F_{HK} is the same for systems with the same number of electrons.

From a practical point of view the HK theorems are not that useful as we do not know the exact form of the $F_{\rm HK}$. Another problem is that the sets \mathcal{A}_N and \mathcal{V}_N are unknown unless one solves the Rayleigh-Ritz variation principle - which is impossible for more than one electron. Furthermore, we would like to lift the restriction of a fixed number of electrons and so have a DFT formally defined over variable number of electrons. This will be important for our optimisers presented in this work. We will first consider how we extend the domain to a larger, known set of densities and potentials.

2.3.2. Constrained Search Formalisms of DFT

In DFT it is essential that there is some way to distinguish between densities, ρ , that are formed from antisymmetric functions from those that are not. The density must be normalised to the number electrons of the system and be non-negative throughout space. Therefore, $\rho \in L_N^1$ where

$$L_N^1 = \left\{ \rho | \rho \in L_1(\mathcal{R}^3), \ \rho \ge 0, \ \int_{\mathcal{R}^3} \rho(\mathbf{r}) \mathrm{d}\mathbf{r} = N \right\}.$$
(2.49)

Fortunately, one can prove that each $\rho \in L_N^1$ is the density of some *N*-electron antisymmetric wavefunction, which one can take as a Slater determinant. In DFT we are actually interested in the subset of $\rho \in L_N^1$ which can be obtained from $\Psi \in \mathcal{W}_N$. This set is known as the *N*-representable densities and is characterised by

$$\mathcal{I}_N = \{\rho | \Psi \mapsto \rho, \Psi \in \mathcal{W}_N\}.$$
(2.50)

But its seems like we have run into the same issue as for the v-representable densities — given a density how do we ascertain whether it's an element of \mathcal{I}_N ?

Firstly we can express the density $\rho \in \mathcal{I}_N$ as an inner product of $\Psi_{\rho} \in \mathcal{W}_N$

$$\rho = N \left\langle \Psi_{\rho} | \Psi_{\rho} \right\rangle. \tag{2.51}$$

Differentiation of the density with respect to the Cartesian basis denoted by ∂_{α} is

$$|\partial_{\alpha}\rho| \le 2\sqrt{N\rho} [\langle\partial_{\alpha}\rho|\partial_{\alpha}\rho\rangle]^{\frac{1}{2}}.$$
(2.52)

This can be written as

$$\frac{|\boldsymbol{\nabla}\rho|^2}{8\rho} \ge \frac{N}{2} \left\langle \boldsymbol{\nabla}\rho | \boldsymbol{\nabla}\rho \right\rangle, \qquad (2.53)$$

and integrating on both sides over \mathcal{R}^3 one gets

$$\int_{\mathcal{R}^3} \frac{|\boldsymbol{\nabla}\rho|^2}{8\rho} \le \langle \Psi_p | T_N | \Psi_p \rangle \tag{2.54}$$

where the left hand side defines the **von-Weizsäcker** functional $T_{\rm VW}$: $L_N^1 \to (0, +\infty]$.³⁴ One immediately arrives at a useful theorem

Theorem 2.3.3. The von Weizsäcker kinetic energy of $\rho \in \mathcal{I}_N$ is a lower bound to the kinetic energy of all N-electron wave functions with the same density:

$$T_{VW}(\rho) \leq \inf_{\Psi \mapsto \rho} \langle \Psi | T_N | \Psi \rangle < \infty, \quad \rho \in \mathcal{I}_N.$$

Furthermore, using $T_{\rm VW}$ we can explicitly define \mathcal{I}_N

Theorem 2.3.4. The set of N-representable densities is defined as

$$\mathcal{I}_N = \left\{ \rho | \rho \in L_1(\mathcal{R}^3), \ \rho \ge 0, \ \int_{\mathcal{R}^3} \rho = N, T_{VW}(\rho) < +\infty \right\}$$

As the set is explicitly defined we want to rewrite the HK variation principle over this set. This means we need to extend the $F_{\rm HK}$ to $F_{\rm LL} : \mathcal{I}_N \to \mathcal{R}$ — this is known as the **Levy-Lieb constrained search functional**.^{35–37} The analytical form is given by

$$F_{\rm LL}(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi | T_N + W_N | \Psi \rangle . \qquad (2.55)$$

One can show that a minimising wavefunction does exist for all $\rho \in \mathcal{I}_N$ and so one can replace the infimum with a minimum. Furthermore, one can show that F_{LL} is always positive and always has a minimising *N*representable density. Without going into the details we will now restrict the set of potentials \mathcal{V}_N to the subset of such potentials which have a Coulombic like form, \mathcal{V}_C .

The ground-state energy of a system in an external Coulomb potential may now be expressed as a HK variation principle over all N-representable densities

Theorem 2.3.5. Hohenberg-Kohn variation principle with F_{LL}

$$E_0(\upsilon) = \inf_{\rho \in \mathcal{I}_N} (F_{LL}(\rho) + (\upsilon|\rho)), \quad \forall \upsilon \in \mathcal{V}_C.$$

We are now in a position that we have an optimisation problem over welldefined sets. However, $F_{\rm LL}$ is non-convex and thus is quite difficult to use in variational approaches. How can we solve this? The answer is that we alter our approach to formal DFT from considering pure states to considering ensemble states which we will discuss in Section 2.3.4. However, we first need to consider the topological properties of our new spaces of densities and potentials.

2.3.3. Topology of Density and Potential Sets

In the HK variation principle using F_{LL} we have introduced the sets \mathcal{I}_N and \mathcal{V}_C but we do not yet know their topological structure. We will discover that both are Banach spaces with very simple definitions. Firstly, we consider \mathcal{I}_N . In order for the normalisation and the kinetic energy of the density to be finite we require

$$\rho \in L_1 \quad |\nabla \rho^{\frac{1}{2}}| \in L_2. \tag{2.56}$$

We would like to transform the restriction on $|\nabla \rho^{1/2}|$ to a restriction on ρ . Using Holders inequality — see Section 1.2.3 — one can show

$$\left\| |\boldsymbol{\nabla} \rho^{\frac{1}{2}}| \right\|_{2}^{2} \ge \frac{1}{36} ||\rho||_{3}.$$
 (2.57)

By comparing with the definition of $T_{\rm VW}$ we see that

$$T_{\rm VW}(\rho) \ge \frac{1}{72} ||\rho||_3,$$
 (2.58)

and hence $\rho \in L_3$ as T_{VW} is bounded from below. Therefore, one can conclude that \mathcal{I}_N must be a subset of the Banach space

$$\mathcal{X} = L_{(3,1)} = L_3 \cap L_1, \tag{2.59}$$

but it must be a strict subset as $\int \rho = N$ and $\rho \ge 0$ also. So one defines the following set

$$\mathcal{R}_N = \left\{ \rho | \rho \in \mathcal{X}, \rho \ge 0, \int \rho = N \right\},$$
(2.60)

and so

Theorem 2.3.6. The set of N-representable densities is characterised through

$$\mathcal{I}_N = \{\rho | \rho \in \mathcal{R}_N, T_{VW}(\rho) < +\infty\}$$

To characterise \mathcal{V}_C first note that in order for the HK variation principle to yield a finite value one must have that $(v|\rho)$ is finite. Let us now define

$$T_{\upsilon}(\rho) := (\upsilon|\rho), \tag{2.61}$$

such that $T_v(\rho)$ is a bounded linear functional, and so is a member of the dual space \mathcal{X}^* . From the theory of Lebesgue spaces, see Section 1.2.5, we can write

$$\mathcal{X}^* = (L_{(3,1)})^* = L_{(3/2,\infty)} = L_{3/2} + L_{\infty}.$$
(2.62)

 \mathcal{V}_C is a strict subset of \mathcal{X}^* as only potentials in \mathcal{X}^* which support a ground state can be also in \mathcal{V}_C . Therefore, \mathcal{V}_C is characterised through

Theorem 2.3.7. The set of admissible Coulomb-like potentials for an Nelectron system is given by

$$\mathcal{V}_C = \{ v | v \in \mathcal{X}^*, H_N(v) \text{ supports a ground state} \}$$

2.3.4. Lieb Ensemble DFT

As stated previously the $F_{\rm LL}$ functional is non-convex. It will transpire that generalising the Levy-Lieb DFT formalism from pure to ensemble states will yield a functional that is convex.

So far our treatment of quantum systems has been using the pure state $|\Psi\rangle$ and its density matrix $|\Psi\rangle\langle\Psi|$. However, for systems which have degenerate ground states — like the O atom — this is not an appropriate treatment. Instead one should be thinking of treating the system using an ensemble density matrix, γ which is a convex combination of pure state density matrices. The set of all possible ensemble density matrices which can describe an N-electron system is denoted by \mathcal{D}_N . So, for any potential $\upsilon \in L_{3/2} + L_{\infty} = \mathcal{X}^*$ one can define the ensemble state energy, $E_{\rm DM} : \mathcal{X}^* \to \mathcal{R}$ as

$$E_{\rm DM}(\upsilon) = \inf_{\gamma \in \mathcal{D}_N} \operatorname{tr} \gamma H_N(\upsilon).$$
(2.63)

It is not difficult to show that the ensemble ground state energy equals the pure state energy,¹³ and so we will denote $E_{\rm DM}$ as E_0 from now on. To extend the HK theorems to ensemble states from pure states we denote \mathcal{B}_N as the set of all ensemble *v*-representable densities - i.e. a map exists between a *N*-electron ground state density matrix and an element in \mathcal{B}_N . It should be stated that all pure-state *v*-representable densities are ensemble *v*-representable but not vice versa. We can extend the HK existence theorem by stating: There is a bijective map between \mathcal{B}_N and \mathcal{V}_N modulo a constant. The HK variational principle is extended to include variations over all ensemble v-representable densities. The details aren't important. What is important is that these extensions restore the problem of unknown domains — so once again we turn to the Levy-Lieb constrained search.

We begin noting that the set \mathcal{I}_N is a convex set containing all ensemble densities which are obtained from some $\gamma \in \mathcal{D}_N$. One can show¹³ that \mathcal{B}_N is dense in \mathcal{I}_N and hence that all *N*-representable densities may be approximated, to any desired accuracy, by an ensemble *v*-representable density. We finally observe that since each ensemble *v*-representable density is *N*representable and, therefore, obtainable from a pure state wave function, they cannot easily be distinguished from a pure-state *v*-representable density. Therefore, a natural requirement for a satisfactory universal density functional is that it should work equally well for pure-state and ensemble densities. We can now split the ground state energy into two minimisations

$$E_0(\upsilon) = \inf_{\rho \in \mathcal{I}_N} \inf_{\gamma \mapsto \rho} \left(\operatorname{tr} \gamma H_N(0) + (\upsilon|\rho) \right)$$
(2.64)

where we can now introduce the Lieb ensemble state constrained-search functional, F_{DM} : $\mathcal{I}_N \to \mathcal{R}$

$$F_{\rm DM}(\rho) = \inf_{\gamma \mapsto \rho} \operatorname{tr} \gamma H_N(0) \tag{2.65}$$

which allows us to re-write the HK variation principle using $F_{\rm DM}$.

Theorem 2.3.8. Hohenberg–Kohn variation principle with F_{DM} $\forall v \in \mathcal{X}^*$ one can define the ground state energy $E_0 : \mathcal{X}^* \to \mathcal{R}$ via

$$E_0(\upsilon) = \inf_{\rho \in \mathcal{I}_N} \left(F_{DM}(\rho) + (\upsilon|\rho) \right)$$

One can also prove¹³ that for every N-representable density a minimising density matrix exists for $F_{\rm DM}$. If this density is also ensemble vrepresentable this density matrix minimises the expectation value of the Hamiltonian. In the definitions above its not entirely clear what the difference between $F_{\rm DM}$ and $F_{\rm LL}$ is. Even though the domains of $F_{\rm DM}$ and $F_{\rm LL}$ are the same the density matrix search defining $F_{\rm DM}$ is more exhaustive than the wavefunction search defining $F_{\rm LL}$ and so $F_{\rm DM} \leq F_{\rm LL}$. Moreover, we can say more. Consider the following:

$$F_{\rm DM}(\rho) = \inf_{\gamma \mapsto \rho} \operatorname{tr} \gamma H_N(0);$$

$$= \inf_{\lambda_i, \Psi_i} \left\{ \sum_i \lambda_i \langle \Psi_i | H_N(0) | \Psi_i \rangle : \rho = \sum_i \lambda_i \rho_i, \ \lambda_i \ge 0, \ \sum_i \lambda_i = 1 \right\};$$

$$= \inf_{\lambda_i, \rho_i} \left\{ \inf_{\Psi \mapsto \rho_i} \left[\sum_i \lambda_i \langle \Psi_i | H_N(0) | \Psi_i \rangle \right] : \rho = \sum_i \lambda_i \rho_i, \ \lambda_i \ge 0, \ \sum_i \lambda_i = 1 \right\};$$

$$= \inf_{\lambda_i, \rho_i} \left\{ \sum_i \lambda_i F_{\rm LL}(\rho_i) : \rho = \sum_i \lambda_i \rho_i, \ \lambda_i \ge 0, \ \sum_i \lambda_i = 1 \right\}.$$

$$(2.66)$$

Then — using our knowledge of convex analysis (see Section 1.3.1) — it is clear that $F_{\rm DM}$ is the convex hull of $F_{\rm LL}$, so $F_{\rm DM}$ is convex and the closest convex approximation of $F_{\rm LL}$. $F_{\rm DM}$ is very satisfactory for us as it is convex and defined for every *N*-representable density - pure or ensemble. This means that we enforce the existence of only global minima in the HK variation principles.

2.3.5. Lieb's Convex-Conjugate Theory

The pinnacle of formal DFT takes form using work by Lieb⁸ who applied convex analysis to derive the Lieb functional $F_{\rm L}$ which is the Γ -regularisation of all possible universal functionals. In other words $F_{\rm L}$ is the closed convex hull of all possible universal functionals. As $F_{\rm L}$ has to be bounded from below it cannot ever equal $-\infty$ over its domain and, therefore, is an element of Γ_0 , the set of all proper semi-continuous convex functions. See Section 1.4 for details. This allows for genuine convex analysis of DFT which is crucial when we look to the optimisations presented in this work. The corresponding HK variation principle is

Theorem 2.3.9. Let E_0 be a weak-star closed concave function on \mathcal{X}^* and let $F_L \in \Gamma_0(\mathcal{X})$ then

$$E_0(\upsilon) = \inf_{\rho \in \mathcal{X}} \left(F_L(\rho) + (\upsilon|\rho) \right)$$
$$F_L(\rho) = \sup_{\upsilon \in \mathcal{X}^*} \left(E_0(\upsilon) - (\upsilon|\rho) \right)$$

This means that the ground state energy and $F_{\rm L}$ are Legendre-Fenchel conjugates of each other. So, $F_{\rm L}$ is a weak-star lower semi-continuous function. Furthermore, $F_{\rm L} = F_{\rm DM}$ when the density is N-representable.

However, only $F_{\rm L}$ is defined for every $\rho \in \mathcal{X}$ and so allows for a more rigorous analysis using the Lebesgue spaces. For example, this allows use of such terms as semi-continuity for $F_{\rm L}$ on the underlying Lebesgue spaces and so provides the rigorous foundation for the formulation of DFT in terms of saddle functions which is discussed in Section 3.4.1.

It can be shown that the Lieb functional¹³ is bounded from below by the von Weizsäcker functional and equality between the two is achieved when the system has one electron. $F_{\rm L}$ is also strictly positive and one can prove the following statement about the optimality conditions and representability

Theorem 2.3.10. For $(\rho, v) \in \mathcal{X} \times \mathcal{X}^*$ then

$$E_0(\upsilon) = \inf_{\gamma \in \mathcal{D}_N} \operatorname{tr} \gamma H_N(\upsilon) \Leftrightarrow -\upsilon \in \partial F_L(\rho) \Leftrightarrow \rho \in \bar{\partial} E_0(\upsilon)$$

Hence $\bar{\partial}E_0(v)$ contains ensemble ground-state densities supported by vand hence that the solutions of the variational problem in Theorem 2.3.10 are exactly the same as the ensemble states arising from the Rayleigh-Ritz variational problem. One can then re-write the existence theorem in terms of $F_{\rm L}$ as

Theorem 2.3.11. The sub-differential $\partial F_L(\rho)$ determines the potential uniquely up to a constant

$$\begin{aligned} -\partial F_L(\rho) &= \upsilon + \mathcal{R}, \ \rho \in \mathcal{B}_N \\ &= \emptyset \ \rho \notin \mathcal{B}_N. \end{aligned}$$

The effective domain of $F_{\rm L}$ is \mathcal{I}_N . The algebraic interior of \mathcal{I}_N is empty which leads to problems when studying the derivatives of $F_{\rm L}$ with respect to the density. Because of this topological property of \mathcal{I}_N $F_{\rm L}$ is **nowhere differentiable**, and so when deriving the equations for approximate DFT models, for example see Section 2.4.1, we should be aware that this is mathematically unsound. Recent work by Kvaal et al.³⁸ has shown that one can use a special type of regularisation^{*} that allows us to get a functional which is the closest approximation to $F_{\rm L}$ whilst being differentiable. This method relies on moving from the infinitely dimensional Banach spaces we have used so far to finite ones, this allows one to formulate DFT in a Hilbert space and so changes the topology greatly. The downside to this is that in optimisation procedures one has now allowed the electron number to vary— which we have not allowed so far. As we shall see later we can allow

^{*} Moreau-Yoshida Regularization

this without changing the topological structures of the variational spaces. In this work, however, the approximate forms of the universal functional are differentiable and so these differentiability problems are less of an issue. The exception to this is in Chapter 7 where accurate densities are used to define accurate forms of $F_{\rm L}$, in which case regularization must be applied.

2.4. Kohn-Sham Density Functional Theory

So far we have discussed only formal DFT — we still don't know the explicit form of $F_{\rm L}$. The most common application of the formal theory is known as Kohn-Sham DFT (KS-DFT).⁵ This has been very successful and is the flavour of applied DFT used ubiquitously by computational chemists. The essence of KS-DFT is that we consider a fictitious non-interacting system which has the same density as the physical system. For this non-interacting system, the kinetic energy (KE) is easily evaluated exactly using the curvature of orbitals. The remaining contributions to $F_{\rm L}$ of the physical, interacting system may then be modelled explicitly, though approximately, in terms of the density at sufficiently high accuracy to be useful for chemical applications.

Kohn and Sham decomposed $F_{\rm L}$ in the following manner

$$F_{\rm L}(\rho) = T_{\rm s}(\rho) + J(\rho) + E_{\rm XC}(p), \quad \forall \rho \in \mathcal{I}_N, \tag{2.67}$$

where $T_s : \mathcal{I}_N \to \mathcal{R}$ is the non-interacting kinetic energy functional $(KEF), J : \mathcal{I}_N \to \mathcal{R}$ is the Hartree functional and $E_{XC} : \mathcal{I}_N \to \mathcal{R}$ is the exchange-correlation functional (XC). They are defined through the following

$$T_{\rm s}(\rho) = \min_{\gamma \mapsto \rho} \operatorname{tr} T\gamma, \qquad (2.68)$$

$$J(\rho) = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2, \qquad (2.69)$$

$$E_{\rm XC}(\rho) = F_{\rm L}(\rho) - T_{\rm s}(\rho) - J(\rho).$$
 (2.70)

 $E_{\rm XC}$ contains all the effects of exchange and correlation, including the correlation correction to the KEF. Furthermore, it is non-convex, unlike J and $F_{\rm L}$, and its exact form is unknown. A large body of research has gone into developing accurate forms of $E_{\rm XC}$.³⁹

The KS HK variational principle is given by

$$E_0(\upsilon) = \inf_{\rho \in \mathcal{I}_N} \left(T_{\mathrm{s}}(\rho) + J(\rho) + E_{\mathrm{XC}}(\rho) + (\upsilon|\rho) \right), \quad \forall \upsilon \in \mathcal{V}_C$$
(2.71)

$$T_{\rm s}(\rho) = \sup_{\upsilon_{\rm s} \in \mathcal{V}_C} \left(E_0(\upsilon_{\rm s}) - (\upsilon_{\rm s}|\rho) \right), \quad \forall \rho \in \mathcal{I}_N,$$
(2.72)

where $T_{\rm s}$ is expressed as the Legendre-Fenchel transform of the non-interacting ground state energy. The KS HK variational principle is similar to Theorem 2.3.10 but we have changed the domains of the density and potential. The first variation principle describes the physical, fully interacting system in the external potential v; the second variation principle describes a fictitious non-interacting system in an effective potential $v_{\rm s}$ defined such that the densities of the interacting and non-interacting systems are the same.

2.4.1. Solving the KS-DFT Equations

To proceed we assume that $v \in \mathcal{V}_C$ and that the density is not only interacting ensemble v-representable \mathcal{B}_N but also non-interacting ensemble v-representable \mathcal{B}_N^0 . With these assumptions the infimums and supremums in Eqs. (2.71) and (2.72) are achieved when the density and potentials vand v_s satisfy

$$-\upsilon \in \partial(T_{\rm s}(\rho) + E_{\rm XC}(\rho) + J(\rho)) \tag{2.73}$$

$$-v_{\rm s} \in \partial(T_{\rm s}(\rho)). \tag{2.74}$$

Assuming the differentiability of the above we obtain the following differential equations

$$\frac{\delta T_{\rm s}(\rho)}{\delta \rho(\mathbf{r})} = (-\upsilon - \upsilon_{\rm J}(\rho) - \upsilon_{\rm XC}(\rho))(\mathbf{r})$$
(2.75)

$$\frac{\delta T_{\rm s}(\rho)}{\delta \rho(\mathbf{r})} = -\upsilon_{\rm s}(\mathbf{r}) \tag{2.76}$$

where

$$\upsilon_{\mathrm{J}}(\rho)(\mathbf{r}) := \frac{\delta J(\rho)}{\delta \rho(\mathbf{r})}$$
$$\upsilon_{\mathrm{XC}}(\rho)(\mathbf{r}) := \frac{\delta E_{\mathrm{XC}}(\rho)}{\delta \rho(\mathbf{r})}.$$

This means the effective potential v_s is equal to the sum of the Hartree potential, v_J , external potential, v, and the XC potential, v_{XC} . We can compute v_s through the ground state density yet we require v_s to calculate

54

the ground state density through

$$E_0(v_s) = \min_{\rho \in \mathcal{I}_N} (T_s(\rho) + (v_s|\rho)).$$
 (2.77)

So like the HF model we resort to solving this minimisation problem using the self-consistent field. To see how one does this in application it's convenient to write the minimisation problem using the ensemble formulation discussed in the previous section where γ is an ensemble density matrix,

$$E_0(\upsilon_s) = \min_{\gamma \mapsto N} \operatorname{tr} H_0(\upsilon_s)\gamma = H_0(\upsilon_s)\hat{\gamma}_s.$$
(2.78)

The minimising density matrix, $\hat{\gamma}_s$ is a convex combination of *N*-electron Slater determinants, Φ_k which are degenerate for $H_0(v_s)$. This can be expressed as

$$\hat{\gamma}_{s} = \sum_{k=1}^{m} p_{k} |\Phi_{k}\rangle \langle \Phi_{k}|,$$

$$p_{k} \ge 0, \quad \sum_{k=1}^{m} p_{k} = 1.$$
(2.79)

Therefore, we can rewrite Eq. (2.77) as

$$E_0(\upsilon_s) = \sum_k \lambda_k \left\langle \Phi_k | H_0(\upsilon_s) | \Phi_k \right\rangle, \lambda_k \ge 0, \sum_k \lambda_k = 1.$$
(2.80)

This is just a convex combination of expectation values for the set of degenerate determinants. As Slater determinants are comprised of a set of orthonormal spin orbitals, $\{\phi_p\}$ we can expand the expectation values as expectation values of one-electron operators over these spin orbitals. This yields,

$$\langle \Phi_k | H_0(\upsilon_{\rm s}) | \Phi_k \rangle = \sum_{p=1}^m n_{k_p} \langle \phi_p | f | \phi_p \rangle ,$$

$$f(\mathbf{r}) = -\frac{1}{2} \nabla^2 + \upsilon_{\rm s}(\mathbf{r}),$$

$$\langle \phi_q | \phi_r \rangle = \delta_{qr}$$

$$(2.81)$$

where m is the total number of spin orbitals used in constructing the Slater determinant's and n_{k_p} is the occupation number of ϕ_p in Φ_k . By introducing the spin occupation numbers $n_p = \sum_k n_{k_p} p_k$ one arrives at

$$E_0(\upsilon_s) = \sum_{p=1}^m n_p \langle \phi_p | f | \phi_p \rangle ,$$

$$n_p \in [0, 1], \sum_p n_p = N,$$

$$\langle \phi_q | \phi_r \rangle = \delta_{qr}.$$
(2.82)

The density can be found through summing over all orbital contributions weighted by the occupation number via

$$\rho(\mathbf{r}) = \sum_{p=1}^{m} n_p \phi_p^*(\mathbf{r}) \phi_p(\mathbf{r}).$$
(2.83)

In KS-DFT one wishes to find the set of MOs $\{\phi_i\}$ which minimises $E_0(v_s)$. Using the theory of Lagrange multipliers it is easy to show that this minimising set satisfies the following KS equations

$$f_{\rm KS}\phi_i = \epsilon_i\phi_i \tag{2.84}$$

where $f_{\rm KS}$ is the Kohn-Sham operator and is given by

$$f_{\rm KS} = -\frac{1}{2}\boldsymbol{\nabla}^2 + \upsilon_{\rm s}(\mathbf{r}). \qquad (2.85)$$

Therefore, by the reintroduction of orbitals we can know the exact noninteracting kinetic energy through the Laplacian operator. Like in HF theory one expands the KS equations in a basis set and solves the SCF procedure using algorithms like DIIS, which is what we have done in this work.

2.4.2. Exchange-Correlation Functionals

In KS-DFT the only error left in approximating $F_{\rm L}$ is in the choice of $E_{\rm XC}$. There are a vast array of choices but we will utilise only two in this work. One is the **Dirac exchange**^{40*} and the PBE XC functional.⁴¹ We will discuss the Dirac functional in depth in the next chapter. The PBE XC functional is constructed using ρ and $\nabla \rho$. Therefore, PBE is an example of a generalised gradient approximation (GGA) functional. The Dirac exchange is a functional of the ρ only and is an example of a local density approximation (LDA) functional. In comparison with LDAs, GGAs tend to

^{*} This is sometimes reffered to as the Slater exchange
improve total energies, atomization energies and energy barriers.^{42–47} The PBE XC functional was developed by improving the PW91 XC functional⁴² in order to yield an accurate description of the linear response of ρ and have a smoother potential. The PBE XC functional is formed from adding the PBE exchange (PBEX) and PBE correlation (PBEC) functional.

The PBEC functional takes the form

$$E_{\rm c}(\rho_{\alpha},\rho_{\beta}) = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \big[\varepsilon_{\rm c}^{\rm unif}(r_{\rm s}(\mathbf{r}),\zeta) + H(r_{\rm s}(\mathbf{r}),\zeta,t)\big]$$
(2.86)

where $r_{\rm s}$ is the local Wigner-Seitz radius, $r_{\rm s}^3(\mathbf{r}) = 3/(4\pi\rho(\mathbf{r})), \zeta = 2a(\beta_1 r_{\rm s}^{1/2} + \beta_2 r_{\rm s} + \beta_3 r_{\rm s}^{3/2} + \beta_4 r_{\rm s}^2 +)$ with $\beta_1 = 7.5957, \beta_2 = 3.5876, \beta_3 = 1.6382$, and $\beta_4 = 0.49294$.

 $t = |\nabla \rho|/(2k_{\rm s}\rho)$ is a dimensionless density gradient factor with $k_{\rm s} = 2(3\pi^{-1}\rho)^{1/6}$. $\varepsilon_{\rm c}^{\rm unif}(r_{\rm s},\zeta)$ is the exchange correlation energy per particle of a uniform electron gas and whose form in the PBE C functional is

$$\varepsilon_{\rm c}^{\rm unif}(r_{\rm s},\zeta) = -2a(1+\alpha_1 r_{\rm s})\ln[1+\frac{1}{\zeta}],\qquad(2.87)$$

where a = 0.0310907 and $\alpha_1 = 0.21370$. In Eq. (2.86) we have introduced the concept of spin polarisation in the guise of ρ_{α} and ρ_{β} which are the density profiles of the α spin electrons and β spin electrons respectively.

The analytic form of H was derived by considering: the slowly varying density limit $(t \to 0)$; the rapidly varying density limit $(t \to \infty)$; and that under the uniform scaling to the high-density limit^{*} the correlation energy must scale to a constant.⁴⁸

These conditions lead to the analytical form

$$H = \gamma \ln \left\{ 1 + \frac{\beta}{\gamma} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2 t^4} \right] \right\}$$
(2.88)

where

$$A = \frac{\beta}{\gamma} \left[\exp\left\{ -\frac{\varepsilon_{\rm c}^{\rm unif}}{\gamma} \right\} - 1 \right]^{-1} \tag{2.89}$$

with $\beta = 0.066725$ and $\gamma = (1 - \ln 2)/\pi^2$.

PBEX was constructed by considering four further conditions. Firstly under the uniform density scaling condition the exchange energy, $E_{\rm X}$, must scale⁴⁹ like λ . Secondly, the exchange energy must obey

$$2E_{\mathcal{X}}(\rho_{\alpha},\rho_{\beta}) = E_{\mathcal{X}}(2\rho_{\alpha}) + E_{\mathcal{X}}(2\rho_{\beta}).$$
(2.90)

 $\overline{\ }^{*} \ \rho(\mathbf{r}) \rightarrow \lambda^{3} \rho(\lambda \mathbf{r})$

57

The third condition was formulated by considering the fact that the linear response of the spin un-polarised uniform electron gas is well respresented by LDA exchange functionals^{50,51} but not GGAs.⁵² And finally the Lieb-Oxford bound⁵³

$$E_{\rm X}(\rho) \ge -1.679 \int \mathrm{d}\mathbf{r} \rho^{4/3}.$$
 (2.91)

In order to satisfy these conditions the PBEX functional is defined through

$$E_{\rm X}(\rho_{\alpha},\rho_{\beta}) = C_{\rm PBEX} \int d\mathbf{r} \rho^{4/3} F_{\rm X}(s)$$
(2.92)

where $s = |\nabla \rho|/(2(3\pi^2)^{1/3}\rho^{4/3})$ which is another dimensionless gradient term, $C_{\text{PBEX}} = 3^{4/3}\pi^{2/3}/4$, and the enhancement factor $F_{\text{X}}(s)$ given explicitly by

$$F_{\rm X}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \tag{2.93}$$

where $\kappa = 0.804$ and $\mu = (\beta \pi^2)/3$.

2.5. Summary

In this chapter we have discussed a range of various optimisation problems in quantum chemistry. We first began by discussing the variation principle in quantum mechanics as a means of finding the energy of a system. However, the appropriate Hamiltonian for chemical systems yields a variational problem which cannot be solved analytically. This led to introducing electronic structure theory where we discussed HF and CC theory. We then moved on to discussing various formulations of DFT in terms of optimisation problems cumulating in Lieb's ensemble DFT and Lieb's convex conjugate theory. These were introduced in Section 2.3.4 and Section 2.3.5 respectively. We arrived at $F_{\rm L}$ which has the following important properties: a domain of \mathcal{I}_N ; is nowhere differentiable; is convex; is bounded below by $T_{\rm VW}$. In the following chapters we will be treating the OF-KEFs and XC functionals as approximations to $F_{\rm L}$. We will also show how one can generalise Lieb's ensemble formulation to a DFT which allows variation in particle number. This will turn out to be crucial to understanding the optimisation scheme introduced in Chapter 6

3. Orbital Free Density Functional Theory

3.1. Orbital Free Density Functional Theory

Method development in electronic structure theory consists of reaching a suitable compromise between accuracy and cost. Accuracy is evaluated by comparing the predicted values from our theoretical methods against experimental observations. The term cost encapsulates both the time it takes to calculate values and the amount of computing power required to get this value. There is a reasonably simple and good measure of cost; how does the cost of the method scale with basis set size, $N_{\rm b}$? For example, the most expensive set of integrals to compute in the Hartree-Fock equations are the two-electron-four-centre (2e-4c) integrals

$$(\mu\nu|\rho\sigma) = \iint \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \eta^*_{\mu}(\mathbf{r}_1) \eta_{\nu}(\mathbf{r}_1) \frac{1}{\mathbf{r}_{12}} \eta^*_{\rho}(\mathbf{r}_2) \eta_{\sigma}(\mathbf{r}_2)$$
(3.1)

which are expressed over a given basis $\{\eta_{\mu}\}$. The subscript notation indicates that each function in the basis set can be used in four different 'slots' to form 2e-4c integrals. Consequently we have $\mathcal{O}(N_b^4)$ different 2e-4c integrals to compute. We, therefore, say that the cost of the HF method formally scales $\mathcal{O}(N_b^4)$. Without going into the details CCSD's computational cost scales as $\mathcal{O}(N_b^6)$ and CCSD(T)'s as $\mathcal{O}(N_b^7)$.

Usually, as one increases the accuracy of the method one increases the cost of the method. This is why finding the appropriate balance of accuracy and cost is at the core of method development in electronic structure theory. This is also why KS-DFT has become the most utilised electronic structure theory method in theoretical and experimental research. Its popularity is due to the fact that whilst KS-DFT is in general more accurate than HF[†], KS-DFT scales formally as $\mathcal{O}(N_b^3)$. KS-DFT is cheaper than HF because whilst one does need to solve the 2e-4c integrals techniques like integral

 $^{^{\}dagger}\,$ HF treats the exchange energy exactly and ignores the correlation energy. KS-DFT resorts to an approximate treatment of both the exchange and correlation energies

screening,⁵⁴ resolution of the identity^{54,55} and density fitting^{55,56} reduce the cost of the integrals to $\mathcal{O}(N_b^2)$. The most expensive step in the SCF procedure for solving the KS equations is in the diagonalisation of the KS Fock matrix, \mathbf{F}_{KS} ,

$$(\mathbf{F}_{\mathrm{KS}})_{\mu\nu} = \langle \eta_{\mu} | f_{\mathrm{KS}} | \eta_{\nu} \rangle = \langle \eta_{\mu} | -\frac{1}{2} \nabla^{2} + \upsilon_{\mathrm{s}}(\mathbf{r}) | \eta_{\nu} \rangle.$$
(3.2)

This diagonilisation step is equivalent to finding the eigenvalues of \mathbf{F}_{KS} , which are the KS orbital energies. This step has a cost $\mathcal{O}(N_{\text{b}}^3)$.

Theoretical and computational studies on biochemical and soft-matter systems have become more frequent over the last few decades. To get accurate descriptions of such systems we need to study their electronic structure. This of course means we would ideally use electronic structure methods based upon quantum theory. But for systems of these sizes we need millions of basis functions, and so, even standard KS-DFT codes have too high of a computational cost. There has been an intensive research effort to reduce the cost of already established computational algorithms for solving the KS equations. These can be classified as linear scaling approaches.^{57–66}

There are many different approaches to this but they all suffer with one problem. They have a larger prefactor than the traditional approaches to solving the KS equations. The prefactor is just a constant factor with which the scaling behaviour is multiplied by. To see how the prefactor can affect the cost of the calculation, let us consider two models with costs of $c_1 = 0.1N_b^3 + 0.1N_b^2$ and $c_2 = 10^6N_b$. We can then say a model with c_1 has cost $\mathcal{O}(N_b^3)$ and a model with c_2 is an $\mathcal{O}(N_b)$ method with a pre-factor of 10^6 . From the above discussions we would immediately say c_2 is a cheaper method. But remember that the big-O-notation means that only in the limit of an infinite basis is $\mathcal{O}(N_b^3) > \mathcal{O}(N_b)$ guaranteed. But what about for finite basis sets? For the linear scaling model above to have a lower cost than the cubic model we need a total of 3,163 basis functions. So by the time the linear model becomes cost effective it is already an expensive calculation.

This extreme example introduces the concept of the cross-over point. This point is the minimum number of basis functions for which the linear version of the model is cheaper than the original. It is not uncommon for this number to be in the thousands for linear scaling versions of KS-DFT codes. This means systems with millions of atoms are still out-of-reach for linear scaling KS-DFT codes using computational resources available to the majority of computational and experimental chemists. So we want to develop linear scaling methodologies with very low prefactors.

In the previous chapter we discussed the fact that the ground state energy of any system can be found by minimising a functional of just the electron density of the system, ρ . If we take the same partition of $F_{\rm L}$ as in KS-DFT, Equation (2.67), but instead model the non-interacting kinetic energy (KE) using the density — rather than the KS orbitals — we arrive at orbital free DFT (OF-DFT). OF-DFT scales linearly with a very low prefactor compared to linear scaling KS-DFT methods.⁶⁷ The reason is that in OF-DFT we construct the KE using just the density rather than a set of Norbitals as is done in KS-DFT. The question remains of how does one accurately model the KE as a functional of the electron density? We start by reviewing the Thomas-Fermi and von Weizsäcker KE functionals.

3.2. Thomas Fermi and von Weizsäcker Models

The Thomas-Fermi (TF)^{68,69} and the von Weizsäcker (VW) KE functionals³⁴ (KEFs) were proposed in the late 1920's and early 1930's. The TF KEF was introduced through a theoretical study of the ground state density of a gas with a uniform electron density. The analytic form of the TF functional can be derived as follows. Firstly we can divide the uniform gas into small cubes, each with sides of length l and volume $\delta V = l^3$. In each cell place δN electrons. We treat each cell as having an infinite 3D potential energy well. This means that inside the cell the potential energy is zero and at the boundaries of the cell the potential energy is infinite. By solving the corresponding Schrödinger equation, using this potential and using a unit system where the mass of the electron is one, we get the following orbital energies

$$\epsilon(n_x, n_y, n_z) = \frac{h^2}{8l^2} \left(n_x^2 + n_y^2 + n_z^2 \right)$$
(3.3)

$$=\frac{h^2}{8l^2}\mathbf{R}^2.$$
(3.4)

Where n_i are the quantum numbers and one can treat **R** as a vector in the space spanned by the quantum numbers. For large **R** we can approximate the number of energy levels smaller than ϵ by the volume of one octant of a sphere defined through **R**. Explicitly this number is given by

$$\mathcal{N}(\epsilon) = \frac{1}{8} \left(\frac{4\pi \mathbf{R}^3}{3} \right) = \frac{\pi}{6} \left(\frac{8l^2 \epsilon}{h^2} \right)^{3/2}.$$
 (3.5)

61

Therefore, the number of energy levels between ϵ and $\epsilon + \delta \epsilon$ can be expressed through

$$g(\epsilon)d\epsilon = \mathcal{N}(\epsilon + \delta\epsilon) - \mathcal{N}(\epsilon)$$

= $\frac{\pi}{4} \left(\frac{8l^2}{h^2}\right)^{3/2} \sqrt{\epsilon}\delta\epsilon + \mathcal{O}((\delta\epsilon)^2)$ (3.6)

where $g(\epsilon)$ is the density of states with energy ϵ . To compute the total energy for a cell containing δN electrons we use the following step function, which is strictly only valid for a system with a temperature of 0 Kelvin,

$$f(\epsilon) = 1, \ \epsilon < \epsilon_F$$
$$= 0, \ \epsilon > \epsilon_F$$

where ϵ_F is the Fermi energy of the system. Then to find the total energy of the electrons in a given cell, dE, it follows — noting Fermi-Dirac statistics and that we are treating spins separately— that

$$dE = 2 \int d\epsilon f(\epsilon) g(\epsilon) \epsilon$$

= $4\pi \left(\frac{2}{h^2}\right)^{3/2} l^3 \int_0^{\epsilon_F} d\epsilon \epsilon^{3/2}$
= $\frac{8\pi}{5} \left(\frac{2}{h^2}\right)^{3/2} l^3 \epsilon_F^{5/2}.$ (3.7)

Furthermore, the total number of electrons in the cell, dN, is

$$dN = 2 \int d\epsilon f(\epsilon) g(\epsilon)$$

= $\frac{8\pi}{3} \left(\frac{2}{h^2}\right)^{3/2} l^3 \epsilon_F^{3/2}$ (3.8)

Combining Equations (3.7) and (3.8) gives

$$\mathrm{d}E = C_F l^3 \left(\frac{\mathrm{d}N}{l^3}\right)^{\frac{5}{3}} \tag{3.9}$$

where $\frac{dN}{l^3} = \rho$ and $C_F \approx 2.871$. The $\rho^{5/3}$ term is very important as it satisfies the co-ordinate scaling — which we will introduce shortly — requirement for KEFs. As a result most OF-KEFs in the literature contain this $\rho^{5/3}$ quantity.

The energy in Equation (3.9) is comprised of only the kinetic energy of the electron density inside the cell because we set the potential energy to

zero. To find the total kinetic energy of the system we sum over all cells with an infinitesimal volume and over each volume treat the density as a function which yields the Thomas Fermi KEF

$$T_{\rm TF}(\rho) = C_F \int \rho^{5/3}(\mathbf{r}) \mathrm{d}\mathbf{r}.$$
 (3.10)

The Thomas Fermi universal functional, $F_{\rm TF}: \mathcal{S} \to \mathcal{R}$ is given by

$$F_{\rm TF}(\rho) = T_{\rm TF}(\rho) + J(\rho), \qquad (3.11)$$

with the energy counterpart, $\varepsilon_{\rm TF} : \mathcal{S}^* \to \mathcal{R}$ given by

$$\varepsilon_{\rm TF}(\upsilon) = \inf_{\rho \in \mathcal{S}} \left(F_{\rm TF}(\rho) + (\upsilon|\rho) \right) \tag{3.12}$$

for for a system defined with an external potential v.

In order to arrive at a complete identification of $F_{\rm TF}$ one needs to characterise S. Without going into details $S = L_{5/3} \cap L_1$ which is a superset of the Banach space $\mathcal{X} = L_3 \cap L_1$, and so, takes finite values for non-admissible densities. Note this is different from $F_{\rm L}$, and so , is the first sign $F_{\rm TF}$ is not a good approximation of $F_{\rm L}$ which is infinite for non-admissible densities. Another disparity is that $F_{\rm TF}$ is differentiable everywhere unlike $F_{\rm L}$. However, it is convex, non-negative and weakly lower semi-continuous like $F_{\rm L}$.⁷⁰

The above issues are important but technical. Historically in KS-DFT, where we develop approximations to the XC part of $F_{\rm L}$, technical issues such as domain dependence and differentiability are ignored if the final answer is accurate. The problem with $F_{\rm TF}$ is that it is not an accurate approximation of $F_{\rm L}$. Firstly the ground state density which minimises Equation (3.12) has an unphysical divergence at the nuclei rather than a finite cusp^{71} and the radial density solutions of Equation (3.12) for atoms do not have shell structure.⁷⁰ But the condemning fact about the TF model is that it predicts that molecules immediately dissociate into their atomic fragments. This is clearly problematic for chemical applications. The problem here is that the TF functional is dependent on the electron density only and not on its spatial derivatives of any order. The following theorem was first shown to be plausible by Teller⁷² on the back of some numerical findings by Sheldon.⁷³ However, Teller's work lacked the rigour required by contemporary mathematics. The rigorous proof of this important theorem was built by Lieb and Simon⁷⁴ and Balazas.⁷⁵

Theorem 3.2.1. Let the non-interacting kinetic energy be modelled in such a way that its integrand is a function of the electron density only, i.e. does not include spatial derivatives of any order. Then the resulting model predicts that atomic subsystems are more stable than the molecule formed from these atoms.

This theorem is known as Teller's non-binding theorem. We first introduced $T_{\rm VW}$ as a lower bound to $F_{\rm L}$ but historically the $T_{\rm VW}$ was first introduced to induce molecular stability into the TF theory. We do this by combining the $T_{\rm TF}$ and $T_{\rm VW}$ in a linear combination, $F_{\rm TFVW}: \mathcal{G} \to \mathcal{R}$,

$$F_{\rm TFVW}(\rho) = T_{\rm TF}(\rho) + T_{\rm VW}(\rho) + J(\rho).$$
 (3.13)

We shall explore this model more than we did for the TF model. In particular we will look at its optimisation problem which yields the energy of the system through

$$\varepsilon_{\mathrm{TFVW}}(\upsilon) = \inf_{\rho \in \mathcal{G}} \left(F_{\mathrm{TFVW}}(\rho) + (\upsilon|\rho) \right) \quad \forall \upsilon \in \mathcal{G}^*.$$
(3.14)

To ease the up-coming analysis we will express $T_{\rm VW}$ as

$$T_{\rm VW}(\rho) = \frac{1}{2} \int \left| \boldsymbol{\nabla} \rho^{1/2} \right|^2 \mathrm{d}\mathbf{r}$$
(3.15)

which is equivalent to our original definition, see Eq. (2.54), as the density is zero as $|\mathbf{r}| \to \infty$. The function space of the Thomas Fermi von Weizsäcker (TFVW) optimisation problem, for an N electron system, is

$$\mathcal{G} = \left\{ \rho^{1/2} \middle| \boldsymbol{\nabla} \rho^{1/2} \in L_2, \ \rho^{1/2} \in L_6 \cap L_{10/3}, \ \int \rho = N \right\}$$
(3.16)

where we have used the Holders inequality.^{*} In addition $\mathcal{G} \supset \mathcal{X}$. The TFVW optimisation problem has the following properties.

Theorem 3.2.2. $\varepsilon_{TFVW}(\rho)$ is strictly convex in \mathcal{G}

Theorem 3.2.3. $\varepsilon_{TFVW}(\rho)$ has a minimum ρ on the set \mathcal{G} and is unique. **Theorem 3.2.4.** Any minimizing $\rho^{1/2} \in \mathcal{G}$ satisfies

$$\left[-\frac{1}{2}\nabla^2 + W_{\rho^{1/2}}(\mathbf{r})\right]\rho^{1/2}(\mathbf{r}) = -\mu\rho^{1/2}(\mathbf{r})$$

^{*} For the mathematically inclined reader we have transposed the condition $\rho \in L_1 \cap L_{5/3}$ from TF theory for $\rho^{1/2}$. We have also used the same technique to transform the $|\nabla \rho| \in L_2$ requirement for T_{VW} into a condition on $\rho^{1/2}$ using Holders theorem and its lemma — the Sobolev inequality.

with

$$W(\rho^{1/2})(\mathbf{r}) = C_{TF}(\rho^{1/2})^{4/3} - \upsilon + |\mathbf{r}|^{-1} * (\rho^{1/2})^2$$

Where * denotes the convolution operation. Proofs of these theorems are given by Lieb.⁵³ Both Theorem 3.2.3 and Theorem 3.2.4 allow for the possibility of solving the optimisation problem using algorithms utilised in KS-DFT by swapping the Kohn-Sham potential with $W(\rho^{1/2})$ and instead of solving for a set of MOs we solve for one quantity, $\rho^{1/2}$. $F_{\rm TFVW}$ cannot be $F_{\rm L}$ as it is differentiable everywhere. However, it is a model which predicts bound states and has a minimising density which is finite at the nucleus. Furthermore, $F_{\rm TFW}$ is a weakly lower semi-continuous convex function.⁷⁰

In the TFVW model we have only so far included the KE, the Coulomb energy, Eq. (2.69), and the electron-nuclei energy, $(\upsilon|\rho)$, in the model. What about the exchange and correlation energies? In the original formulation of the TF and TFVW models one ignores the correlation energy and treats the exchange energy using the Dirac exchange functional, $K_{\rm D}: L_{4/3} \cap L_1 \to \mathcal{R}$,

$$K_{\rm D}(\rho) = C_{\rm D} \int \rho^{4/3}(\mathbf{r}) \mathrm{d}\mathbf{r}, \quad C_{\rm D} = \frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3}.$$
 (3.17)

One can combine this with F_{TFVW} to create $F_{\text{TFDW}} = F_{\text{TFVW}} - K_{\text{D}}$. Once again we find that F_{TFDW} cannot be the F_{L} as its differentiable everywhere and is not convex nor lower semi-continous⁷⁰ but is nevertheless an important functional for historic and scientific reasons. It was the first F_{L} approximation which could predict bound molecular states as well as incorporating the effects of electronic interaction.

The energy functional for the Thomas-Fermi von-Weizsäcker with Dirac exchange (TFDW) model is given by

$$\varepsilon_{\text{TFDW}}(\rho) = T_{\text{VW}}(\rho) + \int B(\rho(\mathbf{r})) d\mathbf{r} + (\upsilon|\rho) + J(\rho), \qquad (3.18)$$

where

$$B(\rho) = C_{\rm TF} \rho^{5/3} - C_{\rm D} \rho^{4/3}.$$
 (3.19)

The function space for TFDW theory is the same as TFW theory i.e. \mathcal{G} . However, TFDW is not convex nor weakly lower-semi continuous due to the $-\int \rho^{4/3}$ term in B. This means the analysis of the optimisation problem of TFDW is more tricky than for TF or TFVW.⁵³

However, we would still like to know what the energy predicted using the

TFDW model is for a given potential $v \in Q$. What is this Banach space Q? Firstly we note that if $\nabla \rho^{1/2} \in L_2$ then $\rho^{1/2} \in L_6^{\dagger}$. It is easiest to construct Q by noting that if $\rho^{1/2} \in \mathcal{G}$ then $\rho \in \mathcal{G}'$ defined through

$$\mathcal{G}' = \left\{ \rho | \rho \in L_3 \cap L_{5/3}, \int \rho = N \right\}.$$
(3.20)

Due to the definition of $\varepsilon_{\text{TFDW}}$ then $\mathcal{Q} = (\mathcal{G}')^*$, and so,

$$Q = \{ v | v \in L_{3/2} + L_{5/2} \}.$$
(3.21)

 $\mathcal{Q} \supset \mathcal{X}^*$, and so, \mathcal{Q} contains all Coulomb like potential functions. Then the ground state energy is

$$E_{\text{TFDW}} = \inf_{\rho \in \mathcal{G}} \left\{ F_{\text{TFDW}}(\rho) + (\upsilon | \rho) \right\} \quad \forall \upsilon \in \mathcal{Q}.$$
(3.22)

One can show that ρ minimizes $F_{\text{TFDW}}(\rho) + (\upsilon|\rho)$ for $\int \rho = N$ if and only if $\rho^{1/2}$ minimizes $F_{\text{TFDW}}(\rho^{1/2}) + (\upsilon|\rho)$. This shows that one can treat the square-root of the density as the objective target rather than the density. We will use this fact in the optimisers presented in this thesis where we solve the variational equation by expanding the square root of the density in a finite basis. There is a similar theorem to Theorem 3.2.4.⁵³ The only difference is a technical constant factor introduced to deal with the non-convexity of the TFDW functional. The eigenvalue equations in Theorem 3.2.4 give the justification of the work by Lopez-Acevedo et al.¹ which is introduced and discussed in detail in Section 3.3.2.

3.3. The Non-interacting Kinetic Energy Functional

As stated previously OF-DFT is desirable as it allows for a fully quantum mechanical treatment of systems with millions of atoms. This has already been done for solid state systems.^{67,76} However, for chemical systems it has proven to be very difficult to derive good enough orbital free approximations to the KE functionals (OF-KEFs).

[†] Again we have used Sobolov's inequality

3.3.1. A Review on Functional Development

The Thomas Fermi functional introduced in the previous section is an LDA functional for the KE and is exact for a uniform electron gas. The $\rho^{5/3}$ in the integrand of the functional ensures that the correct coordinate scaling

$$T_{\rm s}(\rho_{\lambda}) = \lambda^2 T_{\rm s}(\rho), \qquad (3.23)$$

where

$$\rho_{\lambda}(\mathbf{r}) = \lambda^3 \rho(\lambda \mathbf{r}), \qquad (3.24)$$

is adhered to. Therefore, $\rho^{5/3}(\mathbf{r})$ is present in all the OF-KEF's in the literature, usually multiplied by a dimensionless enhancement factor. As discussed all LDA OF-KEF's will predict unstable molecular states. The first GGA OF-KEF was $T_{\rm VW}$ which we have already introduced. It is exact for one electron systems and two electron singlet states and is exact in the rapidly varying density limit.⁷⁷ As seen it is also a lower bound to the true KE. On the basis of Teller's non-binding theorem, when developing OF-KEF's, one would want to include $\nabla \rho(\mathbf{r})$, which forms a GGA OF-KEF or in addition $\nabla^2 \rho(\mathbf{r})$, to form Laplacian level OF-KEF's. It has been found that the gradient expansion of the kinetic energy density^{78,79} is the most systematic way of achieving this. One can write the non-interacting kinetic energy as

$$T_{\rm s}(\rho) = \int t(\rho(\mathbf{r})) \,\mathrm{d}\mathbf{r}. \tag{3.25}$$

For densities that vary slowly over space the kinetic energy density, $t(\rho)(\mathbf{r})$, has the asymptotic gradient expansion (to fourth order) of the form derived initially by Jennings et al.⁷⁹ then simplified by Hodges⁸⁰

$$t = t_0 + t_2 + t_4, \tag{3.26}$$

where

$$t_0 = C_{\rm TF} \rho^{5/3}(\mathbf{r}), \tag{3.27}$$

$$t_2 = \frac{1}{72} \frac{|\boldsymbol{\nabla}\rho(\mathbf{r})|^2}{\rho(\mathbf{r})},\tag{3.28}$$

$$t_4 = \frac{(3\pi^2)^{-2/3}}{540}\rho^{1/2}(\mathbf{r}) \left[\left(\frac{\boldsymbol{\nabla}^2 \rho(\mathbf{r})}{\rho(\mathbf{r})}\right)^2 - \frac{9}{8} \frac{\boldsymbol{\nabla}^2 \rho(\mathbf{r})}{\rho(\mathbf{r})} \left(\frac{\boldsymbol{\nabla} \rho(\mathbf{r})}{\rho(\mathbf{r})}\right)^2 + \frac{1}{3} \left(\frac{\boldsymbol{\nabla} \rho(\mathbf{r})}{\rho(\mathbf{r})}\right)^4 \right]$$
(3.29)

It is accepted in the literature that the gradient expansion is exact for the slowly varying density limit.

If one truncates after the second term one gets the second order gradient expansion (SGA). One commonly combines this functional with Dirac exchange. In general we can talk about a class of functionals called the **Thomas-Fermi-Dirac-** λ **-von-Weizsäcker** functional (TFD λ W). Popular choices of λ include: $\lambda = 1$ is the TFDW model introduced in the previous section; $\lambda = 1/9$ is the SGA model just discussed; $\lambda = 1/5$ which was chosen to minimise the error of energies predicted by the TFD λ W functional over a set of molecules and atoms.⁸¹ There are two main classes of approximate T_s forms in the literature; one-point (local) functionals

$$T_{\rm s}(\rho) = \int t_s(\rho(\mathbf{r})) \mathrm{d}\mathbf{r}$$
(3.30)

and two-point (non-local) functionals

$$T_{\rm s}(\rho) = \int f_{1,\rm s}(\rho(\mathbf{r}))\chi(\mathbf{r},\mathbf{r}')f_{2,\rm s}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}',\qquad(3.31)$$

where $f_{1,s}$ and $f_{2,s}$ are weighting functionals and $\chi(\mathbf{r}, \mathbf{r'})$ is a type of linear response function. For example, both the Wang-Teter (WT) OF-KEF^{82,83} and the Wang-Govind-Carter (WGC) OF-KEF^{84,85} are based on the form

$$T_{\rm s}(\rho) = T_{\rm TF}(\rho) + T_{\rm VW}(\rho) + \sum_{\alpha} \lambda_{\alpha} T_{\rm X}(\rho)^{\alpha}, \qquad (3.32)$$

where the correction $T_{\rm X}$ has the form

$$T_{\rm X}(\rho) = C_{\rm TF} \left\langle \rho_{\alpha}(\mathbf{r}) \right| \omega_{\alpha}(\mathbf{r} - \mathbf{r}') \left| \rho_{\alpha}(\mathbf{r}') \right\rangle.$$
(3.33)

Here $\alpha \in \mathcal{R}$ takes on different values for different functionals. X is a dummy label for the functional acronym e.g. WT or WGC. The kernel $\omega_{\alpha}(\mathbf{r} - \mathbf{r'})$ is a Green's function, which is determined by the requirement that the $T_{\rm X}(\rho)$ must be compatible with Lindhard linear response theory.⁸⁶ They are reasonably accurate for very large metallic systems.^{87–90}

There are two issues with this class of functionals which meant they were not studied in this thesis. One was an implementation issue because the Green's function is usually found by solving an equation in reciprocal (momentum) space using fast Fourier transforms. This is suited to solid state applications, where one uses plane wave basis sets, but is less convenient in molecular applications, where one uses Gaussian basis sets. This also means that the cost of these calculations scale as $\mathcal{O}(N_{\rm b} \ln N_{\rm b})$ which is not linear but is better than formal KS-DFT. Secondly non-local functionals of the form in Eqs. (3.32) and (3.33) with a density independent kernel, e.g. WT OF-KEF, were shown to be 'non-linearly unstable' by Blanc and Cancés.⁹¹ This means these OF-KEF's are not bounded from below i.e. these OF-KEF's never predict an energy minima. It is a matter of fact that non-linear stability is a strong indicator of the ability of a particular functional's proficiency at correctly describing the electronic structures of a diverse range of chemical systems. Although it has not been mathematically shown, there is a strong likelihood that non-local functionals using density dependent kernels, e.g. WGC OF-KEF, will also be non-linearly unstable. Due to these issues, this thesis will from now on only be concerned with local (one-point) functionals. What follows is a brief discussion on the different approaches in developing local functionals.

As the literature is vast we will focus on the functionals tested in this work. The exact KEF is homogeneous of degree 2 with respect to coordinate scaling. To show this we can use the scaled wavefunction for an N electron system

$$\psi(\mathbf{r}_i) \mapsto \psi_{\lambda}(\mathbf{r}_i) = \lambda^{\frac{3N}{2}} \psi(\lambda \mathbf{r}_i)$$
(3.34)

where \mathbf{r}_i is the position vector for the *i*-th electron. This is equivalent to Eq. (3.24). This is because

$$\rho_{\lambda}(\mathbf{r}) = \lambda^{3N} \int |\psi(\lambda \mathbf{r}_{1} \cdots \lambda \mathbf{r}_{N})|^{2} \mathrm{d}(\lambda \mathbf{r}_{2}) \cdots \mathrm{d}(\lambda \mathbf{r}_{N}),$$

$$= \lambda^{3} \int |\psi(\lambda \mathbf{r}_{1}, \lambda \mathbf{r}_{2} \cdots \lambda \mathbf{r}_{N})|^{2} \mathrm{d}\mathbf{r}_{2} \cdots \mathrm{d}\mathbf{r}_{N},$$

$$= \lambda^{3} \rho(\lambda \mathbf{r}).$$

(3.35)

Then one can compute the KE using the scaled wavefunction through

$$\begin{split} \langle \psi_{\lambda} | \hat{T}_{N} | \psi_{\lambda} \rangle &= \\ \int \psi^{*} (\lambda \mathbf{r}_{1} \cdots \lambda \mathbf{r}_{N}) \left(-\frac{\lambda^{2}}{2} \sum_{i} \frac{\partial^{2}}{\partial (\lambda \mathbf{r}_{i})^{2}} \right) \psi (\lambda \mathbf{r}_{1} \cdots \lambda \mathbf{r}_{N}) \mathrm{d}(\lambda \mathbf{r}_{1}) \cdots \mathrm{d}(\lambda \mathbf{r}_{N}) \\ &= \lambda^{2} \int \psi^{*} (\mathbf{r}_{1} \cdots \mathbf{r}_{N}) \left(-\frac{1}{2} \sum_{i} \frac{\partial^{2}}{\partial \mathbf{r}_{i}^{2}} \right) \psi (\mathbf{r}_{1} \cdots \mathbf{r}_{N}) \mathrm{d}\mathbf{r}_{1} \cdots \mathrm{d}\mathbf{r}_{N}, \\ &= \lambda^{2} \left\langle \psi | \hat{T}_{N} | \psi \right\rangle. \end{split}$$
(3.36)

Therefore, the exact KEF in DFT should have a homogeneity degree of

M. S. Ryley Variational Solutions in OF-DFT

69

two with respect to co-ordinate scaling. However, the full interacting case is not homogeneous of any degree, as was shown by Levy and Perdew. 49

Therefore, most GGA functionals are of the form

$$T_{\rm s}(\rho) = C_{\rm TF} \int \rho^{5/3} H(s) \mathrm{d}\mathbf{r}$$
(3.37)

where s is the dimensionless reduced gradient and is defined through

$$s(\rho) = \frac{1}{2(3\pi)^{1/3}} \frac{|\nabla \rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}$$
(3.38)

and is dimensionless with respect to co-ordinate scaling such that Eq. (3.37) automatically satisfies the quadratic response to this scaling. H(s) is known as the enhancement factor. One very popular method to construct this is using the conjointness hypothesis. Put simply, it postulates that enhancement factors used in an exchange functional can be used (if slightly modified by optimisation of parameters) as the enhancement factor for an approximate $T_{\rm s}$ functional. This hypothesis has been proved not to be fruitful in developing accurate OF-KEFs though it can be a suggestive template to explore future possible functional forms. Examples include: conj-B86A⁹²

$$H^{\text{B86A}}(s) = 1 + 0.00387 \frac{(bs)^2}{1 + 0.004(bs)^2},$$
(3.39)

with $b = 2(6\pi^2)^{1/3}$; conj-B86B⁹²

$$H^{\text{B86A}}(s) = 1 + 0.00403 \frac{(bs)^2}{(1 + 0.007(bs)^2)^{4/5}};$$
(3.40)

conj-PW91⁹²

$$H^{\text{PW91}}(s) = \frac{1 + 0.1946s \operatorname{arcsinh}(7.7956s) + \left(0.2743 - 0.1508e^{-100s^2}\right)s^2}{1 + 0.19645s \operatorname{arcsinh}(7.7956s) + 0.004s^4};$$
(3.41)

conj-PW86⁹²

$$H^{\text{PW86}}(s) \left(1 + 1.296s^2 + 14s^4 + 0.2s^6\right)^{1/15}; \qquad (3.42)$$

 $LC94^{93}$

$$H^{\rm LC94}(s) = \frac{1 + 0.093907s \operatorname{arcsinh}(76.32s) + \left(0.26608 - 0.08096e^{-100s^2}\right)s^2}{1 + 0.093907s \operatorname{arcsinh}(76.32s) + 5.7767 \times 10^{-5}s^4}$$
(3.43)

70

 $TW02^{94}$

$$H^{\text{TW02}}(s) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu}{\kappa}s^2}$$
 (3.44)

where $\kappa = 0.8438$ and $\mu = 0.2319$; T92⁹⁵

$$H^{\text{T92}}(s) = 1 + \frac{0.0055(bs)^2}{1 + 0.0253bs \operatorname{arcsinh}(bs)} - \frac{0.072bs}{1 + 2^{5/3}bs};$$
(3.45)

 $VT84^{96}$

$$H^{\rm VT84}(s) = 1 - \frac{\beta s^2 e^{-\alpha s^2}}{1 + \beta s^2} + (1 - e^{-\alpha s^4})(s^{-2} - 1) + \frac{5s^2}{3}$$
(3.46)

where $\alpha = 1.2965$ and $\beta = 2.778$.

Using the conjointness hypothesis as a template for OF-KEF construction has not yielded kinetic functionals with the same successes as the equivalent exchange functionals. One reason could be that the Dirac functional is a better approximation to the exchange energy than the Thomas-Fermi functional is to the KE. Another could be that, from the virial theorem, because the non-interacting kinetic term is the dominant term in the total energy, any inaccuracies in the enhancement factor become more noticeable when used in calculating the KE.

The Pauli potential is defined as the exact non-interacting kinetic energy potential minus the von Weizsäcker potential. An alternative explanation for the errors using the above functionals was given by Karasiev et al.⁹⁷ They noted that the exact Pauli potential is non-negative over all space and then showed that for the functionals they analysed based on the conjointness hypothesis their associated Pauli potential is negative at some points in space. This is not a problem for exchange functionals (there is no corresponding Pauli potential requirement), but this violates an important constraint for the exact $T_{\rm s}$. Averill and Painter,⁹⁸ derived an expression which explained how one could use $T_{\rm s}$, to calculate the forces in a molecule using the virial theorem and the Hellmann-Feynman theorem. Karasiev et al.^{97,99} used this as a starting point and in using the requirement of the positivity of the Pauli term to formulate the PBE-n functionals,

$$H_{\text{PBEn}}(s) = \left[1 + \sum_{i=1}^{n-1} C_i \left[\frac{s^2}{1+as^2}\right]^i\right]$$
(3.47)

 C_i and *a* are parameters which were found optimising the error in potential energy curves and are displayed in Table 3.1.

As well as considering coordinate scaling one can also consider the density

Table $3.1.$:	Parameters	for the	PBEn	OF-NIKEFs
T (0)10 0110	I GIGINIOUOID	TOT OTTO	1 2 2 11	OT THILDED

	a	C_1	C_2	C_3
PBE2	0.2942	2.0309	0	0
PBE3	4.1335	-3.7425	50.258	0
PBE4	1.7107	-7.2333	61.645	-93.683

scaling relationship

$$\rho_{\zeta}(\mathbf{r}) = \zeta \rho(\mathbf{r}). \tag{3.48}$$

It is important to note that density scaling, unlike its coordinate scaling counterpart, does not maintain the normalization of the electronic density to N (number of electrons). An important concept is homogeneity which can be applied to all types of functional scaling. A functional $L(\rho)$ is said to be homogeneous of degree k in density scaling if it satisfies

$$L(\rho_{\zeta}) = \zeta^k L(\rho). \tag{3.49}$$

For $k \neq 0$, the degree of homogeneity with respect to density scaling is given by

$$k = \frac{\int \rho(\mathbf{r}) \frac{\delta L}{\delta \rho(\mathbf{r})} \mathrm{d}\mathbf{r}}{L(\rho)}.$$
(3.50)

One can also say that if a functional $L(\rho)$ is homogeneous of degree m under coordinate scaling, then the degree of homogeneity in coordinate scaling is given by

$$m = -\frac{\int \rho(\mathbf{r})\mathbf{r} \cdot \boldsymbol{\nabla} \frac{\delta L}{\delta \rho(\mathbf{r})} \mathrm{d}\mathbf{r}}{L(\rho)}.$$
(3.51)

If an LDA functional is homogeneous of degree m under coordinate scaling then it is also homogeneous of degree $k = \frac{m+3}{3}$ under density scaling. However, for GGA functionals this relationship breaks down, and so, one can use density scaling as another additional constraint. Unlike for coordinate scaling, there seems to be no unique k that satisfies Eq. (3.50). In their initial work Liu and Parr¹⁰⁰ concluded that the KE is homogeneous of degree one under density scaling (k = 1). Their proof was shown to be incorrect by Gal,¹⁰¹ furthermore, Chan and Handy¹⁰² also showed that a functional that is homogeneous of degree one under density scaling does not satisfy the Lieb-Thirring bound¹⁰³

$$T_{\rm s}(\rho) \ge \frac{3}{5} \left(\frac{3\pi}{4}\right)^{2/3} \int \rho^{5/3}(\mathbf{r}) \mathrm{d}\mathbf{r}.$$
 (3.52)

They also established that the KE is inhomogeneous (system dependent)

under density scaling. This was further highlighted by Borgoo, Teale and Tozer¹⁰⁴ who concluded that the functional was not homogeneous of any degree but that the effective homogeneity, $k_{\rm eff}$, associated with the potential that averages over the integer discontinuity — which is most appropriate for a continuum functional such as a GGA — did not exhibit significant system dependence for systems with more than a few electrons. The integer discontinuity in question refers to the plot of the exact total electronic energy as a function of electron number which comprises of a series of straight line segments, with derivative discontinuities at integer number of electrons.¹⁰⁵

Borgoo and Tozer also built upon the results from previous research, 104,106,107 by defining an average value of $k_{\rm eff}$, $k_{\rm eff,av}$, for a set of closed shell atoms and molecules. This average value was used to determine

$$q_{\rm av} = 5 - 3k_{\rm eff,av} = 0.3434125, \tag{3.53}$$

and defined the enhancement factor through

$$H_{\rm BT2}(x) = \beta x^{q_{\rm av}}(\rho) \mathrm{d}\mathbf{r} \tag{3.54}$$

where x is yet another dimensionless quantity defined as

$$x(\rho) = \frac{|\boldsymbol{\nabla}\rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})},\tag{3.55}$$

and β was optimised to 1.990328 from the zero intercept linear regression between the quantity $T_{\rm s}^{BT2}/\beta$ and the Kohn-Sham $T_{\rm s}$ for a set of closed shell molecules. They also went on to define the following functional

$$H_{\rm BT3}(x) = \gamma x^{q_{\rm opt}}(\rho) \mathrm{d}\mathbf{r} \tag{3.56}$$

where $q_{\rm opt} = 0.2$ was computed by considering an intermediate exponent to yield near optimal energies for their training set of molecules. γ was optimised to 2.397888 using an analogous procedure to the optimisation of β . It was found that BT-2 and BT-3 gave good total energies for their training set and it was shown that for a selection of molecules they even predicted binding.

Levy and Ou-Yang considered how the exact exact KEF behaves under non-uniform co-ordinate scaling.¹⁰⁸ Without going into the details they showed that the gradient expansion SGA breaks this requirement, and so, they proposed the functionals 109 OL1

$$H^{\text{OL1}}(s) = 1 + \frac{5}{27}s^2 + 0.00677\frac{20}{3}(3\pi^2)^{-1/3}s \tag{3.57}$$

and OL2

$$H^{\text{OL1}}(s) = 1 + \frac{5}{27}s^2 + \frac{0.0887}{C_{\text{TF}}} \frac{2(3\pi^2)^{1/3}s}{1 + 8(3\pi^2)^{1/3}s}.$$
 (3.58)

It is well known that in the slowly varying density limit $(s^2 \rightarrow 0)$ the enhancement factor should become 1 i.e. TF, and in the rapidly varying density limit $(s^2 \rightarrow \infty)$ the enhancement factor should become $1 + 5s^2/27$ i.e. SGA. The easiest way to enforce an enhancement factor satisifies both these limits is through the use of a Padé approximant. A Padé approximant is just a ratio of two polynomials and is the 'best' approximation to analytic function fixed at two limits. Examples of OF-KEFs derived using a Padé approximant are: P92¹¹⁰

$$H^{\rm P92}(s) = \frac{1 + 88.396s^2 + 16.3683s^4}{1 + 88.2108s^2}; \tag{3.59}$$

 DK^{111}

$$H^{\rm DK}(x) = \frac{1 + 0.95x + 14.28111x^2 - 19.57962x^3 + 26.64765x^4}{1 - 0.05x + 9.99802x^2 + 2.96085x^3}$$
(3.60)

with $x = 5s^2/27$; E00¹¹²

$$H^{\text{E00}}(s) = \frac{135 + 28s^2 + 5s^4}{135 + 3s^2}.$$
(3.61)

All the functionals discussed in this section where implemented as part of this thesis into XCFun,¹¹³ which is a program that computes the partial derivatives of the functional with respect to the density, the norm of the density gradients and the Laplacian of the density. These were then incorporated into the construction of the required matrix elements for selfconsistent application of OF-DFT.

3.3.2. A Review on Solving the Optimisation Problem

The accuracy of the OF-KEFs introduced in the last section was assessed using post Kohn Sham analysis (PKSA). In PKSA one calculates oneshot KEs using an OF-KEF evaluated on a KS density solution. The problem with this approach is that in self-consistent OF-DFT calculations one will be solving the optimisation problems discussed previously where the $F_{\rm L}$ is an approximate form depending on the choice of OF-KEF and XC functional. PKSA gives us no indication whether the functional is well behaved or whether the optimisation problem is well defined i.e. does the approximate $F_{\rm L}$ have a minimiser?

For solid state systems self-consistent OF-DFT codes are available such as: ATLAS,¹¹⁴ which is a real space finite-difference method for the numerical solution of OF-DFT in periodic systems; and PROFESS,⁸² which is a quasi-linear scaling code and uses local pseudo-potentials. PROFESS also relies on solving for the Coulombic energy and the KE in momentum (reciprocal) space through the use of fast Fourier transforms and periodic boundary conditions. In addition both PROFESS and ATLAS use plane wave basis sets, not Gaussian basis sets. For a self-consistent OF-DFT code which can aid the development OF-KEFs for chemical systems one prefers an all electron treatment. This is because the core regions of the density are essential for accurate KE calculations, and so, it is important to treat these electrons explicitly when testing approximate KEFs.

As noted in Section 3.2 the TF λ W models can be approached as a nonlinear eigenvalue problem. This was first realised by Sahni et al.¹¹⁵ who formulated the problem, for the TFDW model, in the following manner

$$h_{\rm eff}(\mathbf{r})\rho^{1/2}(\mathbf{r}) = \mu\rho^{1/2}(\mathbf{r}),$$
 (3.62)

where

$$h_{\rm eff}(\mathbf{r}) = -\frac{1}{2}\boldsymbol{\nabla}^2 + \upsilon_{\rm eff}(\mathbf{r})$$
(3.63)

and

$$\upsilon_{\text{eff}}(\mathbf{r}) = \upsilon(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} + \frac{\delta T_{\text{s}}(\rho)}{\delta\rho(\mathbf{r})} - \frac{\delta T_{\text{VW}}(\rho)}{\delta\rho(\mathbf{r})} + \frac{\delta E_{\text{XC}}(\rho)}{\delta\rho(\mathbf{r})}.$$
 (3.64)

 μ is known as the chemical potential. The chemical potential can be seen as minus the first ionisation energy of the system if the potential goes to zero as $\mathbf{r} \to \infty$. It was claimed this would be easy to implement by modifying existing KS-DFT codes, however, it has been extremely difficult to do so in reality. Lopez-Acevedo et al.¹ made progress using this method by using the TFD λ W model to derive the following eigenvalue equation (referred from now on as the Lopez-Acevedo scheme)

$$\left(-\frac{1}{2}\boldsymbol{\nabla}^2 + \frac{1}{\lambda}v'_{\text{eff}}(\mathbf{r})\right)\rho^{1/2} = \frac{\mu}{\lambda}\rho^{1/2}(\mathbf{r}),\tag{3.65}$$



Figure 3.1.: A flow chart outlining the Lopez-Acevedo scheme

where

$$\upsilon_{\text{eff}}'(\mathbf{r}) = \upsilon(\mathbf{r}) + \int \frac{\rho(\mathbf{r})'}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} + \frac{\delta T_{\text{TF}}(\rho)}{\delta\rho(\mathbf{r})} - \frac{\delta K_D(\rho)}{\delta\rho(\mathbf{r})}.$$
 (3.66)

As part of this thesis I have implemented the Lopez-Acevedo scheme in to QUEST¹¹⁶* using a similar algorithm to a traditional KS-DFT calculation for a one orbital system. For a KS-DFT calculation (and other quantum chemical models which rely on self-consistent methods) to reach

^{*} A Python based code which is a rapid development platform for electronic structure methods in quantum chemistry

self-consistent (Self Consistent (SC)) convergence, one uses convergence accelerators. A common tactic is to construct density matrices to be used in the current iteration using density matrices from previous iterations. For example, Pulay's DIIS^{22,23} is a very common SC acceleration technique using this approach. The set of MO coefficients $\{c_i\}$, is chosen to minimise the commutator [F, D], where F is the Fock matrix and D the density matrix. This is because it is a necessary and sufficient condition that this commutator becomes zero for a SC solution. However, it has been shown to correspond to a projected quasi-Newton method,^{117,118} which means it requires the total energy to be a quadratic function with respect to the MO coefficients in a neighbourhood of the minimiser. When we tried to apply this to the Lopez-Acevedo scheme, DIIS failed because the total energy defined by the universal functional $F_{\text{TFD}\lambda\text{W}}$ is non-convex as discussed previously and varies rapidly with D. This means $F_{\text{TFD}\lambda\text{W}}$ cannot meet the quadratic requirement of DIIS.

We, therefore, had to use a simple mixing scheme for our convergence accelerator. We used this to form the Coulomb, exchange and KE matrices in the current iteration from the corresponding matrices of previous iterations. For example, the Coulomb matrix for the (n + 1)th iteration is formed by

$$\left\langle \eta_i | J | \eta_j \right\rangle_{n+1} = \alpha \left\langle \eta_i | J | \eta_j \right\rangle_n + (1 - \alpha) \left\langle \eta_i | J | \eta_j \right\rangle_{n-1}, \quad (3.67)$$

where n denotes the iteration number,

$$\langle \eta_i | J | \eta_j \rangle = 2 \sum_{\kappa \nu} \sum_{i}^{\text{occ}} c_{\kappa i} c_{\nu i} (ij | \kappa \nu), \qquad (3.68)$$

where η_i denotes basis functions and the damping factor $\alpha = 0.001$. Due to the small magnitude of α we found that there was a large number of iterations, thousands, needed before convergence was achieved. However, if one sets α much higher then this prevents convergence. Furthermore, the damping required changed with the choice of functional and atom type.

Fig. 2.1 is a flowchart of the Lopez-Acevedo scheme which I have implemented into QUEST. The initial guess of MO coefficients $\{c_i\}$ come from the eigenvector corresponding to the lowest energy eigenvalue of the core Hamiltonian. The 'Fock' matrix is given by

$$F_{ij} = \langle \eta_i | -\frac{1}{2} \nabla^2 | \eta_j \rangle + \frac{1}{\lambda} \langle \eta_i | J + \upsilon | \eta_j \rangle + \frac{1}{\lambda} \langle \eta_i | \frac{\delta T_{\text{TF}}}{\delta \rho} + \frac{\delta E_{\text{XC}}}{\delta \rho} | \eta_j \rangle \quad (3.69)$$

where $E_{\rm XC}$ is the Dirac functional. This is just $-K_D$ which is defined in Equation (3.17). Due to the transformations between the AO and MO basis it is clear that our implementation is not linear scaling. However, in principle, the same methods used to achieve linear scaling KS-DFT codes could be applied here. We found that the number of iterations increased greatly when one increased the number of electrons in the system. The tolerance criteria is 10^{-6} Hartree's on the total energy, 10^{-8} a.u. on the chemical potential and 10^{-6} on the norm of the density matrix. This was implemented for both spin polarised and unpolarised formalisms.

The key issue we found with this scheme is that it can only converge the energy for a very small subset of OF-KEFs. We could only converge for calculations involving TFD λ W functionals for $\lambda \neq 0$. Indeed, more recent work using this scheme has focussed on α TFD λ W functionals,¹¹⁹ where the TF OF-KEF is scaled by $\alpha \in \mathcal{R}$. In this most recent work the projector augmented wave method was used to generate molecular densities from the atomic densities. Therefore, the large increase in iteration count was less of a problem than in the way we implemented this scheme. However, due to the restricted range of functionals one can use in this scheme it is not appropriate for general single point OF-KEF development.

A self-consistent scheme, which directly minimises the energy functional without resorting to solving an eigenvalue problem was introduced by Chan, Cohen and Handy (CCH).² This was done with the goal to investigate the self consistent properties of the TFD λ W models. Their method requires a series of minimisations of the following Lagrangian

$$\mathcal{L}(\rho;\mu) = E(\rho) - \mu N_{\mu} \tag{3.70}$$

where μ takes the role of the Lagrange multiplier and each minimisation yields a density normalised to N_{μ} . The minimisation with respect to the density yields the Euler-Lagrange equation

$$\frac{\delta E}{\delta \rho} = \mu. \tag{3.71}$$

Between each minimisation we conduct a one-dimensional root search to find a new guess of μ such that $|N_{\mu} - N_{\text{elec}}|$ decreases, where N_{elec} is the correct normalisation of the density. We will discuss this method in detail in Chapter 4.

3.4. Solving the Optimisation Problem: A Formal Approach

As we will see in Chapter 4 the CCH approach is slow to converge due to the need to solve a series of optimisation problems. To arrive at a more efficient optimiser we need to discuss DFT but for flexible, not fixed, particle number.

Formally we have to solve the variational problem for a given orbital free approximation to $F_{\rm L}, F_{\rm OF} : \mathcal{P} \to \operatorname{cl}(\mathcal{R})$, given through

$$E(\upsilon) = \inf_{\rho \in \mathcal{P}} \left[F_{\text{OF}}(\rho) + (\upsilon|\rho) \right] \quad \forall \upsilon \in \mathcal{P}^*.$$
(3.72)

Where we require that \mathcal{P} at least contains \mathcal{X} and \mathcal{P}^* contains the Coulomb potentials i.e. contains \mathcal{X}^* . In all our formal development so far we have kept the normalisation constraint on the density fixed. This means that if one tried to self-consistently solve the optimisation problems it is not possible to vary the number of electrons described by the density. The only way to change the number of electrons would be outside of the optimisation, Equation (3.72). We will see this is the case for CCH in Chapter 4. We will find that we wish to minimise the energy and vary the density normalisation *simultaneously*. By this we mean we wish to only complete one functional optimisation which ends at the correctly normalised ground state density. This means we need to extend our formal approach we have developed so far for fixed N-particle systems to grand canonical ensembles, encompassing systems of varying N.

3.4.1. Grand Canonical Ensembles in OF-DFT

The goal in this section is to show a formulation of DFT which allows for arbitrary variation of the particle number.¹³ Up to now our variational principles have been defined for a fixed number of electrons. The reason for this is that we have been formulating DFT with reference to a version of quantum mechanics formulated using a single Hilbert space to describe a system. It is well known that this 'first-quantised' version of quantum mechanics cannot treat any system where particle number is not conserved. The improved formulation of quantum mechanics is quantum field theory which allows for the description of systems with variable particle number. It achieves this by having Fock spaces as its underlying mathematical structure rather than a single Hilbert space. Put simply, a Fock space is the sum of a set of Hilbert spaces describing zero particles, one particle, two particles and so forth.

We will be following the formalism outlined by Helgaker et al.¹³ For a system in an external potential $v \in \mathcal{X}^*$ and particle number $N \in \mathcal{N}_+^{\dagger}$ the N-electron ground state energy $\varepsilon_0 : \mathcal{X}^* \times \mathcal{N}_+ \to \operatorname{cl}(\mathcal{R})$ is

$$\varepsilon_0(\upsilon, N) = \inf_{\hat{\gamma} \in \mathcal{D}_N} \operatorname{tr} \hat{\gamma} \hat{H}(\upsilon)$$
(3.73)

where the hat signifies that the operators act on the Fock space not the Hilbert space. Indeed \mathcal{D}_N is defined using a new definition of \mathcal{W}_N which is now the set of admissible *N*-electron Fock space states. We can now state the following theorem

Theorem 3.4.1. The grand canonical ensemble (Grand Canonical Ensemble (GCE)) ground-state energy $\varepsilon_0 : \mathcal{X}^* \times \mathcal{R} \to cl(\mathcal{R})$ is defined through

$$\varepsilon_{0}(v,N) = \begin{cases} +\infty, & N \in (-\infty,0).\\ \inf_{\hat{\gamma} \in \mathcal{D}_{N}} \operatorname{tr} \hat{\gamma} \hat{H}(v), & N \in \mathcal{N}_{+}.\\ (1 - \langle N \rangle) \varepsilon_{0}(v, \lfloor N \rfloor) + \langle N \rangle \varepsilon_{0}(v, \lceil N \rceil), & N \in (0, +\infty)/\mathcal{N}_{+}. \end{cases}$$

The conditions for $\hat{\gamma}_{\rho} \mapsto \rho$ are

$$\varepsilon_0(\upsilon, N) = \operatorname{tr} \hat{\gamma}_{\rho} \hat{H}(\upsilon) \Longleftrightarrow \rho \in \bar{\partial}_1 \varepsilon_0(\upsilon, N), \forall N \in \mathcal{N}_+$$

The map $v \mapsto \varepsilon_0(v, N)$ belongs to $\overline{\Gamma}^*(\mathcal{X}^*)$ and increases monotonically. The map $N \mapsto \varepsilon_0(v, N)$ is lower semi-continuous on \mathcal{R} and continuous on $(0, +\infty)$.

See Section 1.3.4 for the definition of $\bar{\partial}$. Also $\bar{\Gamma}^*(\mathcal{X}^*)$ denotes the set of all weak-star upper semi continuous concave functionals operating on \mathcal{X}^* . See Chapter 1 for a review of the meaning of these terms.

One can also define the GCE density matrix, $\hat{\Gamma}$, as a convex combination of ensemble density matrices of different particle numbers. The set of admissible *GCE density matrices*, \mathcal{D}_+ , is just the convex hull of the union of all sets of admissible canonical density matrices. It should be clear why \mathcal{D}_+ is the density matrix representation of the Fock space, and so, can describe any value of particle number. We can now define the *GCE ground-state energy* $\mathcal{G}_0 : \mathcal{X}^* \to cl(\mathcal{R})$ as being the greatest lower bound to all the GCE expectation values of a Hamiltonian for a particular $v \in \mathcal{X}^*$,

[†] where \mathcal{N}_+ is the set of all positive integers and zero

or more explicitly

$$\mathcal{G}_0(\upsilon) = \inf_{\hat{\Gamma} \in \mathcal{D}_N} \operatorname{tr} \hat{\Gamma} \hat{H}(\upsilon)$$
(3.74)

which tells us that the GCE ground state energy can be found by a convex optimisation over $\{\varepsilon_0(v, N)\}_{N=0}^{\infty}$.

Now we can explain DFT formalised using GCE's. We know that $\mathcal{G}_0 \in \overline{\Gamma}^*(\mathcal{X}^*)$, and so, we can introduce the **GCE Lieb functional**, $\mathcal{F}_{\mathrm{L}} \in \Gamma_0(\mathcal{X})$ as a Legendre-Fenchel conjugate to \mathcal{G}_0 , and so,

Theorem 3.4.2. The GCE energy \mathcal{G}_0 is related to \mathcal{F}_L in the following Hohenberg-Kohn variation principle:

$$\mathcal{G}_0(\upsilon) = \inf_{\rho \in \mathcal{X}} \left(\mathcal{F}_L(\rho) + (\upsilon|\rho) \right)$$

It can be shown that \mathcal{F}_{L} is non-negative and is the Γ -regularisation and lower bound to all admissible density functionals. See Section 1.4 for a review of the Γ_0 sets. What we are saying is that \mathcal{F}_L a weak-star lower semi-continuous function on \mathcal{X} . Its effective domain is the set \mathcal{I}_+ which is defined through

$$\mathcal{I}_{+} = \left\{ \rho | \rho \in \mathcal{R}_{+}, \left\langle \rho^{1/2} \middle| \rho^{1/2} \right\rangle \leq +\infty, \left\langle \boldsymbol{\nabla} \rho^{1/2} \middle| \boldsymbol{\nabla} \rho^{1/2} \right\rangle \leq +\infty \right\}$$
(3.75)

where \mathcal{R}_+ is a generalisation of the set \mathcal{R}_N , defined in Eq. (2.60), where we lift the restriction $\int \rho(\mathbf{r}) d\mathbf{r} = N$. Since all admissible functionals are bounded below by the von-Weizsäcker functional, which is nowhere locally bounded above, we have that

Theorem 3.4.3. An admissible GCE density functional (including the Lieb functional) is nowhere locally bounded above, nowhere continuous, nowhere differentiable

In OF-DFT one is trying to find approximations to the GCE density functional, and so, must have the above properties to be considered admissible from a purely theoretical perspective.

3.4.2. The Importance of Saddle Functions in OF-DFT

We can control the number of electrons in the system by shifting the potential using the chemical potential, μ , which was introduced previously when discussing the Lopez-Acevedo scheme. However, this is fairly indirect so we will concentrate on the more direct approach to controlling the particle number — which is what we require in our numerical optimisations. We can define the GCE energy bifunctional, $\mathcal{G} : \mathcal{X}^* \times \mathcal{R} \to \operatorname{cl}(\mathcal{R})$ and the GCE universal density bifunctional $\mathcal{F} : \mathcal{X} \times \mathcal{R} \to \operatorname{cl}(\mathcal{R})$ as

$$\mathcal{G}(\upsilon,\mu) = \mathcal{G}(\upsilon-\mu,0) = \mathcal{G}_0(\upsilon-\mu), \quad \forall (\upsilon,\mu) \in \mathcal{X}^* \times \mathcal{R},$$
(3.76)

$$\mathcal{F}(\rho, N) = \mathcal{F}_{\mathrm{L}}(\rho), \int \rho = N \quad \forall (\rho, N) \in \mathcal{X} \times \mathcal{R}$$

$$= +\infty, \int \rho \neq N \quad \forall (\rho, N) \in \mathcal{X} \times \mathcal{R}$$
(3.77)

The bifunctionals are closed concave and closed convex functions respectively i.e. $\mathcal{G} \in \overline{\Gamma}^*(\mathcal{X}^* \times \mathcal{R})$ and $\mathcal{F} \in \overline{\Gamma}(\mathcal{X} \times \mathcal{R})$. They are related through

$$\mathcal{G}(\upsilon,\mu) = \inf_{\rho \in \mathcal{X}} \inf_{N \in \mathcal{R}} \left(\mathcal{F}(\rho,N) + (\upsilon|\rho) - \mu N \right), \quad \forall (\upsilon,\mu) \in \mathcal{X}^* \times \mathcal{R}.$$
(3.78)

One can show that the bifunctionals are the parents of an equivalence class of closed concave-convex functions $[\varepsilon] : \mathcal{X}^* \times \mathcal{R} \to \mathrm{cl}(\mathcal{R})$ and parents of an equivalence class of closed convex-concave $[\mathcal{K}] : \mathcal{X} \times \mathcal{R} \to \mathrm{cl}(\mathcal{R})$. It turns out that ¹³ these equivalence classes are singletons consisting of a closed concave-convex energy saddle function ε and a closed convex-concave density saddle function \mathcal{K} , respectively.

One can then set-up a four way correspondence¹³ between GCE energy and density functionals, much in the same way as one does in thermodynamics where we relate the internal energy, enthalpy, Helmotz free energy and Gibbs free energy as Legendre transforms. Indeed Nalewajski and Parr have derived the 'Maxwell relations' for DFT.¹²⁰ The four functionals introduced in this section are related as follows, see Theorem 3.4.4 for the definition of \mathcal{H} ,

$$\mathcal{F}(\rho, N) = \sup_{\upsilon \in \mathcal{X}^*} \left(\varepsilon(\upsilon, N) - (\upsilon|\rho) \right) = \sup_{\mu \in \mathcal{R}} \left(\mathcal{H}(\rho, \mu) + \mu N \right), \tag{3.79}$$

$$\varepsilon(v,N) = \inf_{\rho \in \mathcal{X}} \left(\mathcal{F}(\rho,N) - (v|\rho) \right) = \sup_{\mu \in \mathcal{R}} \left(\mathcal{G}(v,\mu) + \mu N \right), \tag{3.80}$$

$$\mathcal{H}(\rho,\mu) = \inf_{N \in \mathcal{R}} \left(\mathcal{F}(\rho,N) - \mu N \right) = \sup_{\upsilon \in \mathcal{X}^*} \left(\mathcal{G}(\upsilon,\mu) - (\upsilon|\rho) \right), \tag{3.81}$$

$$\mathcal{G}(\upsilon,\mu) = \inf_{\rho \in \mathcal{X}} \left(\mathcal{H}(\rho,\mu) + (\upsilon|\rho) \right) = \inf_{N \in \mathcal{R}} \left(\varepsilon(\upsilon,N) - \mu N \right).$$
(3.82)

We are most interested in the energy saddle function $\varepsilon : \mathcal{X}^* \times \mathcal{R} \to cl(\mathcal{R})$ which can be defined through

$$\varepsilon(\upsilon, N) = \inf_{\rho \in \mathcal{X}} \sup_{\mu \in \mathcal{R}} \left(\mathcal{H}(\rho, \mu) + (\upsilon|\rho) + \mu N \right) = \sup_{\mu \in \mathcal{R}} \inf_{\rho \in \mathcal{X}} \left(\mathcal{H}(\rho, \mu) + (\upsilon|\rho) + \mu N \right)$$
(3.83)

Or in other words, $\varepsilon(v, N)$ is the saddle value of the closed convex-concave function $(\rho, \mu) \mapsto \mathcal{H}(\rho, \mu) + (v|\rho) + \mu N$. For a given $(v, N) \in \mathcal{X}^* \times \mathcal{R}$ a saddle point may not exist in the Hohenberg-Kohn variation principle but a finite saddle value always exists on the effective domain of ε . One can further show that

Theorem 3.4.4. The closed convex-concave saddle function $\mathcal{H} : \mathcal{X} \times \mathcal{R} \rightarrow cl(\mathcal{R})$ is given by

$$\mathcal{H}(\rho,\mu) = \mathcal{F}_L(\rho) - (\mu|\rho) \tag{3.84}$$

Therefore, $\mathcal{H}(\rho, \mu)$ is lower semi-continuous in ρ for fixed μ and is continuous affine in μ for fixed ρ .

We are now in a position to state the constrained and unconstrained grand-canonical Hohenberg-Kohn variation principles

Theorem 3.4.5. For each $(v, N) \in \mathcal{X}^* \times \mathcal{R}$, the GCE ground-state energy may be obtained from the constrained Hohenberg–Kohn variation principle

$$\varepsilon(\upsilon, N) = \inf_{\rho \in \mathcal{I}_N} \left(\mathcal{F}_L(\rho) + (\upsilon|\rho) \right)$$
(3.85)

or from the unconstrained Hohenberg-Kohn (minimax) variation principle:

$$\varepsilon(\upsilon, N) = \sup_{\mu \in \mathcal{R}} \inf_{\rho \in \mathcal{X}} \left(\mathcal{F}_L(\rho) + (\upsilon - \mu | \rho) + \mu N \right)$$

$$= \inf_{\rho \in \mathcal{X}} \sup_{\mu \in \mathcal{R}} \left(\mathcal{F}_L(\rho) + (\upsilon - \mu | \rho) + \mu N \right)$$
(3.86)

The minimax problem corresponds to Lagrange's method for the constrained problem, with multiplier μ .

Proof. The variation principle in Eq. (3.85) can be arrived from the fourway correspondence and by firstly noting that $\mathcal{F}(\rho, N) = \mathcal{F}_{L}(\rho)$ on \mathcal{X}_{N} and $+\infty$ elsewhere. Then we note that dom $(\mathcal{F}_{L}) = \mathcal{I}_{+}$ giving $\mathcal{I}_{N} = \mathcal{I}_{+} \cap \mathcal{X}_{N}$. The mini-max variation principle, Eq. (3.86), is a trivial proof by definition. In order to prove why the minimax problem arises from the Lagrange method we write Eq. (3.85) as

$$\varepsilon(v, N) = \inf_{\rho \mathcal{I}_N} \left\{ \mathcal{F}_{\mathcal{L}}(\rho) + (v|\rho) | N = N_{\rho} \right\}$$
(3.87)

where $N_p = (1|\rho)$. So a minimising density ρ_0 exists if and only if there exists $\mu_0 \in \mathcal{R}$ such that (ρ_0, μ_0) is a saddle point for the Lagrange function

$$\mathcal{L}(\rho,\mu) = \mathcal{F}_{\mathrm{L}}(\rho) + (\upsilon|\rho) + \mu(N - N_{\rho})$$
(3.88)

Thus Eq. (3.86) is a minimax problem for the saddle value of the Lagrange function $\mathcal{L}(\rho, \mu)$.

In the CCH scheme μ is fixed and \mathcal{L} is optimised with respect to ρ . A new value of μ is chosen, we will explain how later, such that $(N - N_{\rho})$ is minimised, and so, their final optimisation is just Eq. (3.85). In the original work CCH² claimed that incorporating the particle number constraint directly into the energy minimisation exacerbates the problems found in optimisation procedures. In Chapter 6 we discuss how we achieved this by utilising the saddle function $\mathcal{L}(\rho, \mu)$. We will present how numerical optimisations realising this fact are superior to the CCH scheme in terms of computational cost.

We have now finished laying the theoretical foundation upon which the rest of this thesis is built. In Chapter 4 we will outline how we improved upon the CCH scheme by means of allowing the testing of a wide range of possible \mathcal{F}_{L} approximations to be done simply and quickly. In Chapter 5 we will analyse the potentials of TFD λ W functionals. In Chapter 6 we will show how one finds the saddle point of a surface defined via an approximate \mathcal{F}_{L} and why this is superior, in terms of computational cost, to the CCH scheme. In Chapter 7 we focus on the variation principles at the interface between DFT and wave-function methodologies.

4. The Importance of Self Consistency in Orbital Free Density Functional Theory

4.1. Introduction

In this chapter we describe in detail the Chan, Cohen and Handy scheme $(CCH)^2$ introduced at the end of Section 3.3.2. We also outline how we expanded the number of functionals that can be used in this scheme. We begin by reviewing the Lagrange optimisation method in DFT as applied to OF-DFT in Section 4.1.1. We will then, in Section 4.2.1, move on to discuss in some detail how this work was implemented into QUEST¹¹⁶ with a detailed outline of the equations of the gradient and Hessian required for this optimisation. In Section 4.4 we present some results which clearly demonstrate that one should be cautious using post Kohn Sham analysis to determine the accuracy of kinetic energy functionals (KEFs). We will find that the errors in the converged energies using KEFs are not comparable with previous analysis using Kohn-Sham densities. We will also discuss the errors in the ground state densities predicted by these KEFs.

4.1.1. Recap of Variation Principles

OF-DFT strictly adheres to Hohenberg and Kohn's first theorem. In OF-DFT the electronic energy is written as a functional of the ground state density ρ . The energy is partitioned in the same way as in KS-DFT,

$$E(\rho; \upsilon) = T_{\rm s}(\rho) + E_{\rm J}(\rho) + E_{\rm XC}(\rho) + E_{\rm ext}(\rho; \upsilon).$$

$$(4.1)$$

The notation $E(\rho; v)$ means the energy is defined for a fixed external potential, v.

The functionals in Equation (4.1) whose explicit analytical form using ρ are unknown are: $T_{\rm s}$ — the non-interacting kinetic energy functional

CHAPTER 4. THE IMPORTANCE OF SELF CONSISTENCY IN ORBITAL FREE DENSITY FUNCTIONAL THEORY

(KEF); and $E_{\rm XC}$ — the exchange-correlation energy. This differs to KS-DFT where the KEF is expressed using orbitals, not ρ . The advantage of using just ρ rather than relying on the re-introduction of orbitals is that one only has to solve one equation in OF-DFT to find the ground state density. This is in contrast to KS-DFT where one has to solve a set of equations. As there is an equation for every orbital the number of equations increases with the number of electrons in the system. In OF-DFT we only have one equation to solve irrespective of the number of electrons in the system. This means OF-DFT could, in theory, be a linear scaling *ab-inito* method with a very low pre-factor such that biochemical and soft matter systems could be studied exclusively using OF-DFT. For example, the PROFESS code⁸² scales quasi-linearly ($\mathcal{O}(n \log(n))$). So that no confusion can arise we will refer to the KEF used in OF-DFT as the orbital free KEFs (OF-KEFs).

The remaining functionals in Equation (4.1) are known exactly. $E_{\rm J}$ is known as the Coulomb or Hartree repulsion and is expressed as

$$E_{\mathbf{J}}(\rho) = \frac{1}{2} \iint d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$
(4.2)

This is the energy arising from the electron density interacting with itself. $E_{\text{ext}}(\rho; \upsilon)$ is given by

$$E_{\text{ext}}(\rho; \upsilon) = \int d\mathbf{r} \rho(\mathbf{r}) \upsilon(\mathbf{r}) := (\rho | \upsilon)$$
(4.3)

and is the energy due to the electron density interacting with a attractive external potential, v. In this work the potential is the sum of the Coulomb potentials generated by a set of nuclei, $\{I\}$,

$$\upsilon(\mathbf{r}) = -\sum_{A \in I} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|}.$$
(4.4)

In this chapter and the subsequent chapters $\mathbf{r}, \mathbf{r}_i \in \mathcal{R}^3$.

4.2. The Chan Cohen Handy Scheme

4.2.1. Outline of the Scheme

We know from the discussion in Section 3.4.2 that the Lagrangian of an N electron system is a saddle function defined by, see Equation (3.87),

$$\mathcal{L}(\rho,\mu) = F_{\rm L}(\rho) + (\upsilon|\rho) + \mu(N - N_{\mu}) \tag{4.5}$$

and the ground state energy, $E_{\text{GS}} : \mathcal{X}^* \times \mathcal{R} \to \mathcal{R}$, is found through the saddle point optimisation

$$E_{\rm GS}(v,N) = \inf_{\rho \in \mathcal{X}} \sup_{\mu \in \mathcal{R}} \mathcal{L}(\rho,\mu).$$
(4.6)

In this chapter we use the ideas of CCH^2 who, to avoid this saddle point optimisation, carried out Equation (4.6) in a stepwise manner.

First one chooses a value of μ which is held fixed and then one performs the optimisation

$$G(\upsilon;\mu) = \inf_{\rho \in \mathcal{X}} \mathcal{L}(\rho;\mu)$$
(4.7)

where the notation $(v; \mu)$ and $(\rho; \mu)$ denotes the fact that the ρ and v are being allowed to change and μ is kept fixed. The optimising density, ρ_{μ} , for Equation (4.7) will be normalised to N_{μ} electrons. This quantity can be calculated through

$$N_{\mu} = \int \rho_{\mu}(\mathbf{r}) \mathrm{d}\mathbf{r}.$$
 (4.8)

In general this will not be the correct number of electrons in the system but it will be either an over or under estimate due to the initial guess of μ . Since μ determines N_{μ} , the optimisation of Equation (4.6) can be repeated for different values μ' until we have $N_{\mu} < N < N_{\mu'}$. To search for the μ which yields a correctly normalised density CCH proposed to bisect a function $B : \mathcal{R} \to \mathcal{R}$

$$B(\mu) = N_{\mu} - N.$$
 (4.9)

We can identify $B(\mu)$ as the particle number error for a given μ and so, we are searching for a μ for which $B(\mu) = 0$.

Since $\mu \in \mathcal{R}$ this can be achieved by a simple bisection. The bisection method we used begins with two values of μ ; μ_{low} and μ_{high} . These values are chosen such that $B(\mu_{\text{low}}) < 0$ and $B(\mu_{\text{high}}) > 0$. Then we choose μ_{mid} which is computed by

$$\mu_{\rm mid} = \frac{\mu_{\rm high} - \mu_{\rm low}}{2}.$$
(4.10)

Then Equation (4.7) is performed using $\mu_{\rm mid}$ which yields $B(\mu_{\rm mid})$ through Equations (4.8) and (4.9). If $B(\mu_{\rm mid}) < 0$ then we set $\mu_{\rm low} = \mu_{\rm mid}$. Or if $B(\mu_{\rm mid}) > 0$ then we set $\mu_{\rm high} = \mu_{\rm mid}$. We then iterate through the bisection until $|B(\mu_{\rm mid})| < 10^{-6}$. In Figure 4.1 we present an example of $B(\mu)$ and the bisection steps from an OF-DFT calculation. In the figure we can see the computed values of $\mu_{\rm mid}$ — represented by dots in Figure 4.1 — become more tightly packed around the final value of μ . Indeed this is a symptom of all calculations in CCH: the bisection is a slow algorithm



Figure 4.1.: $B(\mu)$ generated using the E00 OF-KEF and PBE XC functionals for the Ne atom.

when one nears the solution point.

4.2.2. The Choice of Basis Functions

To ensure that we are optimising in the correct space, i.e. $\rho \in \mathcal{X}$, we follow CCH and expand the square root of the density as follows

$$\rho^{1/2}(\mathbf{r}) = \sum c_i \eta_i(\mathbf{r}). \tag{4.11}$$

The sum is conducted over the entire basis and the $\{c_i\}$ are the basis coefficients. CCH proposed an even tempered basis set of real Gaussians of the form

$$\eta = (x - x_A)^i (y - y_A)^j (z - z_A)^k e^{-\alpha |\mathbf{r} - \mathbf{r}_A|^2} \quad \text{where} \quad \alpha = 3^n \tag{4.12}$$

and are centred around each nuclear centre, A. This basis set will be referred to as the even-tempered orbital free basis (ETOF). For atomic calculations CCH found that using an ETOF basis with n = -4, ..., 14for all the TFD λ W type models yields energies typically within 1 milli Hartree (mH) of the basis set limit. As all atoms have spherically symmetric charge distributions we also use ETOF in our atomic calculations. We conducted an investigation into whether the basis set convergence analysis conducted by CCH carries over in atomic calculations using a wider range of functionals.



Figure 4.2.: Comparison of basis set convergence using a range of OF-KEFs introduced in Section 3.3

CHAPTER 4. THE IMPORTANCE OF SELF CONSISTENCY IN ORBITAL FREE DENSITY FUNCTIONAL THEORY

In Figure 4.2 we plot the energy difference in Hartree $(E_{\rm h})$. The difference is computed from the energy calculated using a basis set of the form in Eq. (4.12), with the maximum value of n given as the greatest value of the abscissa in the figure. We plot these values, for a fixed minimum value of n = -4, as a function of n_{max} using a range of OF-KEFs using the PBE XC. We did this as adding more negative values of n did not change the energy significantly. As one can see the energy rapidly changes until we reach $n_{\rm max} = 12$. These results suggest that for trustworthy energies we need an accurate description of the core in our basis set rather than the tail to close in on the basis limit for multi-electron systems. Therefore, the basis set requires a large amount of sharp Gaussians. We have found that with the ETOF basis $(n_{\text{max}} = 14)$ we can get to within a few µHartree of the basis set limit for all functionals for all atomic systems. The only exception to this was the TF functional which was also found in the original work.² As explained in that work the TFD densities are singular at the nucleus and contribute a significant proportion to the energy and, therefore, one cannot expect to capture such behaviour using a basis set procedure such as the one presented here. For the diatomic systems we studied in this work we must consider adding higher angular momentum Gaussians to the basis set in order to describe the polarisation of the density for molecular systems. We, therefore, conducted a investigation in the convergence of the energy when one increases the number of basis functions and increase the number of functions with different angular momenta.

In Table 4.1 we show the ground state energies of the N_2 molecule using the TFPWOPT model — defined in Equation (4.29) — for $F_{\rm L}$ computed using a wide range of basis sets for the expansion of $\rho^{\frac{1}{2}}$. We began by increasing the number of s-type Gaussian functions, denoted by ETSn. We added a new, sharper Gaussian as n was increased because, as for atoms, adding diffuse functions had limited effect on the energy. As one can see in Table 4.1 the energies of the ETSn series decreases monotonically with basis size. The error measure, ΔE , in Table 4.1 is the difference between the energy computed using a given basis and the energy computed using the largest basis used — ETSPDF4. ΔE initially decreases quite quickly, though when one has around 18 s functions the error stays fixed at about $6 \times 10^{-2} E_{\rm h}$. Therefore, in our molecular basis we choose to have 18 sfunctions. We then constructed the ETSPn series where we fix the number of s functions to 18 and vary the p-functions from 1 to 10. Note one p function actually corresponds to 3 Gaussian's corresponding to p_x, p_y and p_x .

CHAPTER 4. THE IMPORTANCE OF SELF CONSISTENCY IN ORBITAL FREE DENSITY FUNCTIONAL THEORY

Basis	Composition	Energy (E_h)	$\Delta \overline{E}(E_h)$
ETS1	5s	-38.61915	6.973E+01
ETS2	6s	-53.83645	$5.452E{+}01$
ETS3	$7\mathrm{s}$	-71.51687	3.684E + 01
ETS4	8s	-82.95161	$2.540E{+}01$
ETS5	9s	-93.99829	$1.436E{+}01$
ETS6	10s	-100.69089	$7.663E{+}00$
ETS7	11s	-106.58877	$1.765E{+}00$
ETS8	12s	-107.82407	5.299 E-01
ETS9	13s	-108.17835	1.756E-01
ETS10	14s	-108.26202	9.194 E-02
ETS11	15s	-108.27971	7.426E-02
ETS12	16s	-108.28328	7.069E-02
ETS13	17s	-108.28400	6.996E-02
ETS14	18s	-108.28414	6.982E-02
ETS15	19s	-108.28417	6.979 E-02
ETSP1	18s1p	-108.28415	6.981E-02
ETSP2	18s2p	-108.28635	6.761E-02
ETSP4	18s4p	-108.34476	9.197E-03
ETSP5	18s5p	-108.35241	1.600E-03
ETSP6	18s6p	-108.35241	1.551E-03
ETSP7	18s7p	-108.35243	1.535E-03
ETSP8	18s8p	-108.35243	1.533E-03
ETSP9	18s9p	-108.35243	1.532E-03
ETSP10	18s10p	-108.35243	1.532E-03
ETSPD1	18s5p1d	-108.35239	1.575E-03
ETSPD2	18s5p2d	-108.35239	1.576E-03
ETSPD3	18s5p3d	-108.35300	9.579E-04
ETSPD4	18s5p4d	-108.35386	1.033E-04
ETSPD5	18s5p5d	-108.35392	4.598E-05
ETSPDF1	18s5p5d1f	-108.35393	3.476E-05
ETSPDF2	18s5p5d2f	-108.35395	1.208E-05
ETSPDF3	18s5p5d3f	-108.35396	5.000 E-09
ETSPDF4	18s5p5d4f	-108.35396	0

Table 4.1.: Energies in E_h computed for N₂ using TFPWOPT for a range of basis functions. The form of the error measure, ΔE , is discussed in the text

CHAPTER 4. THE IMPORTANCE OF SELF CONSISTENCY IN ORBITAL FREE DENSITY FUNCTIONAL THEORY

Once again we see a monotonic decrease in the energy and we see the error in the energy stabilises around ETSP8. Note by adding p-functions to the basis set we lower the energy by around 50 mH. In Table 4.1 we see that adding higher angular momentum functions does lead to a decrease in energy. We, therefore, chose to use the ETSPDF4 basis set as we are confident we are at least within 0.1 mH of the basis set limit due to monotonic decrease to the ETSPDF4 value and the stability of the error. We repeated this analysis using several more functionals — E00, OL1, P92 and SGA and found the ETSPDF4 was at most 0.1 mH in error. To summarise we are using the ETOF basis developed by CCH for atoms which consists of s-type Gaussians with exponents of 3^n for every *n* between and including -4 and 14. For brevity this can be stated as the exponents are of the form $3^{[-4,14]}$. We are using the ETSPDF4 basis set which we will now call ETOF-MOL. This basis consists of: s-type Gaussians with an exponent of $3^{[-6,11]}$; p-type Gaussians with an exponent of $3^{[-4,0]}$; d-type Gaussians with an exponent of $2.7^{[-3,1]}$; f-type Gaussians with an exponent of $2.7^{[-2,1]}$. We have used the aug-cc-pV5Z basis,¹²¹⁻¹²⁵ utilising spherical harmonics and with no contraction, for the KS-DFT calculations in this chapter for both atoms and molecules

4.2.3. Convergence Properties of the CCH Scheme

Figure 4.3 outlines the CCH approach that has been implemented into QUEST during this work. The tolerance value was set to 10^{-6} a.u. on the bisection. In the minimisation of \mathcal{L} we set a tolerance on the gradient and energy of 10^{-6} a.u. The starting guess of the chemical potential μ_0 was always chosen to be 0. This choice had little impact on the convergence performance of the scheme compared to the choice of the initial basis coefficients, $\{c_i^0\}$. For atomic calculations one can start with a random guess, however due to using approximate forms of $T_{\rm s}$ and $E_{\rm XC}$ the Lagrangian may have many saddle points. The CCH scheme makes it difficult to ascertain whether we have converged to a first order saddle point or some other higher order saddle point. This is in contrast to the TRIM scheme which we introduce in Chapter 6. We have found that using $\{c_i\}$ from the ground state of the core Hamiltonian is a quick method to generate starting guesses which allows convergence to the correct saddle points. However, for molecules this is unsuitable as it yields very inaccurate starting guesses and thus causes problems for the optimisation. For molecules we found that using the supposition of atomic densities (SAD) yields suitable initial guesses.


Figure 4.3.: A flow chart for the self-consistent OF-DFT scheme discussed in this thesis

We run atomic calculations in the ETOF-MOL basis for every atom type in the molecule and then we combine these converged basis coefficients in one coefficient vector which forms our starting guess molecule.

In Section 3.3.2 we discussed the Lopez-Acevedo scheme, ¹ see Fig. 3.1. The CCH scheme, in contrast to the Lopez-Acevedo scheme, is a direct optimisation approach and no diagonalisation step is ever performed. Instead the optimisation of $\mathcal{L}(\rho; \mu)$ is carried out via a quasi-Newton method by differentiating $\mathcal{L}(\rho; \mu)$ with respect to the expansion coefficients, $\{c_i\}$. At each step in the cycle in Figure 4.3 an optimising set of $\{c_i\}$ is obtained. When $|B(\mu)|$ becomes small then using the set of $\{c_i\}$ as the initial guess of the next cycle significantly accelerates the optimisation of $\mathcal{L}(\rho; \mu)$. One of the key advantages of the CCH scheme is that at the core of this scheme is a minimisation problem which is well suited to a robust and simple (quasi) Newton method. However, the major drawback is the nested nature of the scheme which leads to a large number of Lagrangian optimisations to be performed and hence a large iteration count.

4.3. The Gradient and Hessian

We will now discuss how we calculate the gradient and Hessian elements which are used in the CCH scheme. This is where this work diverges from the original scheme developed by CCH. In the original work the gradient and Hessians were explicitly derived and implemented for the TF and VW OF-KEF and the Dirac exchange functional. In this section we will discuss how we have implemented the optimisation scheme in a such a way that the gradient and Hessians corresponding to any OF-KEF or XC functionals can be calculated without the need for explicitly implementing derivatives for each functional. We will show the general form of the gradient and Hessian and discuss how we have used these in conjunction with XCFun¹¹³ to generate any gradient and Hessian just by coding the integrand of a functional into XCFun. This usually involves very little coding on behalf of the user and is a more 'black-box' method than the original implementation.

4.3.1. The Line Search Algorithm

We have implemented a line search algorithm in order to find the set of basis coefficients $\{c_k\}_{opt}$ which minimise Equation (4.7). We have chosen to base this on the Newton Search which relates the next guess of $\{c_k\}_{opt}$, x_{k+1} , from the previous guess x_k through

$$x_{k+1} = x_k + \alpha_k g_k H_k^{-1} \tag{4.13}$$

where a_k is known as the step size, g_k is the gradient of Equation (4.7) evaluated using x_k and H_k is the Hessian of Equation (4.7) evaluated using x_k . $g_k H_k^{-1}$ is known as the descent direction. In Sections 4.3.2 and 4.3.3 we will discuss the gradient and the Hessian respectively. Equation (4.13) states that $x_{k+1} = x_k$ when one has minimised Equation (4.7) and it is an equation we have to solve iteratively due to the non-linear nature of the problem. We must be careful in our choice of α_k as if its too small we would never make progress but if it is too large we could get trapped between two points on the surface with equal values. We have tried two different algorithms to find α_k .

The first algorithm is the Armijo-backtracking search, which is a wellknown algorithm and suitable for most (quasi) Newton searches.¹²⁶ This

algorithm ensures that we chose a step size such that we take a step which reduces the value of the objective function, which in our case is Equation (4.5). In addition to this, the backtracking iteration cycle ensures that from an initial guess of α_0 we pick the largest possible α such that $\mathcal{L}(x_{k+1}) \leq \mathcal{L}(x_k)$. The algorithm to find α_k is set out in Algorithm 1 where $p_k = g_k H_k^{-1}$.

Algorithm 1: Armijo-backtracking search
Choose $\bar{\alpha} > 0, t \in (0, 1), c \in (0, 1);$
Set $\alpha \leftarrow \bar{\alpha}$;
while $f(x_k + \alpha p_k) > f(x_k) + ctg_k^T p_k$ do
$\alpha \leftarrow t\alpha;$
end

After some trial runs we found that setting t = 0.5 and $c = 10^{-4}$ yields the most optimal convergence rates. Whilst this was a robust approach we found that we could reduce the number of steps required to find an appropriate α_k using quadratic and cubic interpolation.¹²⁶ This generates a decreasing sequence of $\{\alpha_k\}_i$ but in such a way that each $\{\alpha_k\}_i$ is not too small. Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a step size $\alpha_k \in \mathbb{R}^n$ the sufficient decrease condition can be written as

$$f(\alpha_k) \le f(0) + c\alpha_k \nabla f(0) \tag{4.14}$$

where $c \in \mathcal{R}$ is a small number, which we have set to 10^{-4} in this work. Let α_0 be the initial guess. We can form a quadratic approximation $f_Q(\alpha)$ to f by interpolating between $f(0), \nabla f(0)$ and $f(\alpha_0)$. Its explicit form is

$$f_Q(\alpha) = \left(\frac{f(\alpha_0) - f(0) - \alpha_0 \nabla f(0)}{\alpha_0^2}\right) \alpha^2 + \nabla f(0) \alpha + f(0).$$
(4.15)

If we minimise this with respect to α we obtain a new trial value, α_1

$$\alpha_1 = -\frac{\nabla f(0)\alpha_0^2}{2[f(\alpha_0) - f(0) - \nabla f(0)\alpha_0]}.$$
(4.16)

If Equation (4.14) is satisfied by setting $\alpha_k = \alpha_1$ we terminate the search. Otherwise we construct a cubic approximation, $f_C(\alpha)$ that interpolates between $f(0), \nabla f(0), f(\alpha_0)$ and $f(\alpha_1)$. Its explicit form is

$$f_C(\alpha) = a\alpha^3 + b\alpha^2 + \alpha \nabla f(0) + f(0), \qquad (4.17)$$

where

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\alpha_0^2 \alpha_1^2 (\alpha_1 - \alpha_0)} \begin{bmatrix} \alpha_0^2 & -\alpha_1^2 \\ -\alpha_0^3 & \alpha_1^3 \end{bmatrix} \begin{bmatrix} f(\alpha_1) - f(0) - \nabla f(0)\alpha_1 \\ f(\alpha_0) - f(0) - \nabla f(0)\alpha_0 \end{bmatrix}.$$
 (4.18)

Minimising Equation (4.17) with respect to α yields the minimiser α_2 defined through

$$\alpha_2 = \frac{-b + \sqrt{b^2 - 3a\nabla f(0)}}{3a}.$$
(4.19)

If necessary this process is repeated i.e. forming f_C using f(0), $\nabla f(0)$ and the two most recent values of f until we find an α to satisfy Equation (4.14). We have implemented the following algorithm

Algorithm 2: Armijo-backtracking with quadratic/cubic interpola-

tion

```
Choose \alpha_0 > 0, c \in (0, 1);
Set \alpha \leftarrow \bar{\alpha};
if f(\alpha_0) \leq f(\alpha) + c\alpha_0 g_k^T p_k then
 return \alpha_k = \alpha_0;
end
Form \alpha_1 using Equation (4.16);
if f(\alpha_1) \leq f(\alpha) + c\alpha_1 g_k^T p_k then
  return \alpha_k = \alpha_1;
end
while \alpha_1 > 0 do
     Set \rho = 0.3;
     Calculate a, b using Equation (4.18);
    if b^2 - 3ag_k^T p_k < 10^{-6} then
      | \alpha_2 = \alpha_1 \rho ;
     else
          Calculate \alpha_2 using Equation (4.19);
     end
     if f(\alpha_2) \leq f(\alpha) + c\alpha_2 g_k^T p_k then
     return \alpha_k = \alpha_2;
     end
     Set \alpha_0 = \alpha_1;
     Set \alpha_1 = \alpha_2;
end
If \alpha_1 < 0 exits while loop with error;
```

Each step in Algorithm 2 compared to Algorithm 1 is more expensive due to the fact that we have to generate the Hessian and gradient twice as many times per step. However, the number of iteration cycles decreases on average by a factor of five and so, we have found Algorithm 2 is much more efficient and has been used to generate the results presented in this work.

4.3.2. The Gradient

In Equation (4.13) we saw that the gradient, g_k , and the Hessian H_k are required to compute the step direction in the optimisation. The gradient of the Lagrangian given in Eq. (4.5) is

$$g = \left(\cdots, \frac{\partial \mathcal{L}}{\partial c_i}, \cdots\right). \tag{4.20}$$

This may be computed, using Equation (4.5) and expanding $\rho^{1/2}$ in a basis set, Equation (4.11),

$$\frac{\partial \mathcal{L}}{\partial c_i} = 2 \int \eta_i \left(\frac{\delta E}{\delta \rho} - \mu \right) \rho^{\frac{1}{2}} \mathrm{d}\mathbf{r}.$$
(4.21)

Noting that $\frac{\delta E}{\delta \rho}$ may be decomposed in the usual manner — as seen in KS-DFT — and so, the evaluation of Equation (4.21) requires construction of the following matrix elements

$$\langle \eta_i | \frac{\delta T_{\rm s}}{\delta \rho} + \frac{\delta E_{\rm XC}}{\delta \rho} + \frac{\delta E_{\rm J}}{\delta \rho} + \upsilon - \mu | \eta_j \rangle .$$
 (4.22)

The matrix elements involving v, $\frac{\delta E_{\rm XC}}{\delta \rho}$ and $\frac{\delta E_{\rm J}}{\delta \rho}$ are evaluated in the same manner as in KS-DFT calculations. For the $\frac{\delta T_{\rm S}}{\delta \rho}$ we have implemented OF-KEFs into the XCFun¹¹³ package, which we interfaced with QUEST, and are then handled by the same numerical integration techniques employed for XC functionals in standard KS-DFT calculations. The most costly part of the evaluation of Equation (4.22) is for $\frac{\delta E_{\rm J}}{\delta \rho}$. However, this may be subject to the same acceleration and approximations techniques utilised in KS-DFT. In particular one can use density-fitting or J-engine type evaluations for these matrix elements, which should be pursued in future work. These approaches have been enabled using existing routines in the QUEST program. In this work we use conventional integral evaluations to allow evaluation of the approximate KEFs. See Appendix A for further details of the derivation of the gradient terms.

4.3.3. Hessian

As highlighted above we are using a (quasi)Newton search which requires a Hessian. Whilst one can use numerical approximations to the Hessian we have decided to derive and implement the analytical Hessian for a general pair of $E_{\rm XC}$ and $T_{\rm s}$ functionals. The full details of the derivation of the equations presented in this section are given in Appendix A. The reason for using analytical rather than numerical approximations is that the inverses of numerical Hessians can be 'spiky' and can be very different to the analytical for non-linear problems¹²⁷ — which our problem is. Furthermore numerical evaluations of the Hessian is a slow process.

The elements of the Hessian of the Lagrangian in Equation (4.5) are:

$$\frac{\partial^{2} \mathcal{L}}{\partial c_{i} \partial c_{j}} = \frac{\partial^{2} T_{s}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{XC}}{\partial c_{\partial} i c_{j}} + \frac{\partial^{2} E_{J}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{ne}}{\partial c_{i} \partial c_{j}} - 2\mu \int \eta_{i} \eta_{j} d\mathbf{r}.$$
(4.23)

The fourth term in Equation (4.23) evaluates to

$$\frac{\partial^2 E_{\rm ne}}{\partial c_i \partial c_j} = 2 \int \frac{\delta E_{\rm ne}}{\delta \rho} \eta_i \eta_j d\mathbf{r}.$$
(4.24)

The third term in Equation (4.23), using Equation (4.2), evaluates to

$$\frac{\partial^2 E_{\mathcal{J}}}{\partial c_i \partial c_j} = 2 \sum_{kl} c_k c_l [(ij|kl) + 2(ik|jl)], \qquad (4.25)$$

where

$$(mn|pq) = \iint \frac{\eta_m(\mathbf{r}_1)\eta_n(\mathbf{r}_1)\eta_p(\mathbf{r}_2)\eta_q(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2.$$
(4.26)

For the first two terms in Equation (4.23) we assume the functionals can be expressed as

$$F(\rho) = \int f(\rho, \nabla \rho) d\mathbf{r} := \langle f(\rho, \nabla \rho) \rangle, \qquad (4.27)$$

then one finds

$$\frac{\partial^{2} F}{\partial c_{i} \partial c_{j}} = \left\langle \frac{\partial^{2} \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \rho} + \frac{\partial^{2} \nabla \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \nabla \rho} \right\rangle
+ \left\langle \frac{\partial^{2} f}{\partial \rho^{2}} \frac{\partial \rho}{\partial c_{i}} \frac{\partial \rho}{\partial c_{j}} + \frac{\partial^{2} f}{\partial \nabla \rho^{2}} \frac{\partial \nabla \rho}{\partial c_{i}} \frac{\partial \nabla \rho}{\partial c_{j}} \right\rangle
+ \left\langle \frac{\partial^{2} f}{\partial \nabla \rho \partial \rho} \left[\frac{\partial \rho}{\partial c_{j}} \frac{\partial \nabla \rho}{\partial c_{i}} + \frac{\partial \rho}{\partial c_{i}} \frac{\partial \nabla \rho}{\partial c_{j}} \right] \right\rangle.$$
(4.28)

In Appendix A we give the explicit equations for the Hessian which I derived and implemented into QUEST. In our framework we have an interface with XCFun which is an automatic differentiator developed for use in DFT. Given the integrand of an OF-KEF (or XC) it will return a numerically exact derivative of the integrand up-to any order at all grid points. The integrands of all functionals discussed in this work have been coded into XCFun.

The derivation and implementation of the Hessian was a long process. In order to test the accuracy of the implementations of the gradient and Hessian we used the Lagrangian and finite difference methods. Every element of the gradient and Hessian were within 10^{-6} a.u. of the corresponding finite difference gradient or Hessian. We did this checking process for atoms and molecules and for basis functions with angular momentum quantum number upto three i.e an f orbital.

4.4. Results

We investigated the following OF-KEFs using the optimiser discussed above: Borgoo and Tozer's BT2 and BT3;¹⁰⁷ conj-B86A;⁹² conj-B86B;⁹² conj-PW91;⁹² conj-PW86;⁹² DePristo-Kress (DK);¹¹¹ E00;¹¹² Lembarki and Chermette (LC94);⁹³ Ou-Yang and Levy's OL1 and OL2;¹⁰⁹ P92;¹¹⁰ Karasiev, Trickey and Harris' PBE2, PBE3 and PBE4;^{97,99} Thakkar (T92);⁹⁵ Tran and Wesolowski (TW02);⁹⁴ and Karasiev, Chakraborty, Shukruto and Trickey's VT84.⁹⁶ The accuracy of these functionals were assessed using post Kohn– Sham analysis (PKSA) — see Section 3.3.2. In this section the KS orbitals were expanded in the uncontracted, spherical aug-cc-pV5Z basis set. In addition to these functionals we have also investigated functionals of the form

$$T_{\rm s}(\rho) = \alpha T_{\rm TF}(\rho) + \beta T_{\rm VW}(\rho). \tag{4.29}$$

The coefficients α and β have the following values; for TFD $\alpha = 1$, $\beta = 0$ and using the Dirac exchange; for SGA $\alpha = 1$, $\beta = \frac{1}{9}$ and using the Dirac exchange; for TFD02W $\alpha = 1$, $\beta = \frac{1}{5}$ and using the Dirac exchange; for TFDW $\alpha = 1$, $\beta = 1$ and using the Dirac exchange; for TFDWOPT $\alpha =$ 0.697, $\beta = 0.599$ and using the Dirac exchange; for TFPWOPT $\alpha = 0.697$, $\beta = 0.599$ and using the PBE XC functionals. To check we had reached a global minima we ran atomic calculations with 26 randomly chosen different sets of starting coefficients. We present the lowest energy found from these sets for functionals with which 10 or less stationary values were found. In the majority of cases we converged to three or less stationary values. The functionals which this occurred for were DK, E00, LC94, OL1, OL2, P92, PBE2 and VT84, for all atoms with $Z = 1, \dots, 10$. The lowest energy solutions found are presented in this chapter.

Calculations using BT2 and BT3 failed to converge because their functional derivatives diverge in some regions of space and, therefore, Equation (4.7) cannot be set-up. For GGA functionals divergences in their functional derivatives are often encountered at the nuclei, however they do not, in general, contribute to the matrix element $\langle \eta_i | \frac{\delta T_s}{\delta \rho(\mathbf{r})} | \eta_j \rangle$. However, additional divergences of this derivative can cause these matrix elements to diverge and calculations cannot be carried out. These additional divergences in the BT2/BT3 OF-KEF were caused by the functional form which has the term $/|\nabla \rho|^k / \rho^{4k/3}$.^{104,106,107} As $k \neq 1$ this term causes non-physical divergences in the potential. This highlights that performing PKSA is not helpful in designing functionals that can work in a self consistent OF-DFT calculation.

For the other OF-KEFs, we could not guarantee we had reached a global minimum as different starting guesses usually yielded different final energies even though the tolerance criterion had been met, which suggests these functionals plus the PBE XC functional are not convex. This is in contrast to our previous discussion on $F_{\rm L}$ which was shown to be convex. This should serve as warning about developing OF-KEFs without consideration of the number of local minima a particular OF-KEFs.

4.4.1. Atomic Systems

In Table 4.2 we present the total energies for the neutral atomic series $Z = 1 \cdots 18$. We have used the ETOF basis sets introduced in Section 4.2.2. In addition we have included the mean absolute percentage error (MAPE) and mean absolute error (MAE) with respect to the appropriate KS-DFT value, see the figure for details. We see that in Table 4.2 the TFD02W functional yields the most accurate energies with a MAPE and MAE of 4.3% and $0.79E_{\rm h}$ respectively. However, it has been shown that TFD02W yields very small binding energies for diatomics.² Whereas the PBE2 and VT84, which were developed to reproduce molecular binding not accurate energies, perform the worst using this measure. The most pleasing results were the TFWOPT numbers. This was a functional that was developed to reproduce bound states but we find that it predicts relatively accurate energies. We also find that the choice of XC functional does not alter the

MAE	9.25	4.10	5.45	4.49	4.72	5.02	16.16	14.39	0.90		18.03	4.98	0.79	15.12	0.90	
MAPE 1	30.8	13.1	18.8	15.1	15.8	16.5	37.7	35.4	5.4		55.6	16.3	4.3	38.1	5.7	
Ne	-151.565628	-139.022632	-142.033067	-139.826653	-140.409304	-141.281533	-83.765636	-89.642398	-131.082243	-128.866098	-173.580144	-139.886483	-128.801430	-85.734277	-129.768777	-127.4904408
Гц	-118.816516	-108.422382	-111.003159	-109.132513	-109.609142	-110.286535	-64.440165	-68.750063	-100.857229	-99.675892	-136.415760	-109.059246	-100.134516	-65.767433	-99.713816	-98.477731
0	-90.528699	-82.136859	-84.301324	-82.751876	-83.131911	-83.642553	-48.076878	-51.114248	-75.234943	-75.014867	-104.247207	-82.579766	-75.576534	-48.883118	-74.256622	-73.999839
Z	-66.555995	-59.980471	-61.745793	-60.500251	-60.79358	-61.164282	-34.507900	-36.541920	-53.964777	-54.535654	-76.892920	-60.262253	-54.942825	-34.908352	-53.145969	-53.709200
D	-46.734596	-41.751999	-43.140648	-42.177684	-42.394650	-42.650842	-23.552941	-24.825370	-36.776582	-37.798676	-54.155754	-41.905213	-38.033249	-23.656810	-36.110959	-37.113587
В	-30.919626	-27.230771	-28.271832	-27.564994	-27.716417	-27.881999	-15.016626	-15.739415	-23.376290	-24.612180	-35.819120	-27.287592	-24.628398	-14.925838	-22.856516	-24.067774
Be	-18.662235	-16.171104	-16.900125	-16.418289	-16.515446	-16.612622	-8.684728	-9.037581	-13.439628	-14.629912	-21.641308	-16.163127	-14.484086	-8.492154	-13.056979	-14.223277
Li	-9.803161	-8.293546	-8.755039	-8.460393	-8.515009	-8.564039	-4.312219	-4.446804	-6.602616	-7.462145	-11.346536	-8.251475	-7.322708	-4.105406	-6.346338	-7.19339
He	-3.998840	-3.269884	-3.524686	-3.366124	-3.390310	-3.409162	-1.641035	-1.659598	-2.445982	-2.892884	-4.609343	-3.222822	-2.818354	-1.477441	-2.302017	-2.723591
Н	-0.910231	-0.692879	-0.783477	-0.732692	-0.738786	-0.742613	-0.344331	-0.321730	-0.465707	-0.499985	-1.022428	-0.666438	-0.566623	-0.261823	-0.413651	-0.457073
OF-KEF	DKa	$E00^{a}$	$LC94^{a}$	$OL1^{a}$	$OL2^{a}$	$P92^{a}$	$PBE2^{a}$	$\rm VT84^{a}$	$TFPWOPT^{a}$	KS-PBE	TFD^{b}	$\rm SGA^b$	$\mathrm{TFD02W^{b}}$	TFDW^b	$TFDWOPT^{b}$	KS-D

able 4.2.: Self consistent Total energies in E_h from solving Equation (4.7) using an OF-KEF and a) PBE XC or b) Dirac X as the XC functional. The MAPE/MAE was computed using the KS-PBE values for the former and the KS-D values for the latter
--

KS-D	TFDWOF	$\mathrm{TFDW}^{\mathrm{I}}$	TFD02W	SGA^{b}	$\mathrm{TFD}^{\mathrm{b}}$	KS-PBI	TFPWOF	$VT84^{a}$	$PBE2^{a}$	$P92^{a}$	$OL2^{a}$	$OL1^{a}$	$LC94^{a}$	$E00^{a}$	$\mathrm{D}\mathrm{K}^{\mathrm{a}}$	OF-KEI	Table 4.3.
-160.64401	$ T^{b} - 164.65631$	-108.94547	$^{\rm /b}$ $ -161.74914$	-175.23453	-215.90910	1 - 162.17239	T^{a} -166.14431	-113.97181	-106.21073	-176.80000	-175.70367	-175.00602	-177.56151	-174.11030	-189.21498	Na	: Self consisten: as the XC fun latter
2 -198.248755	2 -204.597284	1 - 135.553361	8 -199.138784	1 -215.264985	3 -263.560607	3 -199.955032	8 -206.263925	3 -141.908091	4 -131.923977	9 -217.003603	4 -215.652789	1 -214.831592	2 -217.749870	5 -213.847243	1 -231.763581	Mg	t Total energie ctional. The N
-240.358789	-249.800557	-165.702098	-241.122238	-260.130661	-316.683430	-242.236149	-251.649582	-173.611614	-161.046146	-262.044696	-260.407990	-259.455135	-262.752156	-258.386007	-279.268295	Al	s in E _h from : fAPE/MAE v
-287.182679	-300.464501	-199.528676	-287.843428	-309.976134	-375.418376	-289.234029	-302.499382	-209.234504	-193.711239	-312.067641	-310.112648	-309.020440	-312.716094	-307.871100	-331.842789	Si:	solving Equat vas computed
-338.88845	-356.778251	-237.163770	-339.439253	-364.938913	-439.899267	-341.115513	-359.002211	-248.921528	-201.161631	-367.209744	-364.903137	-363.664288	-367.784174	-362.439964	-389.617795	Р	ion (4.7) equations using the KS
-395.523789	-418.922704	-278.732434	-396.040408	-425.150224	-510.253753	-397.952566	-421.338763	-292.810900	-270.176755	-427.602066	-424.909630	-423.517254	-428.093965	-422.223787	-452.777376	S	ation using an PBE values t
-457.345243	-487.071377	-324.354651	-457.772040	-490.735705	-586.603974	-459.974418	-489.682359	-341.034918	-314.217568	-493.370100	-490.256769	-488.704371	-493.779403	-487.348169	-521.737512	CI	1 OF-KEF an for the former
-524.517258	-561.391127	-374.145854	-524.754301	-561.815964	-669.067116	-527.345904	-564.199703	-393.720546	-362.283078	-564.634316	-561.064216	-559.345682	-564.963154	-557.933683	-596.538561	Ar	nd a) PBE X(and the KS-
	4.9	30.3	0.3	7.9	30.6		4.8	27.4	33.9	7.9	7.2	6.8	8.1	6.4	14.6	MAPE	C or b) 1 D values
	17.62	97.32	0.64	25.07	96.84		17.60	88.10	109.91	25.09	22.88	21.69	25.68	20.52	46.60	MAE	Dirac X for the

102



Figure 4.4.: Mean absolute error (MAE) for Z = 1-18 for OF-KEF/PBE using KS-PBE as reference



Figure 4.5.: Mean percentage error (MPE) for Z = 1-18 for OF-KEF/PBE using KS-PBE as reference

errors significantly as demonstrated by the TFPWOPT and TFDWOPT which are the TFWOPT OF-KEF with the PBE XC and Dirac X respectively. This suggests that the $T_{\rm s}$ approximation is the dominant source of the errors we see.

In Table 4.3 we do the same analysis but for the neutral atomic series $Z = 19 \cdots 26$ where as expected, the MAE increases but the MAPE fall. More interestingly the TFD02W is very accurate over this series with a MAE and MAPE of $0.64E_{\rm h}$ and 0.3% respectively.

In Figures 4.4 to 4.7 we have compared the MAE and MPE one would get using a PKSA and a self-consistent OF-DFT analysis (SCOFA) by solving the appropriate variational problem given in Equation (4.7). There is a clear discrepancy between the PKSA and SCOFA and so, one should treat conclusions reached using PKSA with some caution. For the most part both



Figure 4.6.: Mean absolute error (MAE) for Z = 1-18 for OF-KEF/LDA using KS-LDA as reference



Figure 4.7.: Mean percentage error (MPE) for Z = 1-18 for OF-KEF/LDA using KS-LDA as reference

the MAE and MPE increase as one moves from PKSA and SCOFA which is expected by considering that the density in the SCOFA is qualitatively different to the accurate KS density used in the PKSA. In Figure 4.8 we plot $\delta_{\rm rel}$ as a function of the atomic number Z. We have defined this measure by

$$\delta_{\rm rel} = \frac{E_{\rm OF}(\rho^{\rm OF}) - E_{\rm KS}(\rho^{\rm KS})}{|E_{\rm KS}(\rho^{\rm KS})|}$$
(4.30)

where $E_{\rm OF}$ is the energy functional using a given OF-KEF, $E_{\rm KS}$ is the usual KS energy expression, $\rho^{\rm OF}$ is the density obtained by solving the variational problem defined in Equation (4.7) for a given OF-KEF and $\rho^{\rm KS}$ is the density calculated using the KS-DFT equations.

It is known that the VW functional is exact for one and two electron singlet systems and that the TF functional is exact in the infinite nuclear



Figure 4.8.: Relative error in Energy using an OF-KEF/LDA in SCOFA against KS-LDA reference

charge limit.^{70,71} From Figure 4.8 it is clear that the TFD model becomes more accurate for heavier nuclei, however convergence to zero error is slow. Therefore, the infinite nuclear charge limit may be of limited use in developing accurate functionals for chemical systems. What is also clear is that adding the VW functional to the TFD increases the predicted energy. Therefore, to get the best energy one needs to optimise the amount of TF and VW in the functional to get the best average energy. This was done for example — in the development of the TFDWOPT. The radial atomic



Figure 4.9.: Radial densities for H using an OF-KEF/PBE compared with KS-PBE reference

densities predicted by models of the form in Equation (4.29), TF λ W functionals, have been studied since the 1970's.^{71,128} The exact radial atomic densities should have as many maxima as principle quantum shells occupied in its ground state configuration. The reason for this is that each principle quantum shell has different exponential decay and so, the super-



Figure 4.10.: Radial densities for Ne using an OF-KEF/PBE compared with KS-PBE reference

position of these shells leads to a bumpy radial atomic density. As there are no orbitals in OF-DFT this is a hard thing to reproduce. Indeed the studies on TFD λ W have shown that no models exhibit shell structure.⁷¹ In Figures 4.9 and 4.10 we show the ground state radial density for Hydrogen and Neon resulting from solving the Equation (4.7) for these atoms for a representative sample of OF-KEFs. We also find that no shell structure is predicted.

4.4.2. Molecular Systems

OF-KEF	H_2	N_2	CO	O_2	LiH	MAPE	MAE
E00	-1.338813	-119.548317	-123.532251	-163.946769	-9.000562	10.8	7.04
OL1	-1.396608	-120.536981	-124.524472	-165.129629	-10.188673	15.3	7.92
OL2	-1.408265	-121.115127	-125.113651	-165.881530	-10.248937	15.9	8.32
P92	-1.417248	-121.867599	-125.890462	-166.912645	-10.302172	16.6	8.84
VT84	-0.575150	-73.596430	-76.437654	-102.759470	-4.850818	37.5	24.79
TFPWOPT	-0.967572	-108.353962	-112.418561	-150.871517	-8.152645	4.1	0.57
KS-PBE	-1.166681	-109.452999	-113.242362	-150.259157	-8.047317		
SGA	-1.417248	-120.074144	-124.092233	-164.797393	-8.920635	17.1	8.61
TFD02W	-1.078431	-109.520579	-113.297489	-150.873377	-7.903418	2.2	1.28
TFDW	-0.430446	-70.2473023	-72.978565	-98.265035	-4.469638	40.8	25.98
TFDWOPT	-0.854390	-106.7358153	-110.793175	-148.933531	-6.685150	6.7	0.73
KS-D	-1.043670	-107.755887	-111.530478	-148.234074	-7.704249		

Table 4.4.: Self Consistent Total Energies in E_h

We have also conducted a pilot study of how proposed OF-KEFs behave in self-consistent calculations for molecular systems. In all OF-DFT calculations we have used the ETOF-MOL basis. We have studied a small range of diatomics. In Table 4.4 we note that the error measures are not dissimilar to the atomic error measures, which seems to suggest that errors



Figure 4.11.: Predicted $E_{\text{bind}}/E_{\text{h}}$ for CO for some OF-KEFS and KS-DFT.



Figure 4.12.: Predicted $E_{\text{bind}}/E_{\text{h}}$ for H₂ for some OF-KEFS and KS-DFT.

introduced by $T_{\rm s}$ do not accumulate going from atomic to molecular systems. Due to the issues discussed regarding $T_{\rm s}$ approximations and their ability to predict stable molecules we have we have plotted the potential energy surfaces (PES's) for a few OF-KEFs for: CO in Figure 4.11; H₂ in Figure 4.12; and N₂ in Figure 4.13.

We have plotted the binding energy, E_{bind} , as a function of nuclear separation,

$$E_{\text{bind}}(\mathbf{R}) = E(\mathbf{R}) - \sum_{i=1}^{N} E_i$$
(4.31)

where $E(\mathbf{R})$ is the total energy with internuclear distance \mathbf{R} and E_i is the atomic energy for the *i*-th atom in an N atom molecule. We have also given the equilibrium geometries if the PES displays a minimum. We see that OL1 and E00 predict no bound states for any of the diatomics studied and this is typical for most of the functionals we have tested.

The only functionals to predict binding for all molecules in our test set



Figure 4.13.: Predicted $E_{\text{bind}}/E_{\text{h}}$ for N₂ for some OF-KEFS and KS-DFT.

are VT84 and TFPWOPT. This is to be expected considering that both functionals were designed to reproduce accurate binding energies. It seems that the H₂ molecule is the hardest to reproduce accurate binding curves despite the fact we know the exact $T_{\rm s}$ for this system — $T_{\rm VW}$. But for the iso-electronic CO and N₂ both VT84 and TFPWOPT produce qualitatively reasonable curves. But we should note that while VT84 predicts bound states the energies of these states are inaccurate.

For an instructive example in describing why different choices for the OF-KEF yield models which do or do not predict bound states, we look to the N₂ molecule, $X^1\Sigma_g^+$, previously studied using self-consistent codes.² For a covalently bound molecule the density between the atoms, must be greater than it is when the molecule is 'unbound'. To measure this qualitatively one can compute

$$\Delta \rho = \rho_{\rm mol} - \sum \rho_{\rm atoms} \tag{4.32}$$

where $\rho_{\rm mol}$ is, for example, the density of the molecule N₂ and $\sum \rho_{\rm atom}$ is the sum of the non-interacting atomic density profiles where the atoms are located at the correct equilibrium molecular positions. In Figure 4.14 we have plotted $\rho_{\rm bind}$ in the σ_v plane of the molecule for a KS-DFT calculation using PBE XC. As we can see the red regions, which correspond to greater density in the molecule with respect to the sum of the atoms, are where we would expect the covalent bond between the two nitrogen atoms to be. This is why we see KS-DFT predict a bound state in Figure 4.14.

Developing orbital free approximations to the Lieb functional which predict molecular binding is not guaranteed, even in PKSA studies. Take the E00 OF-KEF with the PBE XC functional. We have seen in Figure 4.13 that this functional does not predict a bound state at the correct geometry



Figure 4.14.: ρ_{bind} using KS-DFT (PBE) for the N₂ molecule

- or anywhere in fact. In Figure 4.15 we have plotted $\rho_{\rm bind}$ using the KS-DFT equilibrium geometry. As one can immediately tell this functional does not yield a molecular density which is great enough in the 'bonding' region. Indeed a large part of the density accumulates near the nuclei and this is the same story for any molecule and all functionals which do not predict binding.

So let us look at ρ_{bind} at the equilibrium geometry predicted using TFP-WOPT. In Figure 4.16 we have plotted ρ_{bind} generated using the SC scheme presented in this paper and using the bond length $r_{\text{eq}} = 0.920$ Å, see Figure 4.13 .It is interesting to find that there is indeed a red zone in the bond region which indicates why the functional predicts a lower energy for N₂ at this bond length than the sum of two non-interacting atoms positioned at either end of the bond. However, when one compares Figure 4.16 to Figure 4.14 we see a marked difference. By considering the discussion for the KS-DFT ρ_{bind} one can conclude that although the TFPWOPT does yield a binding molecular density profile it is wholly inadequate physically as it reproduces hardly any of the features predicted by KS-DFT.



Figure 4.15.: ρ_{bind} from calculations using the E00 OF-KEF and the PBE XC for the N₂ molecule

In computational chemistry one is sometimes interested in more than just than the ground state energy and potential energy curves. For example, chemists often use DFT to help predict chemical reaction mechanisms to aid their experimental work. Clearly here we require an accurate representation of the density as looking at the magnitude of the density gives possible sites of nucleophilic attack, for example. OF-DFT is currently not suitable for this kind of analysis but we hope that codes such as the one discussed in this chapter can help progress towards using OF-DFT in computational studies.

4.5. Conclusions

In this chapter we have discussed our implementation of the CCH scheme. Our implementation allowed more OF-KEFs to be tested in an all-electron self-consistent context. We have found that most OF-KEFs predict very poor densities which has effected the predicted energies. The errors are vastly greater than what is found when one uses post-Kohn Sham analysis. Indeed post-Kohn Sham analysis missed the fact that some functionals cannot be converged in self-consistent calculations due to their analytic form. This should serve as a warning before one conducts post-Kohn Sham analysis. We found that various linear combinations of $T_{\rm TF}$ and $T_{\rm VW}$ provided the most accurate functionals. Indeed for molecular systems TFPWOPT

CHAPTER 4. THE IMPORTANCE OF SELF CONSISTENCY IN ORBITAL FREE DENSITY FUNCTIONAL THEORY



Figure 4.16.: ρ_{bind} from calculations using the TFPWOPT functional for the N₂ molecule

and TFDWOPT provided bound states for molecules. However, it is still too inaccurate to be useful in chemical applications. The CCH scheme whilst robust is quite slow in convergence. The reason for this is the nested iterations which converge on the energy and density normalisation separately. In Chapter 6 we will introduce a scheme which converges on both simultaneously.

5. Potentials in OF-DFT

The chemical potential of the system, μ , for a given external potential v can be computed in the CCH scheme, see previous chapter, through

$$\mu = v_{\mathrm{Ts}}(\mathbf{r}) + v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{J}}(\mathbf{r}) + v_{\mathrm{XC}}(\mathbf{r}).$$
(5.1)

This equation is derived from the optimisation problem in Equation (4.7) and is usually called the Euler equation. This section will be focussed on the spatial dependence of the kinetic potential, $v_{\text{Ts}}(\mathbf{r}) = \frac{\delta T_{\text{s}}}{\delta \rho(\mathbf{r})}$, the Coulomb potential, $v_{\text{J}}(\mathbf{r}) = \frac{\delta E_{\text{J}}}{\delta \rho(\mathbf{r})}$ and the exchange-correlation potential $v_{\text{XC}}(\mathbf{r}) = \frac{\delta E_{\text{XC}}}{\delta \rho(\mathbf{r})}$. Because $\mu \in \mathcal{R}$ the spatial dependence of each of the potentials in Equation (5.1) should cancel. Therefore, there is a delicate balance between v_{Ts} and all the other potentials in Equation (5.1). To investigate this delicate balance for the OF-KEFs discussed in this thesis we have derived and implemented into QUEST the v_{Ts} and v_{XC} terms, for both spin unpolarised and spin polarised potentials. In this chapter we will present the equations in the spin unpolarised framework. The equations for these potentials in the spin polarised framework are given in Appendix B.1.

5.1. Equations Implemented for Computing the Potential

To derive the explicit form of the potential we define a functional with the following form

$$F[\rho] = \int f(\rho(\mathbf{r}), G^{\alpha\alpha}(\mathbf{r}), G^{\alpha\beta}(\mathbf{r}), G^{\beta\beta}(\mathbf{r})) d\mathbf{r}, \qquad (5.2)$$

where

$$G^{\sigma\sigma'}(\mathbf{r}) = \boldsymbol{\nabla}\rho^{\sigma}(\mathbf{r}) \cdot \boldsymbol{\nabla}\rho^{\sigma'}(\mathbf{r}).$$
(5.3)

The reason for doing this is that we again use an interface with XCFun which returns partial derivatives of the integrand with respect to G and secondly some correlation functionals, like PBE C, have non-zero $G^{\alpha\beta}$ because of the inclusion of the spin-polarisation factor, ζ . We also know from Section 1.2.6 that the functional derivative of a functional whose form is

$$F(\rho) = \int f(\rho(\mathbf{r}), \nabla \rho(\mathbf{r})) d\mathbf{r}, \qquad (5.4)$$

is given by

$$\upsilon_F(\mathbf{r}) := \frac{\delta F}{\delta \rho(\mathbf{r})} = \frac{\partial f}{\partial \rho(\mathbf{r})} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla} \rho(\mathbf{r})}.$$
 (5.5)

To get a potential expression more suitable for a functional of the form in Equation (5.2) we use

$$\frac{\partial f}{\partial \boldsymbol{\nabla} \rho(\mathbf{r})} = \boldsymbol{\nabla} \rho(\mathbf{r}) \left[\frac{\partial f}{\partial G^{\alpha \alpha}(\mathbf{r})} + \frac{\partial f}{\partial G^{\alpha \beta}(\mathbf{r})} + \frac{\partial f}{\partial G^{\beta \beta}(\mathbf{r})} \right], \quad (5.6)$$

whose derivation is included in Appendix A.1. To reduce notational clutter we write

$$\frac{\partial f}{\partial G^{\alpha\alpha}(\mathbf{r})} + \frac{\partial f}{\partial G^{\alpha\beta}(\mathbf{r})} + \frac{\partial f}{\partial G^{\beta\beta}(\mathbf{r})} = \sum_{\sigma \ge \sigma'} \frac{\partial f}{\partial G^{\sigma\sigma'}}.$$
 (5.7)

Using this we rewrite Equation (5.5) as

$$\upsilon_F(\mathbf{r}) = \frac{\partial f}{\partial \rho(\mathbf{r})} - \boldsymbol{\nabla} \cdot \left(\sum_{\sigma \ge \sigma'} \frac{\partial f}{\partial G^{\sigma \sigma'}} \boldsymbol{\nabla} \rho(\mathbf{r}) \right).$$
(5.8)

Given a scalar valued function $\phi : \mathcal{R} \to \mathcal{R}$ and a vector valued function $\mathbf{A} : \mathcal{R}^n \to \mathcal{R}$ the following relation holds

$$\boldsymbol{\nabla} \boldsymbol{\cdot} (\phi \mathbf{A}) = \boldsymbol{\nabla} (\phi) \boldsymbol{\cdot} \mathbf{A} + \phi \boldsymbol{\nabla} \boldsymbol{\cdot} \mathbf{A}.$$
 (5.9)

Using this we find that

$$\upsilon_F(\mathbf{r}) = \frac{\partial f}{\partial \rho(\mathbf{r})} - \boldsymbol{\nabla} \left(\sum_{\sigma \ge \sigma'} \frac{\partial f}{\partial G^{\sigma \sigma'}} \right) \cdot \boldsymbol{\nabla} \rho(\mathbf{r}) - \sum_{\sigma \ge \sigma'} \frac{\partial f}{\partial G^{\sigma \sigma'}} \boldsymbol{\nabla}^2 \rho(\mathbf{r}).$$
(5.10)

Using the chain rule on the second term we find

$$\upsilon_{F}(\mathbf{r}) = \frac{\partial f}{\partial \rho(\mathbf{r})} - \sum_{\sigma \geq \sigma'} \frac{\partial^{2} f}{\partial G^{\sigma\sigma'} \partial \rho} \nabla \rho(\mathbf{r}) \cdot \nabla \rho(\mathbf{r}) - \sum_{m \geq m'} \sum_{\sigma \geq \sigma'} \frac{\partial^{2} f}{\partial G^{\sigma\sigma'} \partial G^{m,m'}} \nabla G^{m,m'} \cdot \nabla \rho(\mathbf{r}) - \sum_{\sigma \geq \sigma'} \frac{\partial f}{\partial G^{\sigma\sigma'}} \nabla^{2} \rho(\mathbf{r}).$$
(5.11)

To continue we now have to find the analytic form of the terms involving G. $\nabla G^{\text{mm'}}$ is a vector quantity which we define as

$$\boldsymbol{\nabla} G^{\mathrm{mm}'} = \left(G_x^{\mathrm{mm}'}, G_y^{\mathrm{mm}'}, G_z^{\mathrm{mm}'} \right), \tag{5.12}$$

and $G^{\rm mm'}$ can be expressed as, using Equation (5.3),

$$G^{\mathrm{mm'}} = \frac{\partial \rho^{\mathrm{m}}}{\partial x} \frac{\partial \rho^{\mathrm{m'}}}{\partial x} + \frac{\partial \rho^{\mathrm{m}}}{\partial y} \frac{\partial \rho^{\mathrm{m'}}}{\partial y} + \frac{\partial \rho^{\mathrm{m}}}{\partial z} \frac{\partial \rho^{\mathrm{m'}}}{\partial z}.$$
 (5.13)

So, for example, $G_x^{\rm mm'}$ can be written as

$$G_x^{\mathrm{mm'}} = \frac{\partial^2 \rho^{\mathrm{m}}}{\partial x \partial x} \frac{\partial \rho^{\mathrm{m'}}}{\partial x} + \frac{\partial \rho^{\mathrm{m}}}{\partial x} \frac{\partial^2 \rho^{\mathrm{m'}}}{\partial x \partial x} + \frac{\partial^2 \rho^{\mathrm{m}}}{\partial x \partial y} \frac{\partial \rho^{\mathrm{m'}}}{\partial y} - \frac{\partial \rho^{\mathrm{m}}}{\partial x} \frac{\partial^2 \rho^{\mathrm{m'}}}{\partial x \partial y} + \frac{\partial^2 \rho^{\mathrm{m}}}{\partial x \partial z} \frac{\partial \rho^{\mathrm{m'}}}{\partial x} + \frac{\partial \rho^{\mathrm{m}}}{\partial x} \frac{\partial^2 \rho^{\mathrm{m'}}}{\partial x \partial z}.$$
(5.14)

One again, to de-clutter the expressions we use the following notation for the above expression

$$G_x^{\rm mm'} = \rho_{xx}^{\rm m} \rho_x^{\rm m'} + \rho_x^{\rm m} \rho_{xx}^{\rm m'} + \rho_{xy}^{\rm m} \rho_y^{\rm m'} + \rho_y^{\rm m} \rho_{xy}^{\rm m'} + \rho_{xz}^{\rm m} \rho_z^{\rm m'} + \rho_z^{\rm m} \rho_{xz}^{\rm m'}.$$
 (5.15)

We can write every element of $\nabla G^{\rm mm'}$ as

$$G_{i}^{\rm mm'} = \sum_{j} \rho_{ij}^{\rm m} \rho_{j}^{\rm m'} + \sum_{j} \rho_{i}^{\rm m} \rho_{ij}^{\rm m'}$$
(5.16)

where the sum is over all co-ordinate basis vectors. Therefore,

$$\boldsymbol{\nabla} G^{\mathbf{m},\mathbf{m}'} \cdot \boldsymbol{\nabla} \rho = \sum_{ij} \rho_i \rho_{ij}^{\mathbf{m}} \rho_j^{\mathbf{m}'} + \sum_{ij} \rho_i \rho_{ij}^{\mathbf{m}'} \rho_i^{\mathbf{m}}.$$
 (5.17)

So the potential is now given by, suppressing explicit notational position dependence,

$$\upsilon_{F} = \frac{\partial f}{\partial \rho} - \sum_{\sigma \geq \sigma'} \frac{\partial^{2} f}{\partial G^{\sigma \sigma'} \partial \rho} |\nabla \rho|^{2}$$
$$- \sum_{\sigma \geq \sigma'} \sum_{m \geq m'} \sum_{ij} \frac{\partial^{2} f}{\partial G^{\sigma \sigma'} \partial G^{m,m'}} \left(\rho_{i} \rho_{ij}^{m} \rho_{j}^{m'} + \rho_{i} \rho_{ij}^{m'} \rho_{i}^{m} \right) \qquad (5.18)$$
$$- \sum_{\sigma \geq \sigma'} \frac{\partial f}{\partial G^{\sigma \sigma'}} \nabla^{2} \rho.$$

For convenience let us write this as the following, with the summations implied

$$\upsilon_F = \frac{\partial f}{\partial \rho} - \frac{\partial^2 f}{\partial G^{\sigma\sigma'} \partial \rho} |\boldsymbol{\nabla}\rho|^2 - \frac{\partial^2 f}{\partial G^{\sigma\sigma'} \partial G^{\mathrm{mm'}}} \left(\rho_i \rho_{ij}^{\mathrm{m}} \rho_j^{\mathrm{m'}} + \rho_i \rho_{ij}^{\mathrm{m'}} \rho_i^{\mathrm{m}}\right) - \frac{\partial f}{\partial G^{\sigma\sigma'}} \boldsymbol{\nabla}^2 \rho.$$
(5.19)

In the spin unpolarised framework the total density, ρ , is given by $\rho = \rho^{m} + \rho^{m'}$ and $\rho^{m} = \rho^{m'}$ and so, as a function the total density the potential is given by

$$\upsilon_F = \frac{\partial f}{\partial \rho} - \frac{\partial^2 f}{\partial G^{\sigma\sigma'} \partial \rho} |\boldsymbol{\nabla}\rho|^2 - \frac{1}{2} \frac{\partial^2 f}{\partial G^{\sigma\sigma'} \partial G^{\mathrm{mm'}}} \rho_i \rho_{ij} \rho_j - \frac{\partial f}{\partial G^{\sigma\sigma'}} \boldsymbol{\nabla}^2 \rho. \quad (5.20)$$

XCFun returns values for all of the partial derivatives of the integrand we require at every grid point using automatic differentiation techniques, from which Equation (5.20) can be constructed.

5.2. The Basis Set Dependence of Potential Terms

In order to study whether the chemical potential is a constant throughout space when using finite Gaussian basis sets we created four even tempered basis sets. The basis sets, centred on nuclei A, are

$$B1 = \left\{ e^{-3^{m_i}(\mathbf{r} - \mathbf{r}_A)^2} | m_i \in [-4, -2] \right\}$$

$$B2 = \left\{ e^{-3^{m_i}(\mathbf{r} - \mathbf{r}_A)^2} | m_i \in [-4, 2] \right\}$$

$$B3 = \left\{ e^{-3^{m_i}(\mathbf{r} - \mathbf{r}_A)^2} | m_i \in [-4, 6] \right\}$$

$$B4 = \left\{ e^{-3^{m_i}(\mathbf{r} - \mathbf{r}_A)^2} | m_i \in [-4, 14] \right\}.$$

(5.21)

The value of m_i controls how sharp the basis function is. The greater the value of m_i the sharper the function is. This is exhibited in Figure 5.1. Therefore, when going from B1 to B4 we are increasing the sharpness of the basis.

We have explored how the spatial behaviour of the potentials is affected by the choice of Gaussian basis. To do this we optimised the Lagrangian for the He atom using the VW OF-KEF and the PBE XC functionals using the CCH scheme described in the previous chapter. Its important to note that the VW is exact for the He atom and so, any inaccuracies in $v_{\rm Ts}$ is due to the choice of basis.



Figure 5.1.: Plots of $\eta(r) = \exp[-mr^2]$ for a range of m values



Figure 5.2.: The Coulombic potential $v_{\rm J}(\mathbf{r}) = \delta E_{\rm J}/\delta \rho$ for a ground state density of the He atom expanded in different basis sets. Functional details given in the text.

In Figure 5.2 we have plotted the Coulomb potential, $v_{\rm J}$, evaluated on the minimising density using various basis functions. Its explicit form is given by

$$\upsilon_{\mathrm{J}}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$
 (5.22)

It is clear from the figure $v_{\rm J}$ is relatively insensitive to the choice of basis.

In Figure 5.3 we have plotted the von-Weisäcker potential, v_{Ts} evaluated on the minimising density constructed from various basis sets. Its explicit form is

$$\upsilon_{\rm Ts}(\mathbf{r}) = \frac{|\boldsymbol{\nabla}\rho|^2}{8\rho^2} - \frac{\boldsymbol{\nabla}^2\rho}{4\rho}.$$
 (5.23)

As one can see in Figure 5.3 there are significant differences between B1 and the rest. The basis set B2 has large discrepancies near the nuclei but is close

M. S. Ryley Variational Solutions in OF-DFT 117



Figure 5.3.: The von-Weisäcker kinetic potential $v_{\rm Ts}(\mathbf{r}) = \delta E_{\rm Ts}/\delta\rho$ for the ground state density of the He atom expanded in different basis sets. Functional details given in the text.



Figure 5.4.: The PBE XC potential $v_{\rm XC}(\mathbf{r}) = \delta E_{\rm XC}/\delta\rho$ for the ground state density of the He atom expanded in different basis sets. Functional details given in the text.

to the B3 and B4 sets in the tail regions. B3 and B4 are indistinguishable on this scale. It is clear that there is a convergence in the behaviour of $v_{\rm Ts}$ as one approaches the basis set limit. Its clear that one needs sharp functions to accurately describe the potential in the core region.

In Figure 5.4 we have plotted the PBE XC potential, $v_{\rm XC}$ evaluated on the minimising density constructed from various basis sets. For the PBE XC functional $v_{\rm XC}$ is just the sum of the PBE X potential $v_{\rm X}$ and the PBE C potential $v_{\rm C}$. For reference these terms are

$$\upsilon_{\rm X} = \frac{4}{3} C_{\rm PBEX} \rho^{1/3} \bigg[F_{\rm X} + \frac{\mu s}{(1 + \mu s^2 \kappa^{-2})^2} \bigg(\frac{3}{4C_{\rm PBEX} \rho^{1/3}} \frac{\mathrm{d}|\boldsymbol{\nabla}\rho|}{\mathrm{d}\rho} - 2s \bigg) \bigg]$$
(5.24)



Figure 5.5.: The external potential $v_{\text{ext}}(\mathbf{r})$ for the ground state density of the He atom expanded in different basis sets. Functional details given in the text

and

$$v_{\rm C} = \varepsilon_{\rm c}^{\rm unif} + H + \rho \frac{\partial \varepsilon_{\rm c}^{\rm unif}}{\partial \rho} + \left[\frac{\gamma}{(\gamma(1 + At^2 + A^2t^4) + \beta t^2(At^2 + 1))(1 + At^2 + A^2t^4)} \right] \times \left[\frac{\mathrm{d}|\boldsymbol{\nabla}\rho|}{\mathrm{d}\rho} \frac{\beta t}{k_{\rm s}} (1 + 2At^2) - \frac{7\beta t^2}{3} (1 + 2At^2) - (At^2)^3 \rho(At^2 + 2) \exp\left(-\varepsilon_{\rm c}^{\rm unif}/\gamma\right) \frac{\partial \varepsilon_{\rm c}^{\rm unif}}{\partial \rho} \right]$$
(5.25)

with

$$\frac{\partial \varepsilon_{\rm c}^{\rm unif}}{\partial \rho} = -\frac{r_{\rm s}}{3\rho} \bigg[-2a\alpha_1 \ln(1+\gamma^{-1}) + \frac{(2a(1+\alpha_1 r_{\rm s})(a(\beta_1 r_{\rm s}^{-1/2}+2\beta_2+3\beta_3 r_{\rm s}^{1/2}+4\beta_4 r_{\rm s}^2))}{\gamma(\gamma+1)} \bigg].$$
(5.26)

One can refer to Section 2.4.2 for the exact values of all constants in these expressions.

In Figure 5.5 we have plotted the external potential, v_{ext} evaluated on the minimising density constructed from various basis sets. This potential is independent of the ground state density and as one can see in Figure 5.5 is unaffected by the choice of basis.

In Figure 5.6 we have plotted the various $v_{\text{Euler}} = v_{\text{Ts}} + v_{\text{J}} + v_{\text{XC}} + v_{\text{ext}}$ for the range of basis sets we have been using. Due to the fact the $v_{\text{Euler}} = \mu$



Figure 5.6.: $v_{\text{Euler}} = v_{\text{ext}} + v_{\text{XC}} + v_{\text{Ts}} + v_{\text{J}}$ plotted using a range of basis sets

at the minimising density it is clear from these graphs that v_{Euler} is not a constant over all space and tends to μ only in the basis set limit. In other words Equation (5.1) is only valid in the basis set limit when one uses Gaussian basis sets. In the CCH scheme we reach a zero gradient because v_{Euler} oscillates over the region spanned by $\rho^{1/2}$. Furthermore, the value of μ one obtains from these calculations is the average of v_{Euler} over that particular region.

5.2.1. The Implication for Computing Molecular Forces

The above discussion is important for future work when one would wish to use the CCH scheme to run geometry optimisations. In these calculations we require the force at each step in the calculation to compute a new guess at the ground state geometry of the system. In wavefunction based methods and in KS-DFT we expand the wavefunction Ψ in a finite basis which means the force, F, is given by

$$-F = \left\langle \Psi \right| \frac{\partial H}{\partial \mathbf{R}} \left| \Psi \right\rangle + 2 \left\langle \frac{\partial \Psi}{\partial \mathbf{R}} \right| H \left| \Psi \right\rangle$$
(5.27)

where H is the Hamiltonian of a given theory and \mathbf{R} is the molecular geometry. The first term in Equation (5.27) is what we expect from the Hellmann-Feynman theorem and the second term in Equation (5.27) is known as the Pulay force. The Pulay force is zero when one uses a complete basis or basis functions that are not atom centred. In OF-DFT we are not expanding the wavefunction but $\rho^{1/2}$. To compute the force we note that the total energy of the system can be expressed as

$$E_{\rm tot}(\rho) = E(\rho) + E_{\rm ion} \tag{5.28}$$



Figure 5.7.: v_{Euler} using the B4 basis and the final chemical potential, $\mu = -0.5792$

where E_{ion} is the Coulombic repulsion due to the nuclei. We will be most interested in how the electronic interactions contribute to the force. Consider a two nuclei system we wish to compute the electronic contribution to the force on nuclei A with charge Z_A and a position vector \mathbf{R}_A . This is computed through

$$-\frac{\mathrm{d}E}{\mathrm{d}\mathbf{R}_A} = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\frac{Z_A(\mathbf{r}-\mathbf{R}_A)}{|\mathbf{R}_A-\mathbf{r}|^3} - \int \mathrm{d}\mathbf{r}\frac{\delta E}{\delta\rho(\mathbf{r})}\frac{\partial\rho(\mathbf{r})}{\partial\mathbf{R}_A}.$$
(5.29)

The corresponding Euler-Lagrange equations for the optimisation problem state that $\frac{\delta E}{\delta \rho} = \mu$, where μ a fixed number throughout space. Therefore, the total force on atom A is given by

$$F_A = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \frac{Z_A(\mathbf{r} - \mathbf{R}_A)}{|\mathbf{R}_A - \mathbf{r}|^3} - \frac{\mathrm{d}E_{\mathrm{ion}}}{\mathrm{d}\mathbf{R}_A}$$
(5.30)

This equation is just what we expect from Hellmann-Feynman theory so it seems there are no Pulay forces in OF-DFT. However, this is only true if one uses a non-atom centred basis sets — for example, plane wave basis sets. However, we have just seen that using finite atom centred basis sets, such as Gaussian basis sets, our assumption that μ is constant over space is not true. Therefore, when computing forces in OF-DFT using such basis sets we must include a Pulay like term of the form

$$-\int d\mathbf{r}\mu(\mathbf{r})\frac{\partial\rho}{\partial\mathbf{R}} = -\int d\mathbf{r}[\upsilon_{\text{ext}} + \upsilon_{\text{XC}} + \upsilon_{\text{Ts}} + \upsilon_{\text{J}}](\mathbf{r})\frac{\partial\rho(\mathbf{r})}{\partial\mathbf{R}}.$$
 (5.31)

In the optimisation method described in this chapter μ was the Lagrange multiplier which is a number not a function. So is there a clear relationship



Figure 5.8.: The radial atomic densities for the Neon atom. The accurate density is computed using CCSD(T).

between the Lagrange multiplier we converge to in the bisections and v_{euler} when one is using a finite Gaussian basis? In Figure 5.7 we have plotted v_{euler} using the B4 basis and shown the final converged value of the Lagrange multiplier. It is clear that v_{euler} oscillates around the origin and further analysis has shown that the average function value of v_{euler} equals μ . Future work will be focussed on creating a geometry optimisation algorithm for OF-DFT using the potential code we have introduced into the QUEST code and bearing in mind the basis set dependence of the potentials spatial behaviour.

5.3. The Effect of Potentials on the Density

The Euler equation can be expressed in the following form

$$\frac{\delta T_{\rm s}}{\delta \rho} = -\upsilon_{\rm s} + \mu \tag{5.32}$$

where the effective potential v_s is comprised of the Coulomb, external and XC potential. Throughout this section we will be using the PBE XC functional. As μ should just be a number this means that the kinetic and effective potentials must balance precisely. The self consistent procedure leads to densities that are determined by this balance. Many failures of approximate OF-KEFs can be understood in terms of this balancing act. In Figure 5.8 we have plotted the radial atomic density computed using TFVW and TF02W models for the Ne atom. We have presented the accurate density profile using the CCSD(T)/aug-cc-pVTZ density. For all orbital dependent quantities —including KS-DFT calculations — in this section were generated using aug-cc-pVTZ.^{121,122} For all orbital free quantities shown — including in Figure 5.8 — in this section were generated using the ETOF basis. For all density functional calculations the PBE XC are used. In Figure 5.8 we can see that decreasing the amount of $T_{\rm VW}$ in the OF-KEF shifts the density towards the nucleus in the core region. This is usually explained by the argument that the $T_{\rm TF}$ atomic densities are divergent at the nucleus and the $T_{\rm VW}$ dampens this. Hence reducing the proportion of $T_{\rm VW}$ will force the density to accumulate around the nucleus. We can also see from Figure 5.8 that both models yield a density that is too diffuse in the tail regions. In this section we will explore if there is any connection of these density profiles to the potentials in the corresponding Euler equations.

In KS-DFT one computes the KE exactly via

$$T_{\rm s}[\{\phi_i\}] = -\frac{1}{2} \sum_i \langle \phi_i | \boldsymbol{\nabla}^2 | \phi_i \rangle \qquad (5.33)$$

where ϕ_i are the KS-orbitals arising from the KS equations. Starting from this and the Euler equation, Equation (5.32), King and Handy¹²⁹ derived an equation to compute the kinetic potential v_{Ts} exactly. Its explicit form is

$$\upsilon_{\rm Ts}(\mathbf{r}) = \frac{\sum_i \left[-\frac{1}{2} \phi_i(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) - \epsilon_i \phi_i^2(\mathbf{r}) \right]}{\rho(\mathbf{r})} + \mu.$$
(5.34)

In Equation (5.34) ϵ_i is the energy of ϕ_i . We also know that asymptotically $v_{\rm Ts} - v_{\rm s} = \mu$. To demonstrate this in Figure 5.9 we have plotted Equation (5.34) and an accurate $v_{\rm s}$ for the Neon atom. To compute $v_{\rm s}$ we have used the equations and implementations discussed in Section 5.1. From Figure 5.9 we can see that in the asymptotic regions the differences between the two potentials one predicts the ionisation to be 0.59 Hartree. The is slightly lower than the experimental value for the ionisation energy and can be put down to the inaccuracy in the PBE XC functional. For reference the experimental value for the first ionisation is 0.79 Hartree.¹³⁰

In Figures 5.10 and 5.11 we see that the predicted ionisation energies are too small. It seems that as one decreases the amount of $T_{\rm VW}$ included in the functional one decreases $|\mu|$ this suggests that the amount of $T_{\rm VW}$ in the TFW type models controls the predicted ionisation energies. Indeed it is known⁸ that in exact TF theory the predicted chemical potential is zero. The small magnitude of the chemical potential is a symptom of every OF-KEF discussed in this work. One can now rationalise this by the in-



Figure 5.9.: The King Handy kinetic potential and the v_s . The difference between the potentials asymptotically is $\mu = -0.59E_{\rm h}$



Figure 5.10.: The ground state potential generated by TFVW using the PBE XC and the corresponding v_s . The difference between the potentials asymptotically is $\mu = -0.2E_{\rm H}$

accuracies in how the kinetic potentials and the rest of the Euler equation interact. It is interesting that scaling the contribution to the $T_{\rm VW}$ in these OF-KEFs is an easy way to correct for errors in the chemical potential. By comparing Figure 5.9 with Figures 5.10 and 5.11 we see that the potentials decay far too rapidly. The symptom of this is that the density leaks out to large r because the potential is not strong enough to overcome the energy required to stop this leaking and so, decays to zero too rapidly. We now investigate the possible link between potentials and the observation in Figure 5.8 that the density of more compact for TF02W near the nucleus compared to the TFVW functional. In Figure 5.12 we have plotted the



Figure 5.11.: The ground state potential generated by TF02W using the PBE XC and the corresponding v_s . The difference between the potentials asymptotically is $\mu = -0.1E_{\rm H}$



Figure 5.12.: The kinetic potential error, Δv_{Ts} , for the TFVW and TF02W models for the neon atom.

kinetic potential error, $\Delta v_{\rm Ts}$,

$$\Delta v_{\rm Ts} = v_{\rm Ts}^{\rm model} - v_{\rm Ts}^{\rm ref} \tag{5.35}$$

where $v_{\text{Ts}}^{\text{ref}}$ is Equation (5.34) and $v_{\text{Ts}}^{\text{model}}$ is the potential defined by the OF-KEFs. In the figure both functionals potentials are too positive near the nucleus and so, there is a lower density than expected which can be seen in Figure 5.8. In the vicinity of the nucleus the $T_{\text{TF}} + 0.2T_{\text{VW}}$ is better and so the density is more compact in this region. Therefore, there could be a promising avenue in developing OF-KEFs with correct potential radial dependency which could in turn lead to more accurate densities.

5.4. Conclusions

In this chapter we have investigated the potentials of the Euler equation — Equation (5.32) — involved in the optimisations at the heart of this thesis. We have implemented methods to compute these potentials into QUEST using an interface with XCFun which discussed in Section 5.1. We discovered in Section 5.2 that having an accurate basis set is important for computing and using potentials in applications and also in the balancing of the potentials of the Euler equation. In Section 5.3 we continued the investigation into the balance of potentials in the Euler equation and investigated the impact the potentials generated from OF-KEFs on the variational solutions. We found that the chemical potential was affected by the balance and demonstrated that, for atoms at least, the topological features of the radial density are impacted by the kinetic potential. We hope that starting from accurate approximate kinetic potentials will be an alternative angle for developing OF-KEFs. However, one would have to consider how sensitive this accuracy is to the choice of XC functional. Because as we have seen in this chapter the interaction of the potentials in the Euler equation determines how accurate the density is.

6. TRIM Scheme

6.1. Introduction

Developing electronic structure methods with low computational cost is an important strand of modern method development. Orbital-free density functional theory (OF-DFT) is, in principle at least, a linear scaling *ab-inito* quantum mechanical method which has the potential to be used in purely quantum mechanical simulations of very large systems. In Chapter 4 we discussed our work which extended the range of orbital free non-interacting kinetic energy functionals (OF-KEFs) tested by optimising the corresponding Lagrangian. We found that post Kohn Sham analysis (PKSA) is inadequate in assessing the accuracy of OF-KEFs due to the large errors in the densities resulting from the optimisations using OF-KEFs.

We have demonstrated the need to solve the optimisation problem defined by the OF-KEFs before making claims about their accuracy. However, the extended CCH scheme presented in Chapter 4 was not an efficient way of solving the optimisation problems because we were optimising the density and the chemical potential *separately*, leading to an inefficient solver. The reason for this is that the CCH scheme is avoiding a saddle point optimisation using a nested iteration solver, with a higher overhead. In this chapter we lay out a new way of optimising the Lagrangian defined by any OF-KEF by conducting a saddle point optimisation. Therefore, this new scheme optimises the density and the chemical potential *simultaneously* and we demonstrate the dramatic improvement in efficiency compared to our previous scheme.

In Section 6.2.1 we give a brief review of OF-DFT and restate what we showed in Section 3.4.2 that the target functional in the optimisation problems in OF-DFT is a saddle function. We then outline in Section 6.2.2 the optimisation scheme we have developed⁴ to solve the Euler equation, which is based on the Trust-Region Image Method (TRIM).^{131–133} Finally in Section 6.3 we present a convergence illustration demonstrating why the scheme presented in this paper is more efficient than our previous work. We compare TRIM with both CCH and Lopez-Acevedo. We also discuss how the starting guess effects the convergence properties of the TRIM scheme. The basis functions used for the density are the same as discussed in Chapter 4.

6.2. Theory

6.2.1. OF-DFT Variational Recap

There are two ansatze in OF-DFT. The first is that we can partition the energy functional, the existence of which was guaranteed by Hohenberg and Kohn, in the following manner:

$$E(\rho) = T_{\rm s}(\rho) + E_{\rm J}(\rho) + E_{\rm XC}(\rho) + E_{\rm ne}(\rho).$$
(6.1)

Where: $T_{\rm s}(\rho)$ is the KEF, $E_{\rm J}(\rho)$ is the Coulomb (Hartree) energy due to the electron density repulsion with itself, $E_{\rm XC}(\rho)$ is the exchange-correlation energy and $E_{\rm ne}(\rho)$ is the energy due to the interaction between the nuclei and the electron density. The second ansatz is that all these terms should be written as explicit functionals of the electronic density, ρ . This means the first term in Eq. (6.1) is is known as the OF-KEF because this differs to Kohn-Sham DFT (KS-DFT) where the KEF is treated exactly using orbitals. Whilst removal of the orbitals leads to the difficult challenge of accurately treating the KE, by doing so however we arrive at a method which has a much lower cost, in theory. The reason is that the stationary point with respect to variations in ρ , under the constraint $\int \rho d\mathbf{r} = N$ where N is the number of electrons in the system, leads to a *single* non-linear equation - the Euler equation

$$\frac{\delta T_{\rm s}}{\delta \rho} + \frac{\delta E_{\rm J}}{\delta \rho} + \frac{\delta E_{\rm XC}}{\delta \rho} + \frac{\delta E_{\rm ne}}{\delta \rho} = \mu, \qquad (6.2)$$

where μ is the Lagrange multiplier enforcing the constraint on the minimising ρ .

The Euler equation can be thought as the following optimisation problem

$$E_{\rm GS}(\rho) = \inf_{\rho \in \mathcal{X}} \sup_{\mu \in \mathcal{R}} \left[E(\rho) - \mu \left(\int \rho d\mathbf{r} - N \right) \right]$$
(6.3)

with $E(\rho)$ as defined in Eq. (6.1). This unconstrained Hohenberg-Kohn minimax variation principle corresponds to Lagrange's method for the constrained problem, with multiplier μ . See Section 3.4.2 for details. There-
fore, we expect that the CCH and TRIM schemes converge to the same point. We do indeed find that this is the case. Therefore, if one wants to find the ground state energy, $E_{\rm GS}$, the ground state density and chemical potential of a given system in OF-DFT one can find the first-order saddle point of the surface defined by the target functional in Eq. (6.3). This is the idea behind the TRIM scheme.

6.2.2. A Second Order Optimization Scheme

In this section it will be convenient to define the following function, \mathcal{L} : $\mathcal{D} \times \mathcal{R} \mapsto cl(\mathcal{R})$,

$$\mathcal{L}(\rho,\mu) = E(\rho) - \mu \left(\int \rho d\mathbf{r} - N\right).$$
(6.4)

This is different to $\mathcal{L}(\rho; \mu)$ in the CCH scheme as we are allowing μ to be varied whereas in the CCH scheme μ is kept fixed. We can rephrase the optimisation problem as

$$E_{\rm GS}(\rho) = \inf_{\rho \in \mathcal{D}} \sup_{\mu \in \mathcal{R}} \mathcal{L}(\rho, \mu).$$
(6.5)

The infimum is conducted over set of densities, \mathcal{D} , defined by

$$\mathcal{D} = \left\{ \mathcal{X} | \rho \ge 0, \rho^{\frac{1}{2}} \in \mathcal{H}_1^1 \right\}$$
(6.6)

where \mathcal{H}_1^1 is a first order Sobolev space and ensures that the kinetic energy of any $\rho \in \mathcal{D}$ is finite. The easiest way to ensure this infimum is conducted inside this set is to expand the square root of the density in a finite set of real functions, $\{\eta_i\}$

$$\rho^{\frac{1}{2}}(\mathbf{r}) = \sum_{i} c_{i} \eta_{i}(\mathbf{r}) \tag{6.7}$$

where the $\{c_i\}$ are the basis set expansion coefficients.

The first-order saddle point of \mathcal{L} for a general OF-KEFs is usually not possible to find using analytical methods. Therefore, we will be finding this first-order saddle point using numerical optimisation techniques. We use the trust-region image method¹³¹ (TRIM) which was originally developed in order to find transition states on potential energy surfaces.^{132,133}

In the majority of numerical minimisation techniques we construct a model of the surface about the current guess of the target (minimising) point of the surface, x^* . We label this model by $m(x_k)$ where x_k is the k-th guess at the minimiser. The optimisation continues an iterative manner until it is no longer possible to find a step which goes to the minimum of the model m. In other words m must be constructed so that its gradient is zero at x^* and its Hessian is positive definite at x^* .

We can usually partition the optimisation into two regions based on the accuracy of the model. If we are in the *local region* then we can expect the model $m(x_k)$ to accurately represent the surface around some small neighbourhood of x^* e.g. some open ball centred, $B_{\delta}(x^*)$. Once the optimisation enters $B_{\delta}(x^*)$ it will converge quickly to the solution point as the gradient and Hessian of the model is close to the corresponding quantities of the target surface. When this is not the case we say we are in the *global region*. Almost all optimisations start in the global region, which is trickier for optimisation algorithms than the local region where all successful optimisations will end. The reason for why this global region is difficult for optimisation algorithms is that the step generated from the model about a point does not necessarily point towards the minimum of the surface. Therefore, we chose to use two different algorithms in the TRIM scheme. One for when the optimiser is in the local region and one for when its in the global region. The TRIM scheme consists of a switch between the two.

We begin by discussing the optimisation algorithm used in local regions. This algorithm is the famous Newton method. The idea behind the Newton method is that one should expand the surface to second order in displacements, s, from the current point x_k

$$m(s) = f(x_k) + g_k^T s + \frac{1}{2} s^T H_k s$$
(6.8)

where g_k is the gradient of f evaluated at x_k and H_k is the Hessian of f evaluated at x_k . The Newton method searches for a minimum of this model and so finds the stationary point of m(s) which is given through

$$\frac{\partial m}{\partial s} = 0 \implies s = -H_k^{-1}g_k. \tag{6.9}$$

Note that for this point to be a minimum the Hessian needs to be positive definite. Furthermore, to find s the Hessian must be invertible i.e. it cannot have zero eigenvalues. If this is not the case simple regularisation techniques such as identity shift can resolve this. The identity shift is just a small shift of the eigenvalues of a matrix.

This step is then used to calculate the next approximations to x^* by $x_{k+1} = x_k + s$. This is then iterated through until $x_{k+1} = x_k$ or equivalently

s = 0. It is important to note that if $\lim_{k\to\infty} x_k = x^*$ then the Newton method converges quadratically.

To prove this assume $x_{k+1} \to x^*$ and define the error after the k-th iteration through $\epsilon_k = x_k - x^*$. The expansion of $f(x_k)$ around x^* yields

$$G_k \approx G_* + H_* \epsilon_k \tag{6.10}$$

where G_* and H_* are the gradient and Hessian evaluated at x^* . However, by definition

$$\epsilon_{k+1} = \epsilon_k + (x_{k+1} - x_k)$$

$$= \epsilon_k - \frac{f(x_k)}{G_k}$$

$$\approx \epsilon_k - \frac{G_* \epsilon_k + \frac{1}{2} H_* \epsilon_k^2}{G_* + H_* \epsilon_k}.$$
(6.11)

Using the following second order expansion

$$\frac{\alpha\epsilon + \frac{1}{2}\beta\epsilon^2 + \gamma\epsilon^3}{\alpha + \beta\epsilon + \zeta\epsilon^2} \approx \epsilon - \frac{\beta}{2\alpha}\epsilon^2, \tag{6.12}$$

then one arrives at

$$\epsilon_{k+1} = \frac{H_*}{2G_*} \epsilon_k^2. \tag{6.13}$$

Hence Newton's method converges quadratically. This is important for the TRIM scheme as it means this scheme will converge quadratically once it nears the solution point as the TRIM scheme reduces to a Newton step method once we enter the local region. This will be explained later. Compare that to the CCH scheme where we saw that convergence slows down as one nears the solution point due to the reliance on the bisection method

The Newton method becomes unstable in the global region of the optimisation for the obvious reason that the Taylor expansion around the current point is no longer representative of the surface around the optimiser. Indeed this how one practically defines the global region — it is the set of all points where the Newton step does not decrease the value of the target surface.

A simple demonstration of this can be seen in the maximisation of a Gaussian function

$$f(x) = \exp\left(-x^2\right),\tag{6.14}$$

which one can find by analytic methods to be at $x^* = 0$. Whilst the above discussion concentrated on minimisation it's trivial to adapt the Newton



Figure 6.1.: A plot of $f(x) = \exp(-x^2)$ (black line), x_{k+1} (red dotted line) and x_k (blue dashed line)

method to find maxima rather than minima. The Newton step is fairly easy to calculate given

$$g_k = -2x_k \exp\left(-x_k^2\right) \tag{6.15}$$

and

$$H_k = \exp(-x_k^2) \left[4x_k^2 - 2\right].$$
 (6.16)

Therefore, the Newton step computed at point x_k is given by

$$s = \frac{x_k}{2x_k^2 - 1} \tag{6.17}$$

and the new point by

$$x_{k+1} = \frac{2x_k^3}{2x_k^2 - 1}.$$
(6.18)

We see in Fig. 6.1 that the Newton step converges only for points $x_k < 0.5$ For larger values the method diverges and for $x_k = 2^{-1/2}$ it gives infinite steps. This is a classic example of an optimisation problem having local and global regions.¹³⁴

The Newton step works well in the local region but what algorithm should we use for the global region? We have chosen the trust region method. In the trust region we accept that our model can only accurately represent the true surface in a small region — trust region — around the current point. This trust region is a hypersurface and in the trust region method we take a step to the minimiser of the boundary points of this hypersurface if the Newton step fails. When the true minimizer of the target surface is inside this trust region this method becomes the Newton method.



Figure 6.2.: $r(\omega)$ as defined in Eq. (6.21), blue solid line. The black dotted line corresponds to a choice of trust radius. Note that for every trust radius there exists at least two minimising values of step.

We will now move on to the details of the trust region method. We assume our trust region is a hypersphere of radius r. Therefore, our step, s, should satisfy $|s| \leq |r|$. We once again expand the surface to second order in displacements but when we compute the minimising step we must ensure $|s| \leq |r|$. We can achieve this by minimising the following Lagrangian

$$\mathcal{L}(s,\omega) = m(s) - \frac{1}{2}\omega(s^T s - r^2)$$
(6.19)

where m(s) is defined in Eq. (6.8) and $\omega \in \mathcal{R}$ is a Lagrange multiplier. Minimising \mathcal{L} with respect to all possible displacements yields

$$\frac{\partial \mathcal{L}}{\partial s} = 0 \Longrightarrow s(\mu) = -(H_k - \omega \mathbf{1})^{-1} g_k.$$
(6.20)

We see that ω is a 'level shift parameter' in that it shifts the eigenvalues of H_k downwards by ω . This parameter is chosen such that we always step to the boundary given through

$$r(\omega) = \sqrt{g_k^T (H_k - \omega \mathbf{1})^{-2} g_k}.$$
(6.21)

In Figure 6.2 we have plotted Equation (6.21) and chosen a trust radius of 16. What this figure shows is that $r(\omega)$ diverges when μ is equal to an eigenvalue of the Hessian. To find the value of μ which corresponds to the minimum boundary of the trust region one finds the intersection of $r(\mu)$ with this boundary. But as Figure 6.2 shows there exists at least two intersections for every choice of trust radius. Note that in Figure 6.2 it seems that the dotted red line at r = 3 does not intersect with Equation (6.21) but they do intersect twice once at $+\infty$ and once at $-\infty$. This means that for this small trust-region the Newton method will work. Figure 6.2 also shows how to pick the value of μ which corresponds to the minimum at the boundary. We simply pick the smallest value of ω for which Equation (6.21) intersects with the trust-region boundary.

The strategy of the trust radius method now follows the obvious steps. Firstly one calculates at each iteration the Hessian and gradient. Using this information we compute the step s. When we are in local region, $|s| \leq |r|$ we take the step

$$s(0) = -H_k^{-1}g_k \tag{6.22}$$

which is just the Newton step. When we are in the global region, |s| > |r|we take the step

$$s(\omega) = -(H_k - \omega \mathbf{1})^{-1} g_k \tag{6.23}$$

where the level shift parameter is chosen such that $s(\omega)$ corresponds to taking a step to the minimiser on the boundary of the trust region. As we have seen in practice that involves selecting a value of ω smaller than the lowest eigenvalue of the Hessian and it must be negative to ensure this step is not longer than the Newton step. It is common in practical implementations such as the one we have implemented into QUEST one uses either bisection or linear interpolations to find the numerical value of ω which yields $|s(\omega)| = |r|$. This normally requires 3-4 iterations and we preform the bisection or linear interpolation between zero and the lowest eigenvalue of the Hessian.

There is an important part of the trust region algorithm which we have not yet been discussed. It is clear that as we move from the starting guess we should improve our confidence in our local model. This can be achieved by increasing our trust radius. Indeed the size of the trust radius can be interpreted as a measure of confidence in our local second-order Taylor expansion. So between iterations one would like to be able to update the trust radius if we believe our model is more accurate than we believed it to be. To do this we compute the following ratio

$$T = \frac{f(x_{k+1}) - f(x_k)}{g_k^T s + \frac{1}{2} s^T H_k s}$$
(6.24)

If T is close to one then we increase the trust radius if it is much less than one we reduce it. The increase and decrease factors are arrived at by experimenting with various atomic and molecular calculations. We have found that beginning with a trust radius of 0.4 and having an increase factor of 2.0 and a decrease divisor of 1.2.

So far we have discussed minimisations. But we know from the mathematical framework of DFT that the variation principle at the heart of DFT is a first order saddle point (FOSP) optimisation. However, to utilise the trust-region framework above we have to have a surface we wish to minimise. So the question is: can we identify a new surface which has a oneto-one correspondence with our target surface defined by Equation (6.4) such that this new surface has a minima which coincides with our target surface FOSP?

To investigate this new surface, called the image surface, let us consider the properties of the image surface at the minima. We know that in all directions the gradient of the image surface should be ascending from that point and all eigenvalues of the Hessian matrix should be non negative. Contrast that to the FOSP of the original surface where the gradient should be descending in one direction and has a corresponding negative eigenvalue in the Hessian. It is clear all one needs to do is change the sign of the gradient and Hessian element in the diagonal representation corresponding to the increasing direction of the FOSP. In doing so we have the gradient and Hessian of the image surface. But how does one generate the image surface from the original? The answer is that in general you cannot but thankfully it doesn't matter because the trust-region requires only the gradient and Hessian at each iteration which is easy to generate from the original surface as discussed above. TRIM was first used in quantum chemistry by Helgaker¹³¹ to find the transition states on a potential energy surface.

In the application of TRIM for OF-DFT we allow for variations in the chemical potential, μ , and incorporate this into the gradient and Hessian accordingly. We also must note that for our FOSP the increasing direction is along the μ eigenmode. The only modification of an existing trust-region optimiser is in the computation of the step, p, which one computes through

$$p = -H^{-1}g (6.25)$$

where g is the gradient given by

$$g = \left(\cdots, \frac{\partial \mathcal{L}}{\partial c_i}, \cdots, \frac{\partial \mathcal{L}}{\partial \mu}\right)$$
(6.26)

and ${\cal H}$ is the Hessian

$$H = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \cdots & \frac{\partial^2 \mathcal{L}}{\partial c_i \partial c_j} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial c_i \partial \mu} \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \frac{\partial^2 \mathcal{L}}{\partial \mu \partial c_j} & \cdots & 0 \end{pmatrix}$$
(6.27)

where $0 = \partial^2 \mathcal{L} / \partial \mu^2$. The required modification in order to find the minima of the image surface is simply

$$g = \left(\cdots, \frac{\partial \mathcal{L}}{\partial c_i}, \cdots, -\frac{\partial \mathcal{L}}{\partial \mu}\right)$$
(6.28)

and taking the negative of the eigenvalue of H corresponding to the μ eigenmode. We have derived and implemented the analytic gradient and Hessian. The elements corresponding to the basis coefficients, c_i , are the same as in Chapter 4 but we have included them here to be complete. The elements of the gradient are

$$\frac{\partial \mathcal{L}}{\partial \mu} = N - \int \rho \mathrm{d}\mathbf{r},\tag{6.29}$$

$$\frac{\partial \mathcal{L}}{\partial c_i} = 2 \int \eta_i \left(\frac{\delta E}{\delta \rho} - \mu\right) \rho^{\frac{1}{2}} \mathrm{d}\mathbf{r}.$$
(6.30)

The elements of the Hessian are:

$$\frac{\partial^2 \mathcal{L}}{\partial \mu \partial c_i} = -2 \int \eta_i \rho^{\frac{1}{2}} \mathrm{d}\mathbf{r}, \qquad (6.31)$$

and

$$\frac{\partial^{2} \mathcal{L}}{\partial c_{i} \partial c_{j}} = \frac{\partial^{2} T_{s}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{XC}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{J}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{ne}}{\partial c_{i} \partial c_{j}} - 2\mu \int \eta_{i} \eta_{j} d\mathbf{r}.$$
(6.32)

The fourth term in Eq. (6.32) evaluates to

$$\frac{\partial^2 E_{\rm ne}}{\partial c_i \partial c_j} = 2 \int \frac{\delta E_{\rm ne}}{\delta \rho} \eta_i \eta_j d\mathbf{r}.$$
(6.33)

The third term in Eq. (6.32) evaluates to

$$\frac{\partial^2 E_{\mathcal{J}}}{\partial c_i \partial c_j} = 2 \sum_{kl} c_k c_l [(ij|kl) + 2(ik|jl)].$$
(6.34)

136

For the first two terms in Eq. (6.32) we assume the functionals can be expressed as

$$F(\rho) = \int f(\rho, \boldsymbol{\nabla} \rho) d\mathbf{r} := \langle f(\rho, \boldsymbol{\nabla} \rho) \rangle, \qquad (6.35)$$

then one finds

$$\frac{\partial^{2} F}{\partial c_{i} \partial c_{j}} = \left\langle \frac{\partial^{2} \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \rho} + \frac{\partial^{2} \nabla \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \nabla \rho} \right\rangle \\
+ \left\langle \frac{\partial^{2} f}{\partial \rho^{2}} \frac{\partial \rho}{\partial c_{i}} \frac{\partial \rho}{\partial c_{j}} + \frac{\partial^{2} f}{\partial \nabla \rho^{2}} \frac{\partial \nabla \rho}{\partial c_{i}} \frac{\partial \nabla \rho}{\partial c_{j}} \right\rangle \\
+ \left\langle \frac{\partial^{2} f}{\partial \nabla \rho \partial \rho} \left[\frac{\partial \rho}{\partial c_{j}} \frac{\partial \nabla \rho}{\partial c_{i}} + \frac{\partial \rho}{\partial c_{i}} \frac{\partial \nabla \rho}{\partial c_{j}} \right] \right\rangle.$$
(6.36)

We have implemented this method into QUEST. This implementation has an interface with XCFun, an automatic differentiator, which returns exact values for all derivatives involving f at every grid-point. This means when one wants to run this scheme for an OF-KEF one only needs to supply the integrand, f, to XCFun instead of computing by hand the explicit expressions of the required derivatives.

6.3. Results

6.3.1. Comparing Lopez-Acevedo, CCH and TRIM

We discussed the Lopez–Acevedo scheme in Chapter 3, we discussed our implementation of the CCH scheme in Chapter 4 and in this current chapter we have talked about our novel scheme the TRIM scheme. In this section we will compare these schemes and provide a representative example of convergence rates of the schemes. In Table 6.1 we compare and contrast the three schemes discussed in detail in this work. The Lopez-Acevedo scheme is the easiest to implement as all that is required is a minor modification to existing KS codes. However, one should note that the DIIS and other similar convergence accelerator techniques will not work so one has to resort to a simple mixing technique. See Chapter 3 for a discussion on this. What this results in is a scheme which requires a lot of iterations (up to 10000 depending on the system and functional). The CCH scheme is probably the most difficult to implement due to the nested nature of the optimisation algorithm. Recall from Chapter 4 that one is required to wrap a bisection algorithm around a Lagrangian minimisation algorithm. The increase in complexity does lead to a decrease in the number of iterations required before convergence is reached compared to the Lopez-Acevedo scheme.

	Lopez-Acevedo	ССН	TRIM
Ease of Implementation	EASY	DIFFICULT	INTERMEDIATE
Number of Lagrangian Evaluations	HIGH	INTERMEDIATE	LOW
Fixed Particle Number In iterative process	YES	NO	NO
Applicable to general OF-KEF development	NO	YES	YES
Derivative order required	1	1 or 2	2
Quadratic convergence in local region	NO	NO	YES

Table 6.1.: Comparing the three methods discussed developed and/or implemented in this work

Table 6.2.: Number of calls to the objective functional that each method uses for the Ne atom using the ETOF basis and the PBE XC functional. The number in brackets for the Lopez–Acevedo values are the corresponding mixing parameter.

	Number of	of objective	functional cal	ls
Method	$T_{\rm VW}$	TFVW	SGA	TF02W
Lopez–Acevedo	12(0.75)	237(0.05)	2046(0.005)	8593(0.001)
CCH	167	169	249	220
TRIM	12	17	28	20

However, more importantly the CCH scheme is applicable to a wider class of functionals than the Lopez–Acevedo scheme. The Lopez–Acevedo scheme is restricted to functionals which are linear combinations of $T_{\rm TF}$ and $T_{\rm VW}$.

The TRIM scheme is by far the most efficient of the schemes and whilst more complex to implement than the Lopez-Acevedo scheme is simpler than the CCH scheme. The reason why it is much more efficient — in terms of iteration number — than the other schemes is that it quadratically converges once it is in a local region. The main drawback of this method is that in its current implementation it requires diagonalisation in order to find the eigenvalues of the Hessian. However, this can be eliminated since only the lowest eigenvalue is required.

In Table 6.2 we present the number of objective functional evaluations before convergence is reached for the Ne atom using the PBE XC functionals and different OF-KEF models. The basis used was the ETOF basis and for comparison running KS-DFT using the standard DIIS (subspace = 8) acceleration technique took 7 iterations before convergence with compara-



Figure 6.3.: Convergence in the energy for the Ne atom — using the ETOF basis and TFVW for the OF-KEF and PBE XC — run using the Lopez–Acevedo Scheme (Green Line), TRIM Scheme (Orange Line) and CCH Scheme (Blue Line).

ble tight convergence criteria of 10^{-8} a.u. on ||FDS - SDF||. In the first row of Table 6.2 the numbers in brackets are the required mixing factors. For the Lopez–Acevedo scheme adding TF causes convergence to slow down dramatically an increase giving of 225 functional evaluations. In addition it should be clear that as one reduces the amount of $T_{\rm VW}$ present in the functional convergences becomes dramatically slower. Indeed, we cannot converge the energy using the Lopez–Acevedo scheme for $T_{\rm TF}$ alone. Fortunately the CCH and TRIM schemes are relatively constant though there is a small increase in difficulty convergence from just using $T_{\rm VW}$ to using the TF λ W functionals. A similar convergence pattern is seen for all atoms tested. For example, for the TF02W functional for the neutral atomic systems with $Z = 1 \cdots 10$: the Lopez–Acevedo scheme averaged 1981 functional calls before convergence; the CCH scheme averaged 231 functional calls; and the TRIM scheme averaged 23 functional calls. Clearly the TRIM scheme is the more efficient scheme.

To explain why the TRIM scheme is the most efficient of the three schemes tested in this work it is clearest to examine the convergence graphically. In Figure 6.3 we have plotted $\log(|\Delta E|)$ as a function of iteration. ΔE is defined as the difference between the energy of the current iteration and the final converged energy. The data shown in Figure 6.3 was generated using TFVW for the Ne atom. The blue line represents the CCH



Figure 6.4.: Convergence in μ for the Ne atom — using the ETOF basis and TFVW for the OF-KEF and PBE XC — run using the Lopez–Acevedo Scheme (green Line), TRIM Scheme (orange Line) and CCH Scheme (blue Line).

scheme — 169 function evaluations, the orange line represents the TRIM scheme — 17 function evaluations and the green line represents the Lopez–Acevedo scheme — 237 function evaluations. From Figure 6.3 we see that whilst the Lopez–Acevedo scheme monotonically converges it does so very slowly. We can also see the terraced convergence of the CCH scheme due to the nested iterations. This is why it is slower to converge than the TRIM scheme which we can see rapidly converges to the solution.

As can be seen from Figure 6.4 a similar convergence pattern occurs for μ — the chemical potential. In Figure 6.4 the TRIM scheme (orange line) behaves erratically for the first 10 or so iterations. This is because the optimisation starts in the global region and this becomes even more noticeable for molecules. We will discuss how the initial guess effects the starting iterations in Section 6.3.3.

6.3.2. Comparing CCH and TRIM

So far we have been comparing the Lopez–Acevedo, CCH and TRIM schemes for the TF λ W models. We found that TRIM was the most efficient, however, we know that the CCH and TRIM methods can converge the energy for a much wider range of model OF-KEF's. In this section we will compare the convergence rates for GGA OF-KEFS used in this work. Once again all calculations use the PBE XC functional and the ETOF basis. In Table 6.3

	DK	E00	OL1	OL2	P92	PBE2	VT84
Η	64	-69	-144	-166	-112	1	53
He	-387	-79	-113	-117	-189	16	195
Li	-122	-77	-126	-127	-127	-188	31
Be	-87	-75	-116	-97	-143	-194	-3
В	-104	-75	-144	-156	-114	-173	-30
\mathbf{C}	-166	-109	-85	-134	-153	-120	-9
Ν	-215	-90	-123	-156	-187	17	26
Ο	-302	-75	-160	-87	-116	-25	-43
F	-174	-90	-81	-63	-90	-162	-26
Ne	-165	-103	-160	-156	-128	-57	-86
Na	-345	-134	-181	-193	-193	-165	-147
Mg	-269	-173	-182	-206	-204	-323	-72
Al	-202	-86	-167	-139	-139	-195	-46
Si	-288	-185	-193	-249	-249	-614	-149
Р	-285	-129	-234	-86	-196	-262	-80
\mathbf{S}	-283	-186	-127	-125	-193	-340	-95
Cl	-319	-142	-171	-112	-192	-186	-23
Ar	-320	-129	142	-103	-115	-239	-135

Table 6.3.: Difference in number of calls to the objective functional in theTRIM scheme compared to the CCH scheme

we present the change in objective functional calls when one uses the TRIM scheme rather than the CCH scheme. In Table 6.3 negative values indicate that the TRIM scheme is more efficient than the CCH scheme. Its clear that for the majority of atomic systems it is indeed the case that TRIM is much more efficient. The minority of cases where it isn't can be improved, if required, by altering the increase and decrease factors of the trust radius for the optimisation. Though on the whole we have found that computations which are faster using the CCH scheme are outliers. This is shown in Figure 6.5. The VT84 functional shows the least improvement. Further analysis shows that the VT84 Hessian has a large number of negative eigenvalues to begin with and hence spends a large number of iterations trying to get out of the global region.

In Figure 6.6 we show the percentage change in the call to the objective functional for the OF-KEFs used in this section. We can see that for the majority of cases the lines are below zero i.e. TRIM converges faster than CCH. Indeed a good proportion of the test cases are sped up by 50%. There are systems/functionals where the CCH scheme performs better than the TRIM scheme. For example, He/VT84 which has an increase in over 150%. However, these can be reduced by trailing different combinations of increase and decrease factors. The increase and decrease factors stated in Section 6.2.2 were found by ensuring we maximise the number of cases



Figure 6.5.: The mean percentage change in calls to the Lagrangian build function going from the first-order and TRIM scheme. This avergae was computed over neutral atoms with atomic charge $Z = 1 \cdots 18$ using the OF-KEF's DK, E00 and OL1,OL2, P92, PBE2 and VT84 with the PBE XC functional.

where TRIM is faster than CCH. As one can see in Figure 6.6 this pair of values is normally successful at this.

6.3.3. Molecular Calculations and the Initial Guess

For optimisations on larger molecular systems we require a good initial guess of the density to help us out of the local region. As discussed for the CCH scheme we have found that the superposition of atomic densities (SAD) greatly decreased the number of iterations required to converge the energy. We calculate the densities for the constituent atoms with a given OF-KEFs and form a superposition of these densities to begin the molecular optimisation from.

In Figure 6.7 we present how the electronic energy E changes as we go through a calculation. The particular calculation this behaviour is exhibited for in Figure 6.7 is for N₂ using the E00 OF-KEF and an even tempered 21s10p basis set. Without the SAD guess we require 602 iterations for the molecular calculation. With the SAD guess we require 14 iterations and a further 31 for the atomic calculation. Overall we do about 13 times less iterations using a SAD guess. This we found to be true for all systems investigated. Without the SAD guess convergence for the TRIM scheme is



Figure 6.6.: The percentage difference in calls to the objective functional going from the first-order and TRIM scheme. This was done for neutral atoms with atomic charge $Z = 1 \cdots 18$ using the OF-KEF's discussed in this section with the PBE XC functional.

exceptionally slow as it seems to meander in the global region. Using the SAD guess removes this issue. The explanation can be found in Figures 6.8 and 6.9.

In Figures 6.8 and 6.9 we have plotted the number of negative eigenvalues the Hessian contains at each iteration. From Figure 6.8 we see that we start from a point in the optimisation space which has a very undesirable structure and it takes the optimiser a long time to eventually settle in the local region when we consequently achieve a single negative eigenvalue which is what we require — and converges. In Figure 6.9 we see that SAD massively improves the situation. The optimiser takes only 31 iterations for the atomic (N atom) — blue line in Figure 6.9 — to converge. The superposition of the resulting density is then used in the molecular calculation (N₂). We can see that we start from an ideal point with regards to the negative eigenvalue count. Whilst this is not always the case the SAD guess always yields dramatic improvement. The fact that the SAD guess gives a starting point with either a correct or close to correct Hessian structure is why we find that the SAD guess yields fast convergence rates even when taking into account the calculations generating the SAD guess.



Figure 6.7.: Number of iterations required before convergence is acheived for N_2 with E00 and an even tempered 21s10p basis. The orange line is using a SAD guess whilst the blue line is using the core Hamiltonian as a starting point



Figure 6.8.: The number of negative values of the Hessian at every iteration for N_2 with E00 and an even tempered 21s10p basis.



Figure 6.9.: The number of negative values of the Hessian at every iteration for N_2 with E00 and an even tempered 21s10p basis. The blue line is the atomic iterations and the red is from the molecular

6.4. Conclusions

In this chapter we have introduced an OF-DFT scheme which optimises the chemical potential and density simultaneously by utilising the fact that the objective functional in DFT is a saddle function. In this chapter we found the first order saddle point of this functional using the trust-region image method. We found, in Section 6.3.1, that this new scheme is much more efficient than other all electron optimisation schemes highlighting the importance of both recognising and using the underlying saddle point optimisation in DFT. We highlighted the fact that when one closes in on the solution point this new scheme converges quadratically meaning that convergence occurs in 10's of iterations rather than the 100's or 1000's required for convergence in the CCH and Lopez–Acevedo schemes respectively. However, we have had to reintroduce diagonalisation to achieve this. Future work will require implementation of numerical methods to avoid this step.

In Section 6.3.3 we have highlighted the importance of the starting guess in optimisation for molecular systems. We found that using a simple version of SAD improves convergence rates dramatically. A case study showed that the SAD helps provide a good initial Hessian. We found that the initial Hessian formed using SAD had the correct structure at the beginning of molecular calculations. Without SAD this was not the case. This means that calculations using the SAD guess started in regions allowing for full Newton steps to be taken almost immediately. This is in contrast to using the core Hamiltonian or a random guess where we started at regions of the surface where the Hessian structure was incorrect and thus only very small steps were taken in line with the trust region algorithm. We have found that the scheme introduced in this chapter alongside a SAD guess is suitable to test a wide range of OF-KEFs for molecular systems and we hope this will be a conduit for fruitful investigations in the development of accurate OF-KEFs for chemical systems.

7. Connections between variation principles

7.1. Introduction

As we have seen, variation principles lie at the heart of many quantumchemical theories, giving practical prescriptions for how to obtain the best electronic energy, wave function or electron density via optimization. They may also offer insight into the connections between traditional *ab initio* wave-function based approaches and density-functional theory (DFT). In this chapter, we examine a new variation principle, proposed by Gidopoulos in Ref. 6 for the determination of the non-interacting system of relevance to Kohn–Sham theory.

The variation principle proposed by Gidopoulos consists of minimizing the left-hand side of the inequality

$$\langle \Psi | \hat{H}_0(\upsilon) | \Psi \rangle - E_0(\upsilon) \ge 0, \tag{7.1}$$

with respect to the variations in the potential v, for a fixed electronic wave function Ψ corresponding to a system of interest - typically, the physical ground-state wave function for the system. The energy $E_0(v)$ in Eq. (7.1) is the ground-state energy of a non-interacting system, associated with the non-interacting Hamiltonian $\hat{H}_0(v) = \hat{T} + \sum_i v(\mathbf{r}_i)$, where \hat{T} is the kinetic energy operator. As discussed in Ref. 6 the minimization of the left-hand side of Eq. (7.1) yields the Kohn–Sham non-interacting potential v_s associated with a non-interacting system that has the same density as that of the chosen input wave function Ψ . The same variation principle was also described earlier by Davidson in a different context;¹³⁵ here, we refer to Eq. (7.1) as the *Gidopoulos–Davidson variation principle*.

At first glance, the Gidopoulos–Davidson variation principle appears to be markedly different from alternative approaches for determining the Kohn–Sham system corresponding to a reference wave function or density. For example, in Levy's constrained-search approach to DFT,^{136,137} a constraint on the electron density is explicitly applied to determine the Kohn–Sham system. More closely related is the Lieb variation principle, which for a non-interacting system corresponds to maximizing the left-hand side of the inequality⁸

$$E_0(\upsilon) - (\upsilon|\rho) \le T_{\rm s}(\rho) \tag{7.2}$$

with respect to variations in the potential v for a given input electron density ρ . Both the Gidopoulos–Davidson and Lieb variation principles involve an unconstrained optimization over v, yielding the Kohn–Sham potential $v_{\rm s}$ as their optimizer. Furthermore, their functional derivatives are identical up to a sign.^{6,8}

These observations motivate us to explore the connection between the Lieb and Gidopoulos–Davidson variation principles in Eqs. (7.1) and (7.2), respectively. We begin by reviewing standard variation principles in Section 7.2. In Section 7.3, we highlight the connections between the Gidopoulos–Davidson and Lieb variation principles, including extensions to general interaction strengths and to mixed states. A brief review of the adiabatic connection (AC) is then given in Section 7.4, providing a link between the generalized functionals and the exchange–correlation energy DFT.

Having established the close connection between these alternative variation principles, we present some results from numerical implementation in a common framework in Section 7.5 highlighting the equivalent information they yield both in the non-interacting limit and for arbitrary interaction strengths. In Section 7.6, we make some concluding remarks and discuss possible directions for future work.

7.2. Variation Principles

In this section, we review the Rayleigh–Ritz variation principles for pure and mixed electronic states, introduced in Section 2.2.3 and the Hohenberg– Kohn and Lieb variation principles which were discussed in Section 2.3.

7.2.1. Rayleigh–Ritz Variation Principle

Consider an electronic system described by a Hamiltonian of the form

$$\hat{H}_{\lambda}(v) = -\frac{1}{2} \sum_{i} \nabla_{i}^{2} + \sum_{i} v(\mathbf{r}_{i}) + \sum_{i>j} w_{\lambda}(|\mathbf{r}_{i} - \mathbf{r}_{j}|)$$

= $\hat{T} + \hat{V} + \hat{W}_{\lambda}$ (7.3)

where \hat{T} is the kinetic-energy operator, \hat{V} the external potential operator, and \hat{W}_{λ} the electron-electron repulsion operator for a given electronelectron interaction strength $\lambda \in [0, 1]$, such that $w_0 = 0$ (for non-interacting systems) and $w_1 = 1/|\mathbf{r}_i - \mathbf{r}_j|$ (for physical systems). At a given interaction strength λ , the ground-state energy of an *N*-electron eigenfunction Ψ of the Hamiltonian $H_{\lambda}(v)$ can be defined in the context of wave-function theory by varying the wave function Ψ according to the **Rayleigh-Ritz variation principle**,

$$E_{\lambda}(\upsilon) = \inf_{\Psi \in \mathcal{W}_{N}} \langle \Psi | \hat{H}_{\lambda}(\upsilon) | \Psi \rangle$$
(7.4)

where \mathcal{W}_N is the set of all antisymmetric N-electron wave functions with a finite kinetic energy and a unit normalised inner product,

$$\mathcal{W}_{N} = \left\{ \Psi | \langle \Psi | \Psi \rangle = 1; \quad \sum_{i=1}^{N} \langle \nabla_{i} \Psi | \nabla_{i} \Psi \rangle < \infty \right\}$$
(7.5)

The Rayleigh–Ritz variation principle is well defined for all potentials v belonging to the vector space $\mathcal{X}^* = L_{3/2} + L_{\infty}$, which includes all Coulomb potentials.⁸

It is often more useful to work with mixed rather than pure states, giving the canonical-ensemble Rayleigh–Ritz variation principle

$$E_{\lambda}(v) = \inf_{\hat{\gamma} \in \mathcal{K}_N} \operatorname{tr} \hat{\gamma} \hat{H}_{\lambda}(v)$$
(7.6)

where \mathcal{K}_N is the set all admissible ensemble density matrices,

$$\mathcal{K}_N = \left\{ \sum_i \lambda_i \left| \Psi_i \right\rangle \! \left\langle \Psi_i \right| \left| \lambda_i \ge 0, \sum_i \lambda_i = 1, \Psi_i \in \mathcal{W}_N \right\} \! \right\}.$$
(7.7)

Although the ground-state energy can always be defined as the greatest lower bound in either Eq. (7.4) or Eq. (7.6), the formulation in terms of ensembles is more flexible, allowing for mixed-state solutions. This extra flexibility is important to establish correspondence between the optimizers in the Rayleigh–Ritz variation principle commonly used in *ab initio* theory and the Hohenberg–Kohn variation principle used in DFT.¹³⁸

7.2.2. Hohenberg–Kohn and Lieb Variation Principles

In the context of DFT, the ground-state energy is obtained according to the Hohenberg–Kohn variation principle

$$E_{\lambda}(\upsilon) = \inf_{\rho \in \mathcal{X}} \left(F_{\lambda}(\rho) + (\upsilon|\rho) \right)$$
(7.8)

where v is the external potential, ρ the electron density, $(v|\rho) = \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$ and F_{λ} Lieb's convex-conjugate universal density functional, determined from the Lieb variation principle:⁸

$$F_{\lambda}(\rho) = \sup_{v \in \mathcal{X}^*} \left(E_{\lambda}(v) - (v|\rho) \right).$$
(7.9)

The functionals E_{λ} and F_{λ} are a conjugate pair, related by mutual Legendre– Fenchel transforms. The vector spaces of admissible densities and potentials are the Banach spaces — see Section 2.3.3 for details — $\mathcal{X} = L_3 \cap L_1$ and $\mathcal{X}^* = L_{3/2} + L_{\infty}$, respectively, encompassing all *N*-representable densities $\rho \in \mathcal{X}$ and all Coulomb potentials $v \in \mathcal{X}^*$, with which the density has a finite interaction energy.

As shown in Ref. 138, there is a one-to-one correspondence between the ground-state densities obtained from the Hohenberg–Kohn variation principle with the Lieb functional as in Eq. (7.8) and from the Rayleigh– Ritz variation principle with ensembles as in Eq. (7.6) but not with pure states as in Eq. (7.4).

7.3. Gidopoulos-Davidson Variation Principles

The variation principle of Gidopoulos in Ref. 6 allows for the determination of the non-interacting system of relevance to Kohn–Sham theory and may be written in the form

$$D_0(\Psi) = \inf_{\upsilon \in \mathcal{X}^*} \left(\langle \Psi | \hat{H}_0(\upsilon) | \Psi \rangle - E_0(\upsilon) \right), \tag{7.10}$$

where $\Psi \in \mathcal{W}_N$ is an electronic wave function corresponding to the physical system of interest; typically the physical ground-state wave function of $\hat{H}_1(v)$ for some $v \in \mathcal{X}^*$. The energy $E_0(v)$ is the ground-state energy of the non-interacting system, defined according to Eq. (7.4). Note that $D_0(\Psi)$ is well defined since $\langle \Psi | \hat{H}_0(v) | \Psi \rangle - E_0(v) \geq 0$ for each $\Psi \in \mathcal{W}_N$ by the Rayleigh–Ritz variation principle.

7.3.1. Relationship to Lieb Variation Principle

The Gidopoulos–Davidson variation principle is related in a simple manner to the non-interacting Lieb variation principle

$$F_0(\rho) = \sup_{v \in \mathcal{X}^*} (E_0(v) - (v|\rho)).$$
(7.11)

In previous chapters we denoted the Lieb functional by $F_{\rm L}$ however in this work we will be varying interaction strengths. We therefore use the subscript label to denote the value of this. To see the relation, we decompose the non-interacting expectation value $\langle \Psi | \hat{H}_0(v) | \Psi \rangle$ in the manner

$$\langle \Psi | \hat{H}_0(\upsilon) | \Psi \rangle = T(\Psi) + (\upsilon | \rho_\Psi), \qquad (7.12)$$

where $T(\Psi) = \langle \Psi | \hat{T} | \Psi \rangle$ and ρ_{Ψ} are the kinetic energy and density yielded by Ψ , respectively. A comparison of the functionals in Eqs. (7.10) and (7.11) then gives,

$$D_0(\Psi) = T(\Psi) - F_0(\rho_{\Psi}),$$
(7.13)

showing that the Gidopoulos–Davidson functional of a given system is simply the total kinetic energy of this system minus the non-interacting Lieb functional.

Since the non-interacting Lieb functional is the non-interacting Kohn– Sham kinetic energy,

$$F_0(\rho) = T_s(\rho) \tag{7.14}$$

we find that the Gidopoulos–Davidson functional is the Kohn–Sham kineticenergy correlation energy,

$$D_0(\Psi) = T(\Psi) - T_{\rm s}(\rho_{\Psi}).$$
 (7.15)

Introducing the constrained-search formalism,^{8,136,137} we obtain

$$D_0(\Psi) = \langle \Psi | \hat{T} | \Psi \rangle - \inf_{\Phi \mapsto \rho_{\Psi}} \langle \Phi | \hat{T} | \Phi \rangle$$
(7.16)

151

where Φ is a single Slater determinant describing the non-interacting Kohn–Sham system.

The relationship of the Gidopoulos–Davidson functional to the correlation kinetic energy is well known.⁶ Here we see that the non-interacting Gidopoulos–Davidson and Lieb variation principles yield the same Kohn– Sham system from different directions. The Lieb variation principle minimizes the value of the non-interacting kinetic energy $T_{\rm s}$, subject to a density constraint, whilst the Gidopoulos-Davidson variation principle maximizes the correlation kinetic energy $T_{\rm c} = T - T_{\rm s}$ subject to a similar density constraint.

7.3.2. Objective Functions

Being related in such a simple manner, the optimizations of the Gidopoulos– Davidson and Lieb functional are also related in a simple way. Expressing the functionals in terms of their objective functions, we find

$$D_0(\Psi) = \inf_{v \in \mathcal{X}^*} G_0(v, \Psi) \tag{7.17}$$

$$F_0(\rho) = \sup_{v \in \mathcal{X}^*} L_0(v, \rho) \tag{7.18}$$

where

$$G_0(v,\Psi) = \langle \Psi | \hat{H}_0(v) | \Psi \rangle - E_0(v)$$
(7.19)

$$L_0(v,\rho) = E_0(v) - (v|\rho).$$
(7.20)

Hence, we obtain in agreement with Eq. (7.13),

$$G_0(v, \Psi) = T(\Psi) - L_0(v, \rho_{\Psi}).$$
(7.21)

The functional $L_0(v, \rho)$ is concave in v and affine in ρ , whereas $G_0(v, \Psi)$ is convex in v. After a generalization to mixed states, G_0 becomes convex also in the second variable; see Section 7.3.4.

7.3.3. Functional Derivatives of the Objective Functions

To determine the stationary points of the Gidopoulos–Davidson and Lieb variation principles, we note that ground-state energy $E_0(v)$ is differentiable with functional derivative ρ_v if v supports a ground state with a unique density ρ_v . For a given $\Psi \in \mathcal{W}_N$, the expectation value $\langle \Psi | \hat{H}_0(v) | \Psi \rangle$ is always differentiable with respect to v, with functional derivative ρ_{Ψ} . Hence, assuming differentiability of E_0 at v, we have

$$\frac{\delta G_0}{\delta \upsilon(\mathbf{r})} = \rho_{\Psi}(\mathbf{r}) - \rho_{\upsilon}(\mathbf{r}), \qquad (7.22)$$

$$\frac{\delta L_0}{\delta \upsilon(\mathbf{r})} = \rho_\upsilon(\mathbf{r}) - \rho_0(\mathbf{r}). \tag{7.23}$$

When $\rho = \rho_{\Psi}$, the functional derivatives are identical except for the sign difference.

The second derivatives of G_0 and L_0 with respect to the potential v may also be readily evaluated. They are equal to (minus and plus) one half the non-interacting static density response function of the system,

$$\frac{\delta^2 G_0(v,\Psi)}{\delta v(\mathbf{r}')} = -\frac{\delta^2 L_0(v,\rho)}{\delta v(\mathbf{r})\delta v(\mathbf{r}')} = -\frac{1}{2}\chi_0(\mathbf{r},\mathbf{r}').$$
(7.24)

The analytical expression of χ_0 is found using first order perturbation theory for the one electron Schrödinger equation,

$$\chi_0(\mathbf{r}, \mathbf{r}') = \sum_{ia} \frac{\phi_i(\mathbf{r})\phi_i^*(\mathbf{r}')\phi_a(\mathbf{r}')\phi_a^*(\mathbf{r})}{\varepsilon_i - \varepsilon_a} + \text{c.c.}$$
(7.25)

where the indices i and a denote occupied and virtual orbitals, respectively, whose orbital energies are ε_i and ε_a . In Ref. 6 focus is placed on the optimization of G_0 with respect to v. In passing, we note that the non-interacting Hamiltonian readily separates into one-electron contributions $\hat{H}_0(v) = \sum_k \hat{h}_k(v)$ with $\hat{h}_k(v) = -\frac{1}{2}\nabla_k^2 + v(\mathbf{r}_k)$ and that the orbitals entering Eq. (7.24) are the eigenfunctions of this one-electron Hamiltonian. The non-interacting ground-state energy is the sum of the occupied orbital energies, $E_0(v) = \sum_i \varepsilon_i$. We also remark that, although $\pm \frac{1}{2}\chi_0(\mathbf{r}, \mathbf{r}')$ is positive/negative semi-definite, this does not imply that G_0/L_0 are convex/concave in v since the derivatives in Eq. (7.24) are not defined for all potentials.

Throughout this discussion we have assumed the differentiability of $L_0(v, \rho)$ and $G_0(v, \Psi)$. The functional $L_0(v, \rho)$ is not straightforwardly differentiable as discussed by Lammert,¹³⁹ however this issue can be avoided by using a regularized form as discussed in Ref. 38. Since the derivative of $G_0(v, \Psi)$ amounts to taking the derivative of $-L_0(v, \rho_{\Psi})$ (see Eq. (7.21)), the same regularization techniques can be applied to this functional.

In the Gidopoulos–Davidson and Lieb variation principles, the potential plays the role of a Lagrange multiplier. To see this, we rewrite the Lieb variation principle as a minimax problem

$$F_{\lambda}(\rho) = \sup_{\upsilon \in \mathcal{X}^*} \inf_{\Psi \in \mathcal{W}_N} \left(\langle \Psi | \hat{H}_{\lambda}(\upsilon) | \Psi \rangle - (\upsilon | \rho) \right)$$
(7.26)

$$= \sup_{\upsilon \in \mathcal{X}^*} \inf_{\Psi \in \mathcal{W}_N} \left(\langle \Psi | \hat{H}_{\lambda}(0) | \Psi \rangle - (\upsilon | \rho - \rho_{\Psi}) \right)$$
(7.27)

Since the search over v terminates when the stationary condition $\rho(\mathbf{r})$ –

 $\rho_{\Psi}(\mathbf{r}) = 0$ is satisfied, the Lieb variation principle represents a minimization of the expectation value $\langle \Psi | \hat{H}_{\lambda}(0) | \Psi \rangle$ subject to the constraint that $\Psi \mapsto \rho$. The potential plays precisely the same role in the Gidopoulos–Davidson variation principle, which becomes clear in light of the relationship between the two functionals given by Eq. (7.13).

7.3.4. Generalisation to Ensembles

In Section 2.3.5 the Lieb functional $F_0(\rho)^8$ was introduced and we found that it is convex in ρ . It follows that $-F_0(\rho)$ is concave in ρ but not that the Gidopoulos–Davidson functional $D_0(\Psi) = T(\Psi) - F_0(\rho_{\Psi})$ is concave in Ψ since $T(\Psi)$ is not. Generalizing the Gidopoulos–Davidson functional for pure states $\Psi \in \mathcal{W}_N$ to canonical ensembles $\hat{\gamma} \in \mathcal{K}_N$, we obtain the concave functional

$$\mathcal{D}_{0}(\hat{\gamma}) = \inf_{\upsilon \in \mathcal{X}^{*}} \left(\operatorname{tr} \hat{\gamma} \hat{H}_{0}(\upsilon) - E_{0}(\upsilon) \right)$$

= $\mathcal{T}(\hat{\gamma}) - \sup_{\upsilon \in \mathcal{X}^{*}} \left(E_{0}(\upsilon) - (\upsilon | \rho_{\Psi}) \right),$ (7.28)

where $\mathcal{T}(\hat{\gamma}) = \operatorname{tr} \hat{\gamma} \hat{T}$. To show concavity, we select $\hat{\gamma}_1, \hat{\gamma}_2 \in \mathcal{K}_N$ and obtain for each $0 < \nu < 1$ the inequality

$$\mathcal{D}_{0}(v\hat{\gamma}_{1} + (1-v)\hat{\gamma}_{2})$$

$$= v \operatorname{tr} \hat{\gamma}_{1}\hat{T} + (1-v) \operatorname{tr} \hat{\gamma}_{2}\hat{T} - F_{0}(v\rho_{1} + (1-v)\rho_{2})$$

$$\leq v \operatorname{tr} \hat{\gamma}_{1}\hat{T} + (1-v) \operatorname{tr} \hat{\gamma}_{2}\hat{T} - vF_{0}(\rho_{1}) - (1-v)F_{0}(\rho_{2})$$

$$= v\mathcal{D}_{0}(\hat{\gamma}_{1}) + (1-v)\mathcal{D}_{0}(\hat{\gamma}_{2})$$
(7.29)

where in the second step we have used the convexity of the Lieb functional.

Since Ψ occurs quadratically in $D_0(\Psi)$, a similar proof is precluded for the pure state Gidopoulos–Davidson functional, which is indeed not concave. Note that, for pure states $\hat{\gamma}_{\Psi} = |\Psi\rangle\langle\Psi|$, the ensemble Gidopoulos–Davidson functional reduces to the original functional: $\mathcal{D}_0(\hat{\gamma}_{\Psi}) = D_0(\Psi)$.

7.3.5. Generalization to Arbitrary Interaction Strengths

The Gidopoulos–Davidson functional may be extended to interacting systems in the manner

$$D_{\lambda}(\Psi) = \inf_{\upsilon \in \mathcal{X}^{*}} \left(\langle \Psi | \hat{H}_{\lambda}(\upsilon) | \Psi \rangle - E_{\lambda}(\upsilon) \right)$$
(7.30)

which is related to the Lieb Functional via

$$D_{\lambda}(\Psi) = \langle \Psi | \hat{T} + \hat{W}_{\lambda} | \Psi \rangle - \inf_{\Phi \mapsto \rho_{\Psi}} \langle \Phi | \hat{T} + \hat{W}_{\lambda} | \Phi \rangle.$$
 (7.31)

The first derivative of the objective functional, $G_{\lambda}(v, \Psi) = \langle \Psi | \hat{H}_{\lambda}(v) | \Psi \rangle - E_{\lambda}(v)$, is again a simple density difference, simple density difference,

$$\frac{\delta G_{\lambda}(v,\Psi)}{\delta v(\mathbf{r})} = \rho_{\Psi}(\mathbf{r}) - \rho_{v}(\mathbf{r}), \qquad (7.32)$$

and its second derivative can be expressed in terms of the λ -interacting density response function

$$\frac{\delta^2 G_{\lambda}(v, \Psi)}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} = -\frac{1}{2} \chi_{\lambda}(\mathbf{r}, \mathbf{r}').$$
(7.33)

To perform practical optimizations using Eq. (7.30), we therefore require knowledge not only of the kinetic energy associated with the input wave function Ψ but also the λ -interacting two-electron interaction energy, $W_{\lambda}(\Psi) = \langle \Psi | \hat{W}_{\lambda} | \Psi \rangle.$

In practice, these quantities can be computed from the one- and twoparticle reduced density matrices, respectively.

7.4. Adiabatic Connection

The adiabatic connection considers the link between the non-interacting Kohn–Sham auxiliary and physically-interacting systems.^{140–143} In this approach, the interaction strength λ in Eq. (7.3) is varied between 0 and 1, whilst imposing the constraint that, at each interaction strength, the electron density ρ_{λ} remains fixed at that of the physical system ρ_1 . Most frequently, a linear path between these two limits is considered,¹⁴¹ where the Coulomb operator is simply scaled linearly by the value of λ . However, generalized ACs¹⁴⁴ have been explored along non-linear paths.^{145,146} In the present work, only the linear path is considered but the generalization to non-linear paths is straightforward.

For the Hamiltonian in Eq. (7.3), the λ -dependent universal density functional can be written in the constrained-search^{8,136,137} form for canonical ensembles,

$$F_{\lambda}(\rho) = \min_{\hat{\gamma} \to a} \operatorname{tr} \hat{H}_{\lambda}(0)\hat{\gamma} \tag{7.34}$$

where the minimization is over all density matrices $\hat{\gamma}$ associated with the

input electron density ρ . This functional is convex in ρ , concave in λ and non-negative for $\lambda \geq 0$. The λ -interacting functional can be related to its non-interacting counterpart via

$$F_{\lambda}(\rho) = F_0(\rho) + \int_0^{\lambda} \frac{\partial F_{\nu}(\rho)}{\partial \nu} d\nu, \qquad (7.35)$$

where the derivative is well-defined on the real axis as a right- or leftderivative. Evaluation of the derivative and application of the Hellmann– Feynman theorem^{147,148} leads to an **ab initio** expression for the exchange– correlation energy

$$E_{\rm xc}(\rho) = \int_0^1 \mathcal{W}_{\lambda}(\rho) \mathrm{d}\lambda, \qquad (7.36)$$

where $\mathcal{W}_{\lambda}(\rho)$ is the AC integrand

$$\mathcal{W}_{\lambda}(\rho) = \langle \Psi_{\lambda} | \hat{W} | \Psi_{\lambda} \rangle - E_J(\rho).$$
(7.37)

Furthermore, the exchange and correlation energies may be resolved into separate components, resulting in an expression for the correlation energy alone

$$E_{\rm c}(\rho) = \int_0^1 \left\{ \mathcal{W}_{\lambda}(\rho) - \mathcal{W}_0(\rho) \right\} \mathrm{d}\lambda \tag{7.38}$$

For a review of the adiabatic connection, see Ref. 149.

To make practical use of these expressions, approaches for the calculation of the λ -interacting wave functions yielding a chosen electron density are required, see e.g. Refs. 150–152. The constraint that the density is fixed for all λ may be easily enforced by supplying fixed arguments ρ and Ψ to Eqs. (7.9) and (7.30)) for all λ . We now discuss our implementation of the (generalized) Gidopoulos–Davidson variation principle, exploring the close connections to the generalized Lieb functional numerically.

7.5. Results

From the discussion in the Section 7.3, it is evident that the Gidopoulos– Davidson and Lieb optimizations are sufficiently closely related that they may be implemented in a common computational framework. We first discuss some details of our implementation; we then demonstrate the equivalence of the two approaches by performing numerical optimizations for a set of small atomic and molecular systems.

7.5.1. Computational Details

The variation principle given in Eq. (7.30) allows a value to be obtained for the generalized Gidopoulos–Davidson functional by evaluating its infimum with respect to v. If the density yielded by the reference wave function ρ_{Ψ} is v-representable, the infimum becomes a minimum. To vary v such that an optimization over the potential may be carried out in a practical computational scheme, the potential is modelled using the basis-set expansion of Wu and Yang^{153,154}

$$v_{\lambda}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + (1 - \lambda)v_{\text{ref}}(\mathbf{r}) + \sum_{t} b_{t}g_{t}(\mathbf{r}).$$
(7.39)

Here $v_{\text{ext}}(\mathbf{r})$ is the external potential due to interaction of the electrons with the atomic nuclei, $v_{\text{ref}}(\mathbf{r})$ is a fixed reference potential chosen to ensure that $v_{\lambda}(\mathbf{r})$ has the correct asymptotic behaviour, and $\{g_t\}$ are a set of Gaussian basis functions with coefficients $\{b_t\}$. The reference potential employed in the present work is the Fermi–Amaldi potential¹⁵⁵ whose analytical form is given by

$$\upsilon_{\rm ref}(\mathbf{r}) = \frac{N-1}{N} \int \frac{\rho_{\rm in}(\mathbf{r})}{|\mathbf{r}' - \mathbf{r}|} \mathrm{d}\mathbf{r},\tag{7.40}$$

where $\rho_{\rm in}$ is the given input density which in this work will be CCSD densities. With this choice of potential expansion, the derivatives corresponding to Eq. (7.32) and Eq. (7.33) may be readily implemented as described in Refs.,^{151,152,154} allowing the objective functional to be effectively optimized with respect to the set of coefficients $\{b_t\}$.

An un-contracted form of the Gaussian basis set aug-cc-pVTZ^{121,156} in the spherical-harmonic basis is used for both the orbital expansion and for the potential expansion in Eq. (7.39), for all systems. An approximate Newton method is employed to accelerate convergence of the optimization, ¹⁵⁷ in which the Hessian is regularized using a truncated singular value decomposition with a threshold of 10^{-6} a.u. In all calculations, the convergence threshold was set to 10^{-6} a.u. on the L^2 norm of the objective functional gradient. To obtain a reasonably accurate approximation to the Kohn– Sham system, the input quantities for each functional $F_{\lambda}(\rho_{\Psi})$ and $D_{\lambda}(\Psi)$ were determined at the CCSD level of theory. All calculations were carried out with the QUEST rapid development platform;¹¹⁶ an electronic-structure code developed in Python and exploiting just-in-time compilation using the Numba package.^{158,159}

7.5.2. Kohn–Sham Non-Interacting System

In Table 7.1, the optimized values of the non-interacting Lieb functional $F_0(\rho_{\Psi})$ and Gidopoulos–Davidson functional $D_0(\Psi)$ are presented for a series closed–shell atoms and for the hydrogen molecule at several bond lengths. Additionally, Kohn–Sham energy components are presented, including the internuclear repulsion energy $E_{\rm nn}$, the non-interacting kinetic energy $T_{\rm s}$, the electron–nuclear attraction $E_{\rm ne}$, the Coulomb energy $E_{\rm J}$, the exchange energy $E_{\rm x}$, and the correlation energy $E_{\rm c}$. These components have the same definition when computed from $F_0(\rho_{\Psi})$ and $D_0(\Psi)$. For comparison, the total kinetic energy T and total electron–electron interaction energy W are included, along with the total interacting ground-state energy E_1 .

The consistency of the optimizations was verified by comparing the optimized values of $F_0(\rho_{\Psi})$ and $D_0(\Psi)$ presented in Table 7.1 with the energetic components T_s and T_c respectively. The value of T_s was determined from the Kohn–Sham orbitals obtained at $\lambda = 0$ and the value of T_c was obtained by subtraction of T_s from T, where the latter was determined directly from the $\lambda = 1$ calculation. The H₂ molecule provides a simple prototypical system with which the variation between dynamic and static correlation may be explored. At equilibrium geometry, the electron densities of the two hydrogen atoms overlap substantially, thus binding the molecule and leading to both kinetic and potential contributions to the correlation energy. As the interatomic bond is extended, the system approaches that of two isolated hydrogen atoms, with no kinetic correlation energy; see Table 7.1, where the value of the Gidopoulos–Davidson functional D_0 decreases as the interatomic bond length R increases, becoming just 0.0005 a.u. at R = 10.0a.u.

7.5.3. General Interaction Strengths

In Fig. 7.1, results of optimizations pertaining to the generalized Lieb and Gidopoulos–Davidson functionals, according to Eqs. (7.9) and (7.30), respectively, are presented for interaction strength λ in the range 0 to 1. In the right panel, the Lieb functional $F_{\lambda}(\rho_{\Psi})$ is shown as a function of λ for the H₂ molecule with bond length R = 0.7, 1.4, 3.0, 5.0, 7.0 and 10.0 a.u. The variation of $F_{\lambda}(\rho_{\Psi})$ in λ is broadly linear, indicating that $T_{c,\lambda} = T_1 - T_{\lambda}$ is relatively small and reflecting the dominance of the Coulomb and exchange energies in the two–electron energy W, both of which are linear in λ . The slope of $F_{\lambda}(\rho_{\Psi})$ in λ becomes progressively smaller as the bond

l'able 7.1.: Optimiz	ed tunction	nal value	s and en	ergy comf	onents calc	ulated in	the aug-cc-	-pV1Z bas	sis using th	ie variatio	on principles c
Eds. (7.	10) and (7.	11)). All	quantitie	es are in a	tomic units						
	F_0	D_0	$E_{ m nn}$	$T_{\rm s}$	$E_{ m ne}$	$E_{\rm J}$	$E_{\rm x}$	$E_{\rm c}$	T	M	E_1
He	2.8611	0.0355	0.0000	2.8611	-6.7455	2.0464	-1.0232	-0.0756	2.8967	0.9477	-2.9011
Be	14.5835	0.0661	0.0000	14.5835	-33.6945	7.2122	-2.6725	-0.1534	14.6496	4.3863	-14.6586
Ne	128.5050	0.2720	0.0000	128.5050	-310.9007	65.9350	-12.0691	-0.6071	128.7765	53.2589	-128.8654
${ m H}_2 \; (R=0.7)$	1.7263	0.0324	1.4286	1.7263	-4.8614	1.6508	-0.8254	-0.0700	1.7588	0.7554	-0.9187
H ₂ $(R = 1.4)$	1.1390	0.0328	0.7143	1.1390	-3.6469	1.3215	-0.6607	-0.0729	1.1718	0.5879	-1.1729
H ₂ $(R = 3.0)$	0.8279	0.0418	0.3333	0.8279	-2.6181	0.9539	-0.4769	-0.1184	0.8697	0.3586	-1.0564
H ₂ $(R = 5.0)$	0.9520	0.0224	0.2000	0.9520	-2.3809	0.8193	-0.4097	-0.2063	0.9744	0.2033	-1.0033
H ₂ $(R = 7.0)$	0.9919	0.0052	0.1429	0.9918	-2.2826	0.7669	-0.3835	-0.2406	0.9971	0.1429	-0.9998
H ₂ ($R = 10.0$)	0.9981	0.0005	0.1000	0.9982	-2.1983	0.7245	-0.3623	-0.2623	0.9986	0.1000	-0.9997

M. S. Ryley



Figure 7.1.: F_{λ} of Eq. (7.9), right panel, and D_{λ} of Eq. (7.30), left panel, in (a.u.) as functions of the interaction strength λ for the H₂ molecule at internuclear separations R = 0.7, 1.4, 3.0, 5.0, 7.0 and 10.0 a.u.

length is increased. This behaviour reflects the fact that the H₂ molecule dissociates into two one-electron fragments with $\lambda E_{\rm J} + \lambda E_{\rm x} + E_{\rm c,\lambda} \rightarrow 0$ as $R \rightarrow \infty$ (static correlation energy cancelling the Coulomb and exchange energy).

In the left panel of Fig. 7.1, the Gidopoulos–Davidson functional $D_{\lambda} = T_1 - T_{\lambda} + \lambda (W_1 - W_{\lambda})$ is also plotted as a function of interaction strength. This functional adopts the value of T_c at $\lambda = 0$ and decreases with increasing λ to become 0 at $\lambda = 1$. In contrast to the Lieb functional, this small correlation contribution to the energy reveals the higher-order dependence of the correlation energy on λ at increasingly extended bond lengths. As the bond length R increases and the system approaches one of independent atoms, the value of the Gidopoulos–Davidson functional is smaller at $\lambda = 0$, reflecting a decrease in T_c . However, it also exhibits more pronounced curvature, indicating higher-order dependence on λ as static correlation becomes more significant.

7.5.4. Constructing the Adiabatic Connection

As described in Section 7.4, the AC comprises a link between the noninteracting Kohn–Sham auxiliary system and the physically interacting system through variation in interaction strength, modulated by coupling– constant λ , with the density equal to the physical density of $\lambda = 1$ for all λ . The AC integrand is expressed in Eq. (7.37), from which an exact definition of the correlation energy may be constructed according to Eq. (7.38). Given that the exchange energy scales linearly with λ (for the linear–attenuation AC path), the exchange contribution to Eq. (7.37) is simply a constant and may be subtracted to give the correlation component of the AC integrand,

$$\mathcal{W}_{c,\lambda}(\rho) = \mathcal{W}_{\lambda}(\rho) - \mathcal{W}_{0}(\rho). \tag{7.41}$$

The Gidopoulos–Davidson variation principle of Eq. (7.30) and the Lieb variation principle of Eq. (7.9) can both be exploited to calculate this integrand, using the same input ρ_{Ψ} or Ψ but with a range of different values of λ , to construct the AC using Eq. (7.41).

The equivalence of the AC curves constructed from the Lieb and Gidopoulos– Davidson functionals is confirmed numerically for the H₂ molecule at the same geometries considered in Table 7.1, with the AC integrands $W_{c,\lambda}$ plotted as a function of λ in Fig. 7.2. Here, values of $W_{c,\lambda}$ computed with the Lieb functional Eq. (7.9) are represented by solid lines, whilst values obtained from the Gidopoulos–Davidson functional Eq. (7.30) are denoted by the point markers. It is evident from Fig. 7.2 that the AC curves of these two methods agree to within the convergence of the optimization procedures.

The correlation energy can be computed from these curves using Eq. (7.38) and the numerical values of E_c are presented in Table 7.1. The ratio $|E_c|/T_c$ has been used to assess the relative importance of static correlation.¹⁶⁰ The $|E_c|$ corresponds to the area above each curve in Fig. 7.2, whilst T_c corresponds to the area between each curve and a horizontal line defined by its value of $W_1(\Psi)$. As R increases, this ratio grows and the curves approach an L shape characteristic of systems dominated by strong correlation, indicating that the value of T_c is approaching zero, consistent with the effects of hydrogen molecule dissociation discussed in Section 7.5.3.

7.6. Conclusions

The variation principle proposed in different contexts by Gidopoulos⁶ and Davidson¹³⁵ has been examined and shown to be closely linked to the Lieb variation principle.⁸ For the non-interacting system, the two functionals approach the Kohn–Sham system from different directions. The Lieb functional minimizes the non-interacting kinetic energy T_s subject to the constraint that the density is equal to that of the physical system, whereas the Gidopoulos–Davidson functional maximizes the kinetic correlation energy T_c under the same density constraint. In both cases, the optimization can be cast in an unconstrained manner, with the external potential as a Lagrange multiplier to impose the density constraint.



Figure 7.2.: The correlation adiabatic connection integrand values (a.u.) of Eq. (7.41), calculated using the optimization of Eq. (7.9), lines, and Eq. (7.30), point markers, for the H₂ molecule at internuclear separations R = 0.7, 1.4, 3.0, 5.0, 7.0 and 10.0 a.u.

An extension of the Gidopoulos–Davidson functional to ensembles was also presented, for which the associated functional can be shown to be concave with respect to $\hat{\gamma}$. This contrasts the pure-state functional which is not concave with respect to Ψ . The Gidopoulos–Davidson variation principle has also been extended to general interaction strengths λ , as has previously been done with the Lieb functional.^{8,150–152} Utilizing this extension, it was shown that either functional may be used to calculate the adiabatic connection between the Kohn–Sham system of non-interacting electrons and the physically-interacting system, highlighting the close connection between the approaches.

8. Future Work and Summary

8.1. Future Areas of Research

The likelihood of using OF-DFT in applied chemical research in the near future is low. Indeed even by using PKSA analysis one normally finds the proposed OF-KEF yields inaccurate energies and cannot predict binding in molecules. These errors are worsened when one tests how these functionals perform in the optimisation procedures discussed in this work. Or in other words testing these functionals in OF-DFT calculations show that no current OF-KEFs is accurate enough to be used in chemical applications. A damning statement considering we have had OF-KEFs since the 1930's. To conclude this thesis we will discuss some avenues for future functional development and outline any initial findings we have found throughout this work. We re-emphasise that all research undertaken in the following topics should not be done using PKSA. We propose that the TRIM scheme introduced in Chapter 6 is suitable for such studies.

8.1.1. Laplacian Level Functionals

Throughout this thesis we have assessed the performance of LDA and GGA type functionals. If one expands the kinetic energy to fourth order, see Equation (3.26), one gets the GEA4 functional which is a Laplacian dependent functional. We can express a general Laplacian dependent functional through

$$F(\rho) = \int \rho^{5/3}(\mathbf{r}) f(\rho, \nabla \rho, \nabla^2 \rho) \mathrm{d}\mathbf{r}, \qquad (8.1)$$

where f has differing analytic forms depending on the functional. Recent work^{161–163} has investigated the accuracy of new Laplacian functionals. The research was conducted using PKSA which, as this thesis strongly suggests, we should be sceptical of until one tries to optimise these functionals. In QUEST we have implemented the Laplacian level Hessian. For the implemented equations see Appendix A.3.3. In addition to some of the Laplacian functionals introduced in this recent work we have also tested: GEA4; MGGA;¹⁶⁴ mGEA4;¹⁶⁵ PP88;¹⁶⁶ MGEA4;¹⁶⁷ RDA;⁹⁷ and a range of functionals proposed by King and Handy.¹²⁹ In this limited study we found that convergence was much more challenging for a general Laplacian dependent functionals. However, for various modified (parametrised) versions of GEA4 (mGEA4 and MGEA4) we found that convergence was relatively simple. Future work should begin by investigating why Laplacian dependent functionals are more difficult to converge the energy for than their GGA counterparts.

8.1.2. Non-Local Functionals

In this thesis we have concentrated on local functionals. There has been some recent nonlocal functionals which have demonstrated promising results for atoms. Liu and Parr^{168–170} studied the development of DFT in terms of well-behaved functionals. These functionals have the following form

$$F(\rho) = C + \int \rho(\mathbf{r}) \frac{\delta F}{\delta \rho(\mathbf{r})} d\mathbf{r} - \frac{1}{2} \iint \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \frac{\delta^2 F(\rho)}{\delta \rho(\mathbf{r}_1) \delta \rho(\mathbf{r}_2)} d\mathbf{r}_1 d\mathbf{r}_2 + \frac{1}{3!} \iiint \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \rho(\mathbf{r}_3) \frac{\delta^3 F(\rho)}{\delta \rho(\mathbf{r}_1) \delta \rho(\mathbf{r}_2) \delta \rho(\mathbf{r}_3)} d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 - \cdots$$
(8.2)

The problem is that to implement the full expansion in Equation (8.2) one has to know the functional derivative up to all orders

It was shown^{170–173} that the correlation energy, $E_{\rm c}(\rho)$, and $T(\rho) - T_{\rm s}(\rho)$ can be expanded to good accuracy in a power series expansion where each term had a different degree with respect to co-ordinate scaling. However, this cannot be done for $T_{\rm s}$ as it has an exact degree of homogeneity with respect to co-ordinate scaling of 2. Instead Liu and Parr¹⁷⁰ decided to expand $T_{\rm s}$ in a power series of homogeneous functionals with respect to density scaling.

Liu and Parr assumed that the series in Equation (8.2) existed and converges uniformly such that it is sufficiently accurate after n terms. From this assumption they then showed that $F(\rho)$ can be expanded in terms of functionals with homogeneity of (n + 1) with respect to density scaling. They then proposed that the $F(\rho)$ in Equation (8.2) can be expressed as a power series of some local functionals $D_j(\rho)$. These functionals are homogeneous with respect to density scaling with a degree of homogeneity, n_j ,
and are defined through

$$\int \rho(\mathbf{r}) \frac{\delta D_j(\rho)}{\delta \rho(\mathbf{r})} d\mathbf{r} = n_j D_j(\rho).$$
(8.3)

The power series expansion is then given by

$$F(\rho) = \sum_{j=1}^{n} C_j [D_j(\rho)]^j$$
(8.4)

where $C_j \in \mathcal{R}$ are to be determined. In Liu and Parr's work these coefficients were found by a least squares minimisation using the Hartree-Fock reference values for the neutral atoms with $Z = 1 \cdots 18$. These values are: $C_1 = 3.26422$; $C_2 = 0.02631$; and $C_3 = 0.000498$.

The functional

$$F_j(\rho) = C_j [D_j]^j, \tag{8.5}$$

where D_j is homogeneous of degree m with respect to coordinate scaling, has an explicit form of

$$F_j(\rho) = C_j \left[\int \rho^{(1+\frac{m}{3j})}(\mathbf{r}) \mathrm{d}\mathbf{r} \right]^j.$$
(8.6)

Such that F_j is also homogeneous of degree m with respect to coordinate scaling. The Liu-Parr expansion is then approximated through

$$F(\rho) = \sum_{j=1}^{n} F_j.$$
 (8.7)

For $T_{\rm s}$ the degree of homogeneity with respect to coordinate scaling is 2 and, therefore, the Liu-Parr expansion of $T_{\rm s}$ takes the form

$$T_{\rm s}(\rho) = \sum_{j=1}^{n} C_j \left[\int \rho^{(1+\frac{2}{3j})} \right]^j.$$
(8.8)

Liu and Parr truncated this expansion after n = 3. In more recent work using the Liu-Parr expansion Karasiev et al.¹⁷⁴ expanded up-to 7th order. They showed that energies do improve — using post-SCF methods — when one includes higher order terms in the expansion as one expects. Furthermore, they showed that using a shell-by-shell mean square optimisation of the expansion coefficients, $\{C_j\}$, yields a dramatic improvement in the errors.

$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$.848279 -47.46968 .537157 -48.391760 .499453 -48.309774 .067774 -37.113587		$\begin{array}{r} 3 & -62.368489 \\) & -69.218550 \\ 1 & -69.060140 \\ \hline & -53.7092 \end{array}$	3 -62.368489 -85.533084 0 -69.218550 -89.177149 1 -69.060140 -88.902693 -53.7092 -73.999839	3 -62.368489 -85.533084 -113.969828 0 -69.218550 -89.177149 -117.731125 1 -69.060140 -88.902693 -116.082433 -53.7092 -73.999839 -98.477731	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
--	---	--	--	---	--	--	--

М.	S.	Rylev
TAT .	\sim .	10,10,

M. S. Ryley

In Table 8.1 we display the energies predicted by the non-local functionals. Although Karasiev et al. parametrised up to the 7th order we found that the functional potentials were unstable near the nuclei due to the high powers of the density e.g. $\rho^{11/3}$. Thus we could not converge the energy when we include 4th order or higher terms. For the non-local functionals we could converge the energy but, unfortunately we found that the predicted energies were not as good as the TFDW models presented in Chapter 4. Furthermore, if we compare Tables 8.1 and 8.2 we see there is a large increase in the MAPE/MAE's going from PKSA to self-consistent analysis. The increase is around 700-800%. Again we see that we should really be testing these functionals in codes such as TRIM.

Future work for non-local functionals should be focussed on re-parametrization to minimize the errors in Table 8.1 rather than minimising the errors with respect to Hartree-Fock values. Furthermore, it would be interesting to study the potentials as in Chapter 5 to see how the kinetic potential behaves with respect to the KS potential. In addition it was noted by Liu and Parr that these non-local functionals are not size consistent, however recently¹⁷⁵ this issue of size consistency for exchange correlation functionals in KS-DFT was corrected. It would be interesting to do something similar for these kinetic functionals.

8.1.3. Machine Learning in OF-DFT

Machine learning (ML) is a topic which has garnered much interest in the general population. For example, one can find plenty of videos online offering instruction on the topic. Due to the increased computing power available to researchers there has been lots of interest in using ML in computational chemistry. For example, ML has been used to predict properties of structures after being trained from DFT calculations.^{176,177} ML has also been used to study properties of solid state matter.¹⁷⁸ Furthermore, ML has been used to approximate density functionals.^{179,180} Snyder et al.¹⁸⁰ used a type of ML called kernel ridge regression to produce accurate self consistent OF-DFT densities and energies. This OF-DFT scheme was only been demonstrated for 1D systems only. See Li et al.¹⁸¹ for details on this. However, there has been recent work which extends this for 3D molecular systems.¹⁸²

Recent work by Golub and Manzhos¹⁸³ could provide a starting point for future research utilising the TRIM scheme presented in this thesis. In this work they used neural networks to re-parametrise the fourth order gradient

expansion (GEA4) of the kinetic energy density (KED). Their calculations were undertaken using the PROFESS⁸² code. This meant this work was undertaken using a plane wave basis and a pseudopotential. In this work they compared the KED from KS-DFT to the KED's of their re-parametrised gradient expansions for Al, Si, Mg, H_2O and C_6H_6 . They found the reparametrisation of GEA4 using neural networks lead to an improvement in the KED errors with respect to the KS-DFT reference. However, they pointed out that using the pseudopotential caused numerical instabilities near the nuclei. They found that by treating more electrons explicitly they could resolve these issues. As the schemes presented in this work are allelectron and can converge the energy for GEA4 type functionals we are in a good position to attempt the same re-parametrization, using neural networks, of the KEDs (or indeed potentials) for a wide range of chemical systems. It would be a good step forward in OF-DFT research if the findings by Golub and Manzhos can be extended to a wider range of chemical systems.

8.2. Conclusions

Method development in electronic structure theory often strikes a compromise between accuracy and computational cost. In this thesis we have been primarily focussed on OF-DFT which is a method which promises to be a method with ultra low computational cost. This has been used to study metallic systems with millions of atoms using purely quantum mechanical methods.^{67,76} In OF-DFT one computes the total energy using functionals of the ground state density only. This is in contrast to KS-DFT where one uses orbitals. The use of which leads to an increase in computational cost due to the fact in KS-DFT we have a set of non-linear eigenvalue equations to solve. In OF-DFT there is only one non-linear variational problem to solve. There are two main challenges in OF-DFT development for chemical systems. One is in the development of accurate OF-KEFs. However, in this work we have focussed on the second issue which is how one solves for the ground state density and energy for a given OF-KEF for chemical systems.

In Chapters 2 and 3 we outlined the underlining theory of this work. The main conclusion from this chapter is that the optimisation at the heart of OF-DFT is a convex-concave **saddle point optimisation**. In Chapter 4 we discussed an approach to the optimisation problem which breaks down the saddle point optimisation into a series of nested minimisations. This approach was developed by CCH^2 and we have extended it so that we

can converge the energy for and, therefore, investigate, a wide range of OF-KEFs.³ We found that some functionals could not be converged due to their analytical form. This was not noted in the various papers introducing these functionals because they relied on PKSA. For the functionals that we could converge the energy for we found that the energy errors with respect to KS-DFT for atoms and molecules were much poorer compared to the energy errors generated using PKSA. Upon further analysis we discovered this increase in error was due to the inaccurate densities one generates using the KEFs in variational calculations. Once again PKSA did not provide any indication of this. The overriding conclusion is that one should be sceptical of conclusions based on PKSA of proposed KEFs and future research should endeavour to use the optimisation methods of the type discussed in this thesis.

In Chapter 5 we presented a small study on the potentials which are components of the Euler equation generated from the optimisation discussed in Chapter 4. We discovered that the balancing act between the potentials is very susceptible to inaccuracies in the chosen basis. Furthermore, we demonstrated that future work which relies on forces generated by OF-KEFs, for example, geometry optimisations, one must include a Pulay like term. This Pulay term, however, does not appear if one uses plane wave basis sets. In Chapter 5 we also investigated how this balance of potentials can affect the ground state density computed using the optimisation procedures outlined in this work. This study concentrated on a few linear combinations of $T_{\rm TF}$ and $T_{\rm VW}$. We demonstrated how the resulting potentials can affect the predicted ionisation energy and radial density. In future work it would be instructive to develop functionals which reproduce the potentials exhibited in KS-DFT and to explore its effect on the densities and energies for atomic and molecular systems.

In Chapter 6 we introduce an optimisation scheme,⁴ TRIM, which exploits the fact that the ground state density is generated through a saddle point optimisation. Unlike the CCH scheme the TRIM scheme simultaneously optimises the chemical potential and density. This yields a scheme which is far more efficient in terms of iteration count. Indeed when comparing the TRIM, CCH and Lopez–Acevedo¹ schemes we found that the TRIM scheme outperforms the rest when it comes to converging the energy for a wide range of functionals and doing so efficiently. However, our current implementation of the TRIM scheme requires a matrix diagonilisation which is a costly step. Future development of the scheme should focus on reducing the cost of this step. This matrix diagonilisation can be avoided

as we do not need the full spectrum; only the lowest eigenvalue is required. In this case numerical methods such as the inverse iteration method do exist for this task.¹⁸⁴

We also investigated the importance of the initial guess for density in molecular calculations. We found that using superposition of atomic densities (SAD) for this guess improves the convergence rates dramatically. Indeed it is more efficient to conduct atomic optimisations to generate the guess then run the molecular calculation compared to just running a molecular calculation using a random guess. The reason for this was explored by examining the structure of the Hessian in the initial stages of the calculation. We found that by using a SAD guess the Hessian structure is very close to what it should be at the convergence point, that of a first order saddle point, meaning the TRIM scheme rapidly becomes a simple Newton step method. The TRIM scheme is suitable to be used for assisting the development of accurate OF-KEFs and investigating desirable properties such as potentials.

In Chapter 7 we investigated a proposed variation principle from Gidopoulos⁶ for the determination of Kohn–Sham effective potentials was examined and extended to arbitrary electron-interaction strengths and to mixed states. Comparisons were drawn with Lieb's convex-conjugate functional, which allows for the determination of a potential associated with a given electron density by maximization, yielding the Kohn–Sham potential for a non-interacting system. The mathematical structure of the two functionals was shown to be intrinsically related; the variation principle put forward by Gidopoulos was shown to be expressed in terms of the Lieb functional. The equivalence between the information obtained from the two approaches was illustrated numerically by their implementation in a common framework.

Bibliography

- J. Lehtomäki, I. Makkonen, M. Caro, A. Harju and O. Lopez-Acevedo, *J. Chem. Phys.*, 2014, 141, 234102.
- [2] G. K.-L. Chan, A. J. Cohen and N. C. Handy, J. Chem. Phys., 2001, 114, 631.
- [3] M. S. Ryley, M. Withnall and A. M. Teale, In preparation.
- [4] M. S. Ryley, T. J. P. Irons, T. Helgaker and A. M. Teale, In preparation.
- [5] W. Kohn and L. Sham, *Phys. Rev.*, 1965, A140, 1113.
- [6] N. I. Gidopoulos, *Phys. Rev.* A, 2011, 83, 040502.
- T. J. P. Irons, J. W. Furness, M. S. Ryley, J. Zemen, T. Helgaker and A. M. Teale, *J. Chem. Phys.*, 2017, 147, 134107.
- [8] E. H. Lieb, Int. J. Quantum Chem., 1983, 24, 243.
- [9] W. A. Rudin, *Principles of Mathematical Analysis*, McGrawhill New York, 1964, vol. 3.
- [10] W. Rudin, *Functional Analysis*, McGraw-hill, 1991.
- [11] H. S. Bear, *A primer of Lebesgue integration*, Elsevier, 2001.
- [12] W. Rudin, *Real and complex analysis*, Tata McGraw-hill education, 2006.
- [13] T. Helgaker, P. Jorgensen and J. Olsen, *Principles of Density Functional Theory*, In production.
- [14] H. Eschrig, The fundamentals of density functional theory, Springer, 1996.
- [15] R. T. Rockafellar, *Convex analysis*, Princeton university press, 1970, vol. 28.

- [16] I. Ekeland and R. Temam, Convex analysis and variational problems, Siam, 1999, vol. 28.
- [17] D. R. Hartree, Math. Proc. Camb. Phil. Soc., 1928, 24, 89.
- [18] V. Fock, *Zeitschrift für Physik*, 1930, **61**, 126.
- [19] E. H. Lieb and B. Simon, *Comm. Math. Phys.*, 1977, 53, 185.
- [20] P.-L. Lions, Comm. Math. Phys., 1987, 109, 33.
- [21] E. Cances and C. Le Bris, ESAIM: Mathematical Modelling and Numerical Analysis, 2000, 34, 749.
- [22] P. Pulay, Chem. Phys. Lett., 1980, 73, 393.
- [23] P. Pulay, J. Comput. Chem., 1982, 3, 556.
- [24] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.*, 2007, 79, 291.
- [25] F. Coester, *Nuc. Phys.*, 1958, 7, 421.
- [26] F. Coester and H. Kümmel, *Nuc. Phys.*, 1960, 17, 477.
- [27] H. Kümmel, *Nuc. Phys.*, 1961, **22**, 177.
- [28] H. Kümmel, K. H. Lührmann and J. G. Zabolitzky, *Phys. Rep.*, 1978, 36, 1.
- [29] J. Cížek, J. Chem. Phys., 1966, 45, 4256.
- [30] J. Paldus, J. Cížek and I. Shavitt, *Phys. Rev.* A, 1972, 5, 50.
- [31] O. Sinanoglu, J. Chem. Phys., 1962, 36, 706.
- [32] K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.*, 1989, 157, 479.
- [33] P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, 864.
- [34] C. F. von Weizsäcker, Z. Phys., 1935, 96, 431.
- [35] M. Levy, Proc. Natl. Accad. Sci. USA, 1979, 76, 6062.
- [36] M. Levy, *Phys. Rev.* A, 1982, 26, 1200.
- [37] E. Lieb, Int. J. Quantum Chem., 1983, 24, 243.

- [38] S. Kvaal, U. Ekström, A. M. Teale and T. Helgaker, J. Chem. Phys., 2014, 140, 18A518.
- [39] N. Mardirossian and M. Head-Gordon, *Mol. Phys.*, 2017, **115**, 2315.
- [40] P. A. M. Dirac, Proc. Camb. Phil. Soc., 1930, 26, 376.
- [41] J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, 77, 3865.
- [42] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, 46, 6671.
- [43] E. Engel and S. Vosko, *Phys. Rev.* A, 1993, 47, 2800.
- [44] A. D. Becke, J. Chem. Phys., 1992, 96, 2155.
- [45] B. G. Johnson, P. M. Gill and J. A. Pople, J. Chem. Phys., 1993, 98, 5612.
- [46] A. D. Becke, J. Chem. Phys., 1986, 84, 4524.
- [47] S. F. Sousa, P. A. Fernandes and M. J. Ramos, J. Phys. Chem. A, 2007, 111, 10439.
- [48] M. Levy, Int. J. Quantum Chem., 1989, 36, 617.
- [49] M. Levy and J. P. Perdew, *Phys. Rev.* A, 1985, 32, 2010.
- [50] C. Bowen, G. Sugiyama and B. Alder, *Phys. Rev. B*, 1994, 50, 14838.
- [51] S. Moroni, D. M. Ceperley and G. Senatore, *Phys. Rev. Lett.*, 1995, **75**, 689.
- [52] G. Ortiz, *Phys. Rev. B*, 1992, 45, 11328.
- [53] E. H. Lieb and S. Oxford, *Int. J. Quantum Chem.*, 1981, **19**, 427.
- [54] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley-Blackwell, 2000.
- [55] A. Sodt, J. E. Subotnik and M. Head-Gordon, J. Chem. Phys., 2006, 125, 194109.
- [56] H. Werner, F. R. Manby and P. J. Knowles, J. Chem. Phys., 2003, 118, 8149.

- [57] C. Ochsenfeld, J. Kussmann and D. S. Lambrecht, *Rev. Comp. Chem.*, 2007, 23, 1.
- [58] C. Skylaris, P. D. Haynes, A. A. Mostofi and M. C. Payne, J. Chem. Phys., 2005, 122, 084119.
- [59] D. Bowler, T. Miyazaki and M. Gillan, J. Phys.: Condensed Matter, 2002, 14, 2781.
- [60] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker and T. Deutsch, *Phys. Chem. Chem. Phys.*, 2015, 17, 31360.
- [61] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón and D. Sánchez-Portal, *J. Phys: Condensed Matter*, 2002, 14, 2745.
- [62] J. J. Eriksen, P. Baudin, P. Ettenhuber, K. Kristensen, T. Kjærgaard and P. Jørgensen, J. Chem. Theo. Comp., 2015, 11, 2984.
- [63] T. S. Chwee, A. B. Szilva, R. Lindh and E. A. Carter, J. Chem. Phys., 2008, 128, 224106.
- [64] E. Hernández, M. Gillan and C. Goringe, *Phys. Rev. B*, 1996, 53, 7147.
- [65] M. D. Fabian, B. Shpiro, E. Rabani, D. Neuhauser and R. Baer, arXiv preprint arXiv:1809.08307, 2018.
- [66] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev and F. Neese, J. Chem. Phys., 2017, 146, 164105.
- [67] L. Hung and E. A. Carter, *Chem. Phys. Lett.*, 2009, 475, 165.
- [68] L. H. Thomas, Proc. Camb. Phil. Soc., 1930, 23, 376.
- [69] E. Fermi, *Rend. Accad. Lincei.*, 1927, 6, 602.
- [70] E. H. Lieb, *Rev. Mod. Phys.*, 1981, **53**, 603.
- [71] R. Parr and W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, 1st edn., 1994.
- [72] E. Teller, *Rev. Mod. Phys.*, 1962, **34**, 627.
- [73] J. Sheldon, *Phys. Rev.*, 1955, **99**, 1291.

- [74] E. H. Lieb and B. Simon, *Adv. Math.*, 1977, 23, 22.
- [75] N. L. Balàzs, *Phys. Rev.*, 1967, **156**, 42.
- [76] M. Chen, X.-W. Jiang, H. Zhuang, L.-W. Wang and E. A. Carter, J. Chem. Theory Comput., 2016, 2950.
- [77] W. Jones and W. Young, J. Phys. C: Solid State Physics, 1971, 4, 1322.
- [78] J. P. Perdew, V. Sahni, M. K. Harbola and R. K. Pathak, *Phys. Rev. B*, 1986, 34, 686.
- [79] M. Brack, J. B. K. and C. Y. H., *Phys. Lett.*, 1976, 65B, 1.
- [80] C. H. Hodges, *Can. J. Phys.*, 1973, **51**, 1428.
- [81] E. Gross and R. Dreizler, *Phys. Rev.* A, 1979, 20, 1798.
- [82] G. S. Ho, V. L. Lignères and E. A. Carter, Comput. Phys. Commun., 2008, 179, 839.
- [83] L. Wang and M. P. Teter, *Phys. Rev. B*, 1992, 45, 13196.
- [84] Y. A. Wang, N. Govind and E. A. Carter, *Phys. Rev. B*, 1998, 58, 13465.
- [85] Y. A. Wang, N. Govind and E. A. Carter, *Phys. Rev. B*, 1999, 60, 16350.
- [86] Y. A. Wang and E. A. Carter, in *Theoretical methods in con*densed phase chemistry, Springer, 2002, p. 117.
- [87] L. Hung and E. A. Carter, Chem. Phys. Lett., 2009, 475, 163.
- [88] I. Shin, A. Ramasubramaniam, C. Huang, L. Hung and E. A. Carter, *Philosophical Magazine*, 2009, 89, 3195.
- [89] Q. Peng, X. Zhang, L. Hung, E. A. Carter and G. Lu, *Phys. Rev.* B, 2008, 78, 054118.
- [90] L. Hung and E. A. Carter, J. Phys. Chem. C, 2011, 115, 6269.
- [91] X. Blanc and E. Cancès, J. Chem. Phys., 2005, 122, 214106.
- [92] D. J. Lacks and R. G. Gordon, J. Chem. Phys., 1994, 100, 4446.
- [93] A. Lembarki and H. Chermette, *Phys. Rev.* A, 1994, 50, 5328.

- [94] F. Tran and T. A. Wesołowski, Intl. J. Quantum Chem., 2002, 89, 441.
- [95] A. J. Thakkar, Phys. Rev. A, 1992, 46, 6920.
- [96] V. V. Karasiev, D. Chakraborty, O. A. Shukruto and S. B. Trickey, *Phys. Rev. B*, 2013, 88, 161108.
- [97] V. Karasiev, R. Jones, S. Trickey and F. Harris, *Phys. Rev. B*, 2009, 80, 245120.
- [98] F. W. Averill and G. S. Painter, *Phys. Rev. B*, 1981, 24, 6795.
- [99] V. Karasiev, S. Trickey and F. E. Harris, J. Comput.-Aided Mater. Des, 2006, 13, 111.
- [100] S. Liu and R. G. Parr, *Chem. Phys. Lett.*, 1997, 278, 341.
- [101] T. Gál, *Phys. Rev.* A, 2000, 62, 044501.
- [102] G. K.-L. Chan and N. C. Handy, *Phys. Rev.* A, 1999, 59, 2670.
- [103] E. H. Lieb and W. E. Thirring, *Phys. Rev. Lett.*, 1975, **35**, 687.
- [104] A. Borgoo, A. M. Teale and D. J. Tozer, J. Chem. Phys., 2012, 136, 034101.
- [105] J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz Jr, *Phys. Rev. Lett.*, 1982, 49, 1691.
- [106] A. Borgoo, A. M. Teale and D. J. Tozer, *Phys. Chem. Chem. Phys.*, 2014, 16, 14578.
- [107] A. Borgoo and D. J. Tozer, J. Chem. Theory Comput., 2013, 9, 2250.
- [108] H. Ou-Yang and M. Levy, *Phys. Rev.* A, 1990, 42, 155.
- [109] H. Ou-Yang and M. Levy, Intl. J. Quantum Chem., 1991, 40, 379.
- [110] J. P. Perdew, *Phys. Lett.* A, 1992, 165, 79.
- [111] A. E. DePristo and J. D. Kress, *Phys. Rev.* A, 1987, 35, 438.
- [112] M. Ernzerhof, J. Mol. Struct. (THEOCHEM), 2000, 501, 59.

- [113] U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen and K. Ruud, J. Chem. Theory Comput., 2010, 6, 1971.
- [114] W. Mi, X. Shao, C. Su, Y. Zhou, S. Zhang, Q. Li, H. Wang, L. Zhang, M. Miao, Y. Wang and Y. Ma, *Comp. Phys. Commun.*, 2016, 200, 87.
- [115] M. Levy, J. P. Perdew and V. Sahni, *Phys. Rev.* A, 1984, 30, 2745.
- [116] QUEST, a rapid development platform for QUantum Electronic Structure Techniques, 2017.
- [117] T. Rohwedder and R. Schneider, J. Math. Chem., 2011, 49, 9.
- [118] S. Høst, J. Olsen, B. Jansik, L. Thøgersen, P. Jørgensen and T. Helgaker, J. Chem. Phys., 2008, 129, 124106.
- [119] L. E. Leal, A. Karpenko, M. Caro and O. Lopez-Acevedo, *Phys. Chem. Chem. Phys.*, 2015, 17, 31463.
- [120] R. F. Nalewajski and R. G. Parr, J. Chem. Phys., 1982, 77, 399.
- [121] T. H. Dunning, J. Chem. Phys., 1989, 90, 1007.
- [122] D. E. Woon and T. H. Dunning, J. Chem. Phys., 1994, 100, 2975.
- [123] D. E. Woon and T. H. Dunning, J. Chem. Phys., 1993, 98, 1358.
- [124] R. A. Kendall, T. H. Dunning Jr and R. J. Harrison, J. Chem. Phys., 1992, 96, 6796.
- [125] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning and A. K. Wilson, *Theor. Chem. Acc.*, 2011, **128**, 69.
- [126] J. Nocedal and S. Wright, *Numerical Optimization*, Springer Science & Business Media, 2006.
- [127] A. Tarantola, Inverse problem theory and methods for model parameter estimation, siam, 2005, vol. 89.
- [128] W. Yang, *Phys. Rev.* A, 1986, 34, 4575.
- [129] R. A. King and N. C. Handy, Phys. Chem. Chem. Phys., 2000, 2, 5049.

- [130] "NIST Computational Chemistry Comparison and Benchmark Database: NIST Standard Reference Database Number 101, Editor:Russell D. Johnson III", "2018".
- [131] T. Helgaker, Chem. Phys. Lett., 1991, 182, 503.
- [132] C. M. Smith, *Theoretica chimica acta*, 1988, 74, 85.
- [133] C. M. Smith, International Journal of Quantum Chemistry, 1990, 37, 773.
- [134] T. Helgaker, K. Ruud and P. R. Taylor, in *The Reaction Path in Chemistry: Current Approaches and Perspectives*, Springer, 1995, p. 109.
- [135] E. R. Davidson, Reduced Density Matrices in Quantum Chemistry, Elsevier BV, 1976, vol. 6.
- [136] M. Levy, Proceedings of the National Academy of Sciences, 1979, 76, 6062.
- [137] M. Levy and J. P. Perdew, Int. J. Quantum Chem., 1982, 21, 511.
- [138] S. Kvaal and T. Helgaker, J. Chem. Phys., 2015, 143, 1184106.
- [139] P. E. Lammert, International Journal of Quantum Chemistry, 2007, 107, 1943.
- [140] D. Langreth and J. Perdew, *Solid State Commun.*, 1975, 17, 1425.
- [141] D. C. Langreth and J. P. Perdew, *Phys. Rev. B*, 1977, 15, 2884.
- [142] O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B*, 1976, 13, 4274.
- [143] O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B*, 1977, 15, 6006.
- [144] W. Yang, J. Chem. Phys., 1998, 109, 10107.
- [145] J. Toulouse, F. Colonna and A. Savin, *Mol. Phys.*, 2005, 103, 2725.
- [146] A. M. Teale, S. Coriani and T. Helgaker, J. Chem. Phys., 2010, 133, 164112.
- [147] H. Hellmann, in Hans Hellmann: Einführung in die Quantenchemie, Springer Science + Business Media, 2015, p. 19.

- [148] R. P. Feynman, *Phys. Rev.*, 1939, 56, 340.
- [149] A. Savin, F. Colonna and R. Pollet, Int. J. Quantum Chem., 2003, 93, 166.
- [150] F. Colonna and A. Savin, J. Chem. Phys., 1999, 110, 2828.
- [151] A. M. Teale, S. Coriani and T. Helgaker, J. Chem. Phys., 2009, 130, 104111.
- [152] A. M. Teale, S. Coriani and T. Helgaker, J. Chem. Phys., 2010, 132, 164115.
- [153] W. Yang and Q. Wu, *Phys. Rev. Lett.*, 2002, **89**, 143002.
- [154] Q. Wu and W. Yang, J. Chem. Phys., 2003, 118, 2498.
- [155] E. Fermi and E. Amaldi, R. Accad. d'Italia. Memorie, 1934, 6, 119.
- [156] R. A. Kendall, T. H. Dunning and R. J. Harrison, J. Chem. Phys., 1992, 96, 6796.
- [157] Q. Wu and W. Yang, J. Theor. Comp. Chem., 2003, 2, 627.
- [158] S. K. Lam, A. Pitrou and S. Seibert, Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, 2015, p. 7.
- [159] S. K. Lam, A. Pitrou and S. Seibert, Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC - LLVM'15, 2015.
- [160] M. Ernzerhof, J. P. Perdew and K. Burke, Int. J. Quantum Chem., 1997, 64, 285.
- [161] S. Laricchia, L. A. Constantin, E. Fabiano and F. Della Sala, J. Chem. Phys. Comp., 2013, 10, 164.
- [162] S. Smiga, E. Fabiano, L. A. Constantin and F. Della Sala, J. Chem. Phys., 2017, 146, 064105.
- [163] L. A. Constantin, E. Fabiano and F. Della Sala, J. Chem. Theo. Comp., 2017, 13, 4228.
- [164] J. Perdew and L. A. Constantin, *Phys. Rev. B*, 2007, 75, year.
- [165] N. L. Allan, C. G. West, D. L. Cooper, P. J. Grout and N. H. March, J. Chem. Phys., 1985, 83, 4562.

181

- [166] G. I. Plindov and S. K. Pogrebnya, Chem. Phys. Letts., 1988, 143, 535.
- [167] D. Lee, L. A. Constantin, J. P. Perdew and K. Burke, J. Chem. Phys., 2009, 130, 034107.
- [168] S. Liu and R. G. Parr, *Phys. Rev.* A, 1997, 55, 1792.
- [169] R. G. Parr, S. Liu, A. A. Kugler and A. Nagy, *Phys. Rev. A*, 1995, 52, 969.
- [170] S. Liu and R. G. Parr, *Phys. Rev.* A, 1996, 53, 2211.
- [171] A. Görling and M. Levy, Phys. Rev. A, 1995, 52, 4493.
- [172] S. Liu, P. Süle, R. Lopez-Boada and A. Nagy, Chem. Phys. Lett., 1996, 257, 68.
- [173] A. Görling and M. Levy, *Phys. Rev. B*, 1993, 47, 13105.
- [174] E. V. Ludeña, E. X. Salazar, M. H. Cornejo, D. E. Arroyo and V. V. Karasiev, *Int. J. Quant. Chem.*, 2018, **118**, e25601.
- [175] D. J. Sharpe, M. Levy and D. J. Tozer, J. Chem. Theo. Comp., 2018, 14, 684.
- [176] M. Rupp, A. Tkatchenko, K.-R. Müller and O. A. Von Lilienfeld, *Phys. Rev. Lett.*, 2012, **108**, 058301.
- [177] K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. Von Lilienfeld, A. Tkatchenko and K.-R. Müller, J. Chem. Theo. Comp., 2013, 9, 3404.
- [178] G. R. Schleder, A. C. M. Padilha, C. M. Acosta, M. Costa and A. Fazzio, *J. Phys. Mater.*, 2019, 2, 032001.
- [179] K. Vu, J. C. Snyder, L. Li, M. Rupp, B. F. Chen, T. Khelif, K.-R. Müller and K. Burke, *Int. J. Quantum Chem.*, 2015, **115**, 1115.
- [180] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller and K. Burke, *Phys. Rev. Lett.*, 2012, **108**, 253002.
- [181] L. Li, J. C. Snyder, I. M. Pelascher, J. Huang, U.-N. Niranjan, P. Duncan, M. Rupp, K.-R. Müller and K. Burke, *Int. J. Quantum Chem.*, 2016, 116, 819.

- [182] F. Brockherde, L. Vogt, L. Li, M. E. Tuckerman, K. Burke and K.-R. Müller, *Nature communications*, 2017, 8, 872.
- [183] P. Golub and S. Manzhos, Phys. Chem. Chem. Phys., 2019, 21, 378.
- [184] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe and H. van der Vorst, Templates for the solution of algebraic eigenvalue problems: a practical guide, SIAM, 2000.

A. Equations for the gradient and Hessians used in this work

A.1. Preliminaries

We wish to optimise the following Lagrangian:

$$\mathcal{L}[\rho(\mathbf{r})] = E[\rho(\mathbf{r})] - \mu N, \qquad (A.1)$$

where E is the total energy, μ is the chemical potential and N is the number electrons associated with the current ground state density. We expand the square root of the density to enforce positivity,

$$\rho^{\frac{1}{2}}(\mathbf{r}) = N_0^{\frac{1}{2}} \sum_k c_k \eta_k(\mathbf{r}).$$
(A.2)

Where N_0 is the correct number of electrons for a given system. From now on the dependence on **r** will be implied, not explicitly stated. Our goal is to optimise the Lagrangian, for a fixed chemical potential, to find a set of optimised coefficients $\{c_{opt}\}$. We required analytic expressions for the Hessians and gradients of the Lagrangian with respect to the expansion coefficients, c_i .

A.2. Gradient

The chemical potential , μ , is not constructed using a basis and so the variation of the Lagrangian with respect to the basis is given by

$$\frac{\partial \mathcal{L}}{\partial c_i} = \frac{\partial E}{\partial c_i} - \mu \frac{\partial N}{\partial c_i}.$$
 (A.3)

We can also write, using Equation (A.2),

$$N = \int \rho^{\frac{1}{2}} \rho^{\frac{1}{2}} \mathrm{d}\mathbf{r} = N_0 \int \sum_k c_k \eta_k \sum_l c_l \eta_l \mathrm{d}\mathbf{r}.$$
 (A.4)

This means

$$\frac{\partial N}{\partial c_i} = N_0 \int \left(\eta_i \sum_k c_k \eta_k + \eta_i \sum_l c_l \eta_l \right) d\mathbf{r} = 2N_0 \int \eta_i \rho^{\frac{1}{2}} d\mathbf{r}.$$
(A.5)

The derivative of the energy can be split as the following,

$$\frac{\partial E}{\partial c_i} = \frac{\partial T_{\rm s}}{\partial c_i} + \frac{\partial E_{\rm XC}}{\partial c_i} + \frac{\partial E_{\rm J}}{\partial c_i} + \frac{\partial}{\partial c_i} \int \mathrm{d}\mathbf{r} \upsilon_{\rm ext} N_0^{\frac{1}{2}} \sum_k c_k \eta_k N_0^{\frac{1}{2}} \sum_l c_l \eta_l, \quad (A.6)$$

with $T_{\rm s}$ the non-interacting kinetic energy, $E_{\rm XC}$ the exchange-correlation energy, $E_{\rm J}$ the Coulomb/Hartree energy and $v_{\rm ext}$ the electron-nuclei attraction.

Equation (A.6) can be re-written as,

$$\int \frac{\delta T_{\rm s}}{\delta \rho} \frac{\partial \rho}{\partial c_i} d\mathbf{r} + \int \frac{\delta E_{\rm XC}}{\delta \rho} \frac{\partial \rho}{\partial c_i} d\mathbf{r} + \int \frac{\delta E_{\rm J}}{\delta \rho} \frac{\partial \rho}{\partial c_i} d\mathbf{r} + 2N_0 \int \eta_i \upsilon_{\rm ext} \rho^{\frac{1}{2}} d\mathbf{r}. \quad (A.7)$$

We now require the following derivative,

$$\frac{\partial \rho}{\partial c_i} = \frac{\partial}{\partial c_i} N_0 \sum_k c_k \eta_k \sum_l c_l \eta_l = 2N_0 \eta_i \rho^{\frac{1}{2}}.$$
 (A.8)

Combining Equations (A.3), (A.5), (A.7) and (A.8) one gets the final and implemented form of the analytic gradient,

$$\frac{\partial \mathcal{L}}{\partial c_i} = 2N_0 \int \eta_i \left(\frac{\delta T_{\rm s}}{\delta \rho} + \frac{\delta E_{\rm XC}}{\delta \rho} + \frac{\delta E_{\rm J}}{\delta \rho} + \upsilon_{\rm ext} - \mu \right) \rho^{\frac{1}{2}} \mathrm{d}\mathbf{r}.$$
(A.9)

A.3. Hessian

The Hessian is:

$$\frac{\partial^{2} \mathcal{L}}{\partial c_{i} \partial c_{j}} = \frac{\partial^{2} T_{s}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{XC}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2} E_{J}}{\partial c_{i} \partial c_{j}} + \frac{\partial^{2}}{\partial c_{i} \partial c_{j}} \int \mathrm{d}\mathbf{r} \upsilon_{\text{ext}} N_{0}^{\frac{1}{2}} \sum_{k} c_{k} \eta_{k} N_{0}^{\frac{1}{2}} \sum_{l} c_{l} \eta_{l} - \mu \frac{\partial^{2} N}{\partial c_{i} \partial c_{j}}$$
(A.10)

The final term in Equation (A.10) evaluates as:

$$-\mu N_0 \int d\mathbf{r} \frac{\partial}{\partial c_j} 2\eta_i \sum_k c_k \eta_k = -2\mu N_0 \int \eta_i \eta_j d\mathbf{r}.$$
 (A.11)

The second term from the right in Equation (A.10) evaluates similarly

$$2N_0 \int \upsilon_{\text{ext}} \eta_i \eta_j \mathrm{d}\mathbf{r}. \tag{A.12}$$

Now for the first of the slightly tricky terms, the third term in Equation (A.10). Firstly the Coulomb energy can be written as

$$E_{\mathrm{J}}[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r_{1}})\rho(\mathbf{r_{2}})}{|\mathbf{r_{1}} - \mathbf{r_{2}}|} \mathrm{d}\mathbf{r_{1}} \mathrm{d}\mathbf{r_{2}},$$

$$= \frac{N_{0}^{2}}{2} \sum_{l,k,m,n} c_{l}c_{k}c_{m}c_{n} \iint \frac{\eta_{l}(\mathbf{r_{1}})\eta_{k}(\mathbf{r_{1}})\eta_{m}(\mathbf{r_{2}})\eta_{n}(\mathbf{r_{2}})}{|\mathbf{r_{1}} - \mathbf{r_{2}}|} \mathrm{d}\mathbf{r_{1}} \mathrm{d}\mathbf{r_{2}}, \quad (A.13)$$

$$\equiv \frac{N_{0}^{2}}{2} \sum_{l,k,m,n} c_{l}c_{k}c_{m}c_{n}(lk|mn).$$

So now we can begin the differentiation of $E_{\rm J}$.

$$\frac{\partial^2 E_{\rm J}}{\partial c_i \partial c_j} = \frac{N_0^2}{2} \frac{\partial}{\partial c_i} \bigg[\eta_j \sum_{kmn} c_k c_m c_n(jk|mn) \\
+ \eta_j \sum_{lmn} c_l c_m c_n(lj|mn) + \eta_j \sum_{lkn} c_l c_k c_n(lk|jn) + \eta_j \sum_{lkm} c_l c_k c_m(lk|mj) \\
= N_0^2 \frac{\partial}{\partial c_i} \bigg(\eta_j \sum_{kmn} c_k c_m c_n(jk|mn) + \eta_j \sum_{lkn} c_l c_k c_n(lk|jn) \bigg).$$
(A.14)

Evaluating the next derivative leads us to

$$\frac{\partial^{2} E_{\mathrm{J}}}{\partial c_{i} \partial c_{j}} = N_{0}^{2} \bigg[\eta_{i} \eta_{j} \sum_{mn} c_{m} c_{n}(ji|mn) + \eta_{i} \eta_{j} \sum_{km} c_{k} c_{n}(jk|mi) + \eta_{j} \eta_{i} \sum_{kn} c_{k} c_{n}(jk|im) \\ \eta_{i} \eta_{j} \sum_{kn} c_{k} c_{n}(ik|jn) + \eta_{i} \eta_{j} \sum_{kl} c_{k} c_{l}(kl|ji) + \eta_{i} \eta_{j} \sum_{ln} c_{l} c_{n}(li|jn) \bigg] \\ = 2N_{0}^{2} \eta_{i} \eta_{j} \bigg[\sum_{kl} c_{k} c_{l}(ij|kl) + 2\sum_{kl} c_{k} c_{l}(ik|jl) \bigg].$$
(A.15)

The only two terms that remain are the second-derivatives of the noninteracting kinetic energy functional and the exchange-correlation functional. The form of these two change whether they are LDA,GGA or Laplacian level. So, we can write a generic functional as

$$F[\rho, \nabla \rho, \nabla^2 \rho] = \int f(\rho, \nabla \rho, \nabla^2 \rho) d\mathbf{r}.$$
 (A.16)

M. S. Ryley

Variational Solutions in OF-DFT

187

A.3.1. LDA

Using the chain rule and the product rule one finds that:

$$\frac{\partial}{\partial c_j} \left(\frac{\partial F}{\partial c_i} \right) = \int \left(\frac{\partial \rho}{\partial c_i} \left(\frac{\partial}{\partial c_j} \frac{\partial f(\rho)}{\partial \rho(\mathbf{r})} \right) + \frac{\partial f(\rho)}{\partial \rho(\mathbf{r})} \left(\frac{\partial}{\partial c_j} \frac{\partial \rho}{\partial c_i} \right) \right) d\mathbf{r}.$$
(A.17)

Applying the chain rule again gives,

$$\frac{\partial^2 F[\rho]}{\partial c_i \partial c_j} = \int \frac{\partial \rho}{\partial c_i} \frac{\partial^2 f(\rho)}{\partial \rho(\mathbf{r})^2} \frac{\partial \rho}{\partial c_j} + \frac{\partial f(\rho)}{\partial \rho(\mathbf{r})} \left(\frac{\partial^2 \rho}{\partial c_j \partial c_i}\right) \mathrm{d}\mathbf{r}.$$
 (A.18)

Now using Equation (A.2) we can write

$$\frac{\partial \rho}{\partial c_m} = \frac{\partial}{\partial c_m} \left(N_0 \sum_x c_x \eta_x \sum_y c_y \eta_y \right) = 2N_0 \eta_m \sum_k c_k \eta_k.$$
(A.19)

In addition

$$\frac{\partial^2 \rho}{\partial c_n \partial c_m} = \frac{\partial}{\partial c_n} \left(2N_0 \eta_m \sum_k c_k \eta_k \right) = 2N_0 \eta_m \eta_n.$$
(A.20)

So we now find that

$$\frac{\partial^2 F[\rho]}{\partial c_i \partial c_j} = \int 4N_0^2 \eta_i \eta_j \frac{\partial^2 f(\rho)}{\partial \rho^2} \sum_l c_l \eta_l \sum_k c_l \eta_k + 2 \frac{\partial f(\rho)}{\partial \rho} \eta_i \eta_j(\mathbf{r}) d\mathbf{r}$$

$$= \left\langle \eta_i \left| 4N_0^2 \frac{\partial^2 f(\rho)}{\partial \rho^2} \sum_l c_l \eta_l \sum_k c_l \eta_k + 2N_0 \frac{\partial f(\rho)}{\partial \rho} \right| \eta_j \right\rangle$$
(A.21)

In order to cast the derivatives in a form which is callable from XCFun, we do the following:

$$\frac{\partial f}{\partial \rho} = \frac{\partial f}{\partial \rho_{\alpha}} \frac{\partial \rho_{\alpha}}{\partial \rho} + \frac{\partial f}{\partial \rho_{\beta}} \frac{\partial \rho_{\beta}}{\partial \rho} = \frac{1}{2} \left[\frac{\partial f}{\partial \rho_{\alpha}} + \frac{\partial f}{\partial \rho_{\beta}} \right], \tag{A.22}$$

$$\frac{\partial^2 f}{\partial \rho^2} = \frac{1}{4} \left[\frac{\partial^2 f}{\partial \rho_{\alpha}^2} + 2 \frac{\partial^2 f}{\partial \rho_{\alpha} \partial \rho_{\beta}} + \frac{\partial^2 f}{\partial \rho_{\beta}^2} \right].$$
(A.23)

A.3.2. GGA

For a GGA, which is dependent on ρ and $\nabla \rho$. One can write

$$\frac{\partial^2 F[\rho]}{\partial c_i \partial c_j} = \frac{\partial}{\partial c_i} \left\langle \frac{\partial \rho}{\partial c_j} \frac{\partial f}{\partial \rho} + \frac{\partial \nabla \rho}{\partial c_j} \frac{\partial f}{\partial \nabla \rho} \right\rangle, \tag{A.24}$$

where we have used the following notation

$$\langle f \rangle := \int f(\mathbf{r}) \mathrm{d}\mathbf{r}.$$

The right hand side of Equation (A.24) expands to

$$\left\langle \frac{\partial^2 \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \rho} + \frac{\partial \rho}{\partial c_j} \frac{\partial}{\partial c_i} \left[\frac{\partial f}{\partial \rho} \right] + \frac{\partial^2 \nabla \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \nabla \rho} + \frac{\partial \nabla \rho}{\partial c_j} \frac{\partial}{\partial c_i} \left[\frac{\partial f}{\partial \nabla \rho} \right] \right\rangle. \quad (A.25)$$

With some manipulation one arrives at

$$\left\langle \frac{\partial^2 \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \rho} + \frac{\partial^2 \nabla \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \nabla \rho} + \frac{\partial^2 f}{\partial \rho^2} \frac{\partial \rho}{\partial c_i} \frac{\partial \rho}{\partial c_j} \right\rangle + \left\langle \frac{\partial^2 f}{\partial \nabla \rho^2} \frac{\partial \nabla \rho}{\partial c_i} \frac{\partial \nabla \rho}{\partial c_j} + \frac{\partial^2 f}{\partial \nabla \rho \partial \rho} \left[\frac{\partial \rho}{\partial c_j} \frac{\partial \nabla \rho}{\partial c_i} + \frac{\partial \rho}{\partial c_i} \frac{\partial \nabla \rho}{\partial c_j} \right] \right\rangle.$$
(A.26)

The first term in Equation (A.26) is very straightforward as we have derived it already for the LDA. It is given by

$$\left\langle \frac{\partial^2 \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \rho} \right\rangle = 2N_0 \left\langle \eta_i \right| \frac{\partial f}{\partial \rho} \left| \eta_j \right\rangle.$$
(A.27)

Before going any further we need to note three things. Firstly,

$$\nabla \rho = N_0 \nabla \left(\sum_l c_l \eta_l(\mathbf{r}) \sum_k c_k \eta_k(\mathbf{r}) \right)$$

= $N_0 \left(\sum_k c_k \eta_k(\mathbf{r}) \nabla \left(\sum_l c_l \eta_l(\mathbf{r}) \right) + \sum_l c_l \eta_l(\mathbf{r}) \nabla \left(\sum_k c_k \eta_k(\mathbf{r}) \right) \right)$
= $2N_0 \sum_k c_k \nabla \eta_k(\mathbf{r}) \sum_l c_l \eta_l(\mathbf{r}).$ (A.28)

Secondly,

$$\frac{\partial \boldsymbol{\nabla} \rho}{\partial c_j} = 2N_0 \left(\boldsymbol{\nabla} \eta_j(\mathbf{r}) \sum_l c_l \eta_l(\mathbf{r}) + \eta_j \sum_k c_k \boldsymbol{\nabla} \eta_k(\mathbf{r}) \right),$$

$$\frac{\partial^2 \boldsymbol{\nabla} \rho}{\partial c_j \partial c_i} = 2N_0 (\eta_i(\mathbf{r}) \boldsymbol{\nabla} \eta_j(\mathbf{r}) + \eta_j(\mathbf{r}) \boldsymbol{\nabla} \eta_i(\mathbf{r}))$$

$$= 2N_0 \boldsymbol{\nabla} (\eta_i(\mathbf{r}) \eta_j(\mathbf{r})).$$
(A.29)

And finally that XCFun does not return derivatives with respect to gradients but returns derivatives with respect to these two quantities,

$$G_{\alpha\beta} = \boldsymbol{\nabla} \rho_{\alpha} \cdot \boldsymbol{\nabla} \rho_{\beta}$$

$$G_{\alpha\alpha} = \boldsymbol{\nabla} \rho_{\alpha} \cdot \boldsymbol{\nabla} \rho_{\alpha}.$$
(A.30)

The second term from Eq. (A.26) requires,

$$\frac{\partial f}{\partial \boldsymbol{\nabla} \rho} = \frac{\partial f}{\partial \boldsymbol{\nabla} \rho_{\alpha}} \frac{\partial \boldsymbol{\nabla} \rho_{\alpha}}{\partial \boldsymbol{\nabla} \rho} + \frac{\partial f}{\partial \boldsymbol{\nabla} \rho_{\beta}} \frac{\partial \boldsymbol{\nabla} \rho_{\beta}}{\partial \boldsymbol{\nabla} \rho} = \frac{1}{2} \left(\frac{\partial f}{\partial \boldsymbol{\nabla} \rho_{\alpha}} + \frac{\partial f}{\partial \boldsymbol{\nabla} \rho_{\beta}} \right)$$

$$= \frac{1}{2} \left[\frac{\partial G_{\alpha\alpha}}{\partial \boldsymbol{\nabla} \rho_{\alpha}} \frac{\partial f}{\partial G_{\alpha\alpha}} + \frac{\partial G_{\alpha\beta}}{\partial \boldsymbol{\nabla} \rho_{\alpha}} \frac{\partial f}{\partial G_{\alpha\beta}} + \frac{\partial G_{\beta\beta}}{\partial \boldsymbol{\nabla} \rho_{\beta}} \frac{\partial f}{\partial G_{\beta\beta}} + \frac{\partial G_{\alpha\beta}}{\partial \boldsymbol{\nabla} \rho_{\beta}} \frac{\partial f}{\partial G_{\alpha\beta}} \right]$$

$$= \frac{1}{2} \left[2 \boldsymbol{\nabla} \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha\alpha}} + \boldsymbol{\nabla} \rho_{\beta} \frac{\partial f}{\partial G_{\alpha\beta}} + 2 \boldsymbol{\nabla} \rho_{\beta} \frac{\partial f}{\partial G_{\beta\beta}} + \boldsymbol{\nabla} \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha\beta}} \right]$$

$$= \frac{\boldsymbol{\nabla} \rho}{2} \left[\frac{\partial f}{\partial G_{\alpha\alpha}} + \frac{\partial f}{\partial G_{\alpha\beta}} + \frac{\partial f}{\partial G_{\beta\beta}} \right]$$
(A.31)

and so,

$$\left\langle \frac{\partial^2 \boldsymbol{\nabla} \rho}{\partial c_i \partial c_j} \frac{\partial f}{\partial \boldsymbol{\nabla} \rho} \right\rangle = N_0 \left\langle \left[\frac{\partial f}{\partial G_{\alpha \alpha}} + \frac{\partial f}{\partial G_{\alpha \beta}} + \frac{\partial f}{\partial G_{\beta \beta}} \right] \boldsymbol{\nabla} \rho \cdot \boldsymbol{\nabla} (\eta_i(\mathbf{r}) \eta_j(\mathbf{r})) \right\rangle.$$
(A.32)

The third term of Eq. (A.26) can be taken from LDA functional derivative,

$$4N_0^2 \left\langle \eta_i \eta_j \sum_k c_k \eta_k \sum_l c_l \eta_l \frac{\partial^2 f}{\partial \rho^2} \right\rangle.$$
(A.33)

The fourth term from Eq. (A.26) requires some manipulation. Firstly,

$$\begin{split} \frac{\partial^2 f}{\partial \nabla \rho^2} &= \frac{\partial}{\partial \nabla \rho} \left[\frac{\nabla \rho}{2} \left[\frac{\partial f}{\partial G_{\alpha\alpha}} + \frac{\partial f}{\partial G_{\alpha\beta}} + \frac{\partial f}{\partial G_{\beta\beta}} \right] \right] \\ &= \frac{1}{2} \left(\frac{\partial f}{\partial G_{\alpha\alpha}} + \frac{\partial f}{\partial G_{\alpha\beta}} + \frac{\partial f}{\partial G_{\beta\beta}} \right) + \frac{\nabla \rho}{2} \left(\frac{\partial}{\partial \nabla \rho} \frac{\partial f}{\partial G_{\alpha\alpha}} + \frac{\partial}{\partial \nabla \rho} \frac{\partial f}{\partial G_{\alpha\beta}} + \frac{\partial}{\partial \nabla \rho} \frac{\partial f}{\partial G_{\beta\beta}} \right) \\ &= \frac{A}{2} + \frac{\nabla \rho}{4} \left(2 \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\alpha\alpha}^2} + \nabla \rho_{\beta} \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} + \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} \right. \\ &+ 2 \nabla \rho_{\beta} \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} + \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\beta\beta} \partial G_{\alpha\beta}} + 2 \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} \\ &+ \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\beta\beta}^2} + 2 \nabla \rho_{\beta} \frac{\partial^2 f}{\partial G_{\beta\beta} \partial G_{\alpha\beta}} + 2 \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} \\ &+ \nabla \rho_{\beta} \frac{\partial^2 f}{\partial G_{\beta\beta} \partial G_{\alpha\beta}} + \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\beta\beta} \partial G_{\alpha\beta}} + 2 \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial G_{\alpha\alpha}^2} \right) \\ &= \frac{A}{2} + \frac{\nabla \rho \cdot \nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha\alpha}^2} + \frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\beta\beta}^2} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} \right] \\ &+ 2 \frac{\partial^2 f}{\partial G_{\beta\beta} \partial G_{\alpha\beta}} \right] \\ &= \frac{A}{2} + \frac{\nabla \rho \cdot \nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha\alpha}^2} + \frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\beta\beta}^2} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} \right] \\ &= \frac{A}{2} + \frac{\nabla \rho \cdot \nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\beta\beta}^2} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} \right] \\ &= \frac{A}{2} + \frac{\nabla \rho \cdot \nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\alpha\beta}^2} + \frac{\partial^2 f}{\partial G_{\beta\beta}^2} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} + 2 \frac{\partial^2 f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}} \right] \\ &= \frac{A}{2} + \frac{\nabla \rho \cdot \nabla \rho}{4} D. \end{aligned}$$

Therefore,

$$\left\langle \frac{\partial^2 f}{\partial \boldsymbol{\nabla} \rho^2} \frac{\partial \boldsymbol{\nabla} \rho}{\partial c_j} \frac{\partial \boldsymbol{\nabla} \rho}{\partial c_i} \right\rangle = 2N_0^2 \left\langle \boldsymbol{\nabla} \left(\eta_j \sum_l c_l \eta_l \right) \cdot \boldsymbol{\nabla} \left(\eta_i \sum_l c_l \eta_l \right) \left[A + \frac{\boldsymbol{\nabla} \rho \cdot \boldsymbol{\nabla} \rho}{2} D \right] \right\rangle$$
(A.35)

The final part of Eq. (A.26), requires the following derivative,

$$\begin{aligned} \frac{\partial^2 f}{\partial \nabla \rho \partial \rho} &= \frac{\partial}{\partial \nabla \rho} \left[\frac{1}{2} \left(\frac{\partial f}{\partial \rho_{\alpha}} + \frac{\partial f}{\partial \rho_{\beta}} \right) \right] \\ &= \frac{1}{4} \left(2 \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\alpha\alpha}} \nabla \rho_{\alpha} + \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta\alpha}} \nabla \rho_{\beta} + \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta\alpha}} \nabla \rho_{\alpha} + \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta\alpha}} \nabla \rho_{\beta} \right] \\ &+ 2 \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\alpha\alpha}} \nabla \rho_{\alpha} + \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\alpha}} \nabla \rho_{\beta} + 2 \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\beta}} \nabla \rho_{\beta} + \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\alpha}} \nabla \rho_{\alpha} \right) \\ &= \frac{1}{4} \nabla \rho \left[\frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\alpha\alpha}} + \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\alpha}} + \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta\beta}} \right] \\ &+ \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\alpha\alpha}} + \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\alpha}} + \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta\beta}} \right] \\ &:= \frac{G}{4} \nabla \rho. \end{aligned}$$
(A.36)

Therefore

$$\left\langle \frac{\partial^2 f}{\partial \boldsymbol{\nabla} \rho \partial \rho} \left[\frac{\partial \rho}{\partial c_j} \frac{\partial \boldsymbol{\nabla} \rho}{\partial c_i} + \frac{\partial \rho}{\partial c_i} \frac{\partial \boldsymbol{\nabla} \rho}{\partial c_j} \right] \right\rangle = N_0^2 \left\langle G \boldsymbol{\nabla} \rho \cdot \left[\eta_i \sum_n c_n \eta_n \boldsymbol{\nabla} \left(\eta_j \sum_n c_n \eta_n \right) + \eta_j \sum_n c_n \eta_n \boldsymbol{\nabla} \left(\eta_i \sum_n c_n \eta_n \right) \right] \right\rangle.$$
(A.37)

A.3.3. Laplacian Level

We shall discuss the Hessian of Laplacian functional which are dependent on ρ , $\nabla \rho$ and $\nabla^2 \rho$. The general form of a Laplacian functional is given by

$$X[\rho] = \int f(\rho, \boldsymbol{\nabla}\rho, \boldsymbol{\nabla}^2 \rho) d\mathbf{r} = \left\langle f(\rho, \boldsymbol{\nabla}\rho, \boldsymbol{\nabla}^2 \rho) \right\rangle.$$
(A.38)

The first derivative with respect to variations of the expansion coefficients is

$$\frac{\partial X}{\partial c_j} = \left\langle \frac{\partial \rho}{\partial c_j} \frac{\partial f}{\partial \rho} + \frac{\partial \nabla \rho}{\partial c_j} \frac{\partial f}{\partial \nabla \rho} + \frac{\partial \nabla^2 \rho}{\partial c_j} \frac{\partial f}{\partial \nabla^2 \rho} \right\rangle.$$
(A.39)

The second derivative is

$$\frac{\partial^{2} f}{\partial c_{j} \partial c_{i}} = \frac{\partial^{2} \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \rho} + \frac{\partial \rho}{\partial c_{j}} \left[\frac{\partial \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \rho^{2}} + \frac{\partial \nabla \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla \rho \partial \rho} + \frac{\partial \nabla^{2} \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla^{2} \rho \partial \rho} \right]
+ \frac{\partial^{2} \nabla \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \rho} + \frac{\partial \nabla \rho}{\partial c_{j}} \left[\frac{\partial \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \rho \partial \nabla \rho} + \frac{\partial \nabla \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla \rho^{2}} + \frac{\partial \nabla^{2} \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla^{2} \rho \partial \nabla \rho} \right]
+ \frac{\partial^{2} \nabla^{2} \rho}{\partial c_{i} \partial c_{j}} \frac{\partial f}{\partial \rho} + \frac{\partial \nabla^{2} \rho}{\partial c_{j}} \left[\frac{\partial \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \rho \partial \nabla^{2} \rho} + \frac{\partial \nabla \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla^{2} \rho \partial \nabla \rho} + \frac{\partial \nabla^{2} \rho}{\partial c_{i}} \frac{\partial^{2} f}{\partial \nabla^{2} \rho \partial \nabla \rho} \right]
(A.40)$$

Only the fourth, eighth, ninth, tenth, eleventh and twelfth terms in Equation (A.40) are not known already from LDA and GGA level. We require knowledge of how the Laplacian functional is expressed with respect to the basis. We use the fact that

$$(fg)'' = f''g + 2f'g' + fg''$$
(A.41)

Then,

$$\nabla^{2} \rho = N \nabla^{2} \left(\sum_{k} c_{k} \eta_{k} \sum_{l} c_{l} \eta_{l} \right)$$

$$= N \sum_{m=1}^{3} \frac{\partial^{2}}{\partial x_{m}^{2}} \left(\sum_{k} c_{k} \eta_{k} \sum_{l} c_{l} \eta_{l} \right)$$

$$= N \sum_{m=1}^{3} \left[2 \sum_{k} c_{k} \frac{\partial^{2} \eta_{k}}{\partial x_{m}^{2}} \sum_{l} c_{l} \eta_{l} + 2 \sum_{k} c_{k} \frac{\partial \eta_{k}}{\partial x_{m}} \sum_{l} c_{l} \frac{\partial \eta_{l}}{\partial x_{m}} \right] \quad (A.42)$$

$$= 2N \sum_{l} \sum_{k} \sum_{m=1}^{3} c_{k} c_{l} \frac{\partial^{2} \eta_{k}}{\partial x_{m}^{2}} \eta_{l} + \sum_{m=1}^{3} c_{k} c_{l} \frac{\partial \eta_{k}}{\partial x_{m}} \frac{\partial \eta_{l}}{\partial x_{m}} \right]$$

$$= 2N \sum_{l} \sum_{k} c_{k} c_{l} \left[\nabla^{2} \eta_{k} \eta_{l} + \sum_{m=1}^{3} \frac{\partial \eta_{k}}{\partial x_{m}} \frac{\partial \eta_{l}}{\partial x_{m}} \right].$$

We will also need,

$$\frac{\partial \nabla^2 \rho}{\partial c_i} = 2N \left[\sum_l c_l \left(\nabla^2 \eta_i \eta_l + \sum_{m=1}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_l}{\partial x_m} \right) + \sum_k c_k \left(\nabla^2 \eta_i \eta_k + \sum_{m=1}^3 \frac{\partial \eta_k}{\partial x_m} \frac{\partial \eta_i}{\partial x_m} \right) \right] \\ = 2N \left[\sum_l c_l \nabla^2 \eta_i \eta_l + \sum_l c_l \sum_{m=1}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_l}{\partial x_m} + \sum_k c_k \nabla^2 \eta_k \eta_i + \sum_k c_k \sum_{m=1}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} \right] \\ = 2N \left[\sum_l c_l \nabla^2 \eta_i \eta_l + 2\sum_k c_k \sum_{m=1}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} + \sum_k c_k \nabla^2 \eta_k \eta_l \right].$$
(A.43)

For the fourth term we require the following derivative,

$$\frac{\partial^2 f}{\partial \rho \partial \nabla^2 \rho} = \frac{\partial}{\partial \rho} \frac{1}{2} \left[\frac{\partial f}{\partial \nabla^2 \rho_{\alpha}} + \frac{\partial f}{\partial \nabla^2 \rho_{\beta}} \right] \\
= \frac{1}{4} \left[\frac{\partial^2 f}{\partial \nabla^2 \rho_{\alpha} \partial \rho_{\alpha}} + \frac{\partial^2 f}{\partial \nabla^2 \rho_{\alpha} \partial \rho_{\beta}} + \frac{\partial^2 f}{\partial \nabla^2 \rho_{\beta} \partial \rho_{\alpha}} + \frac{\partial^2 f}{\partial \nabla^2 \rho_{\beta} \partial \rho_{\beta}} \right] \\
:= F[\rho].$$
(A.44)

Therefore the fourth term in Equation (A.40) is given by,

$$\left\langle N^2 \eta_j \sum_k c_k \eta_k F[\rho] \left[\sum_l c_l \nabla^2 \eta_i \eta_l + \sum_k c_k \nabla^2 \eta_k \eta_i + 2 \sum_k c_k \sum_{m=1}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} \right] \right\rangle.$$
(A.45)

The eighth term in Equation (A.40) further requires,

$$\frac{\partial^2 f}{\partial \nabla \rho \partial \nabla^2 \rho} = \frac{\partial}{\partial \nabla^2 \rho} \left[\frac{\nabla \rho}{2} \left(\frac{\partial f}{\partial G_{\alpha \alpha}} + \frac{\partial f}{\partial G_{\alpha \beta}} + \frac{\partial f}{\partial G_{\beta \beta}} \right) \right] \\
= \frac{\nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha \alpha} \partial \nabla^2 \rho_{\alpha}} + \frac{\partial^2 f}{\partial G_{\alpha \alpha} \partial \nabla^2 \rho_{\beta}} + \frac{\partial^2 f}{\partial G_{\alpha \beta} \partial \nabla^2 \rho_{\alpha}} \right] \\
+ \frac{\nabla \rho}{4} \left[\frac{\partial^2 f}{\partial G_{\alpha \beta} \partial \nabla^2 \rho_{\beta}} + \frac{\partial^2 f}{\partial G_{\beta \beta} \partial \nabla^2 \rho_{\alpha}} + \frac{\partial^2 f}{\partial G_{\beta \beta} \partial \nabla^2 \rho_{\beta}} \right] \\
:= \frac{G[\rho]}{4}.$$
(A.46)

Therefore the eighth term in Equation (A.40) is,

$$\left\langle N^2 \nabla \rho G[\rho] \nabla \left(\eta_j \sum_l c_l \eta_l \right) \left[\sum_l c_l \nabla^2 \eta_i \eta_l + \sum_k c_k \nabla^2 \eta_k \eta_i + 2 \sum_k c_k \sum_{\substack{m=1 \ k \neq m}}^3 \frac{\partial \eta_i}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} \right] \right\rangle.$$

The ninth term in Equation (A.40) further requires,

$$\frac{\partial^2 \nabla \rho}{\partial c_i \partial c_j} = 2N \left(\eta_i \nabla^2 \eta_j + \eta_j \nabla^2 \eta_i + 2\sum_{m=1}^3 \frac{\partial \eta_j}{\partial x_m} \frac{\partial \eta_i}{\partial x_m} \right), \tag{A.48}$$

and also

$$\frac{\boldsymbol{\nabla}^2 \rho}{2} = \boldsymbol{\nabla}^2 \rho_{\alpha} = \boldsymbol{\nabla}^2 \rho_{\beta}. \tag{A.49}$$

Therefore the ninth term in Equation (A.40) is given by,

$$\left\langle N\left(\frac{\partial f}{\partial \boldsymbol{\nabla}^2 \rho_{\alpha}} + \frac{\partial f}{\partial \boldsymbol{\nabla}^2 \rho_{\beta}}\right) \left(\eta_i \boldsymbol{\nabla}^2 \eta_j + \eta_j \boldsymbol{\nabla}^2 \eta_i + 2\sum_{m=1}^3 \frac{\partial \eta_j}{\partial x_m} \frac{\partial \eta_i}{\partial x_m}\right) \right\rangle.$$
(A.50)

The tenth term in Equation (A.40) requires nothing new, so one can immediately write down this term as

$$\left\langle N^2 \eta_i \sum_k c_k \eta_k F[\rho] \left[\sum_l c_l \nabla^2 \eta_j \eta_l + \sum_k c_k \nabla^2 \eta_k \eta_j + 2 \sum_k c_k \sum_{m=1}^3 \frac{\partial \eta_j}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} \right] \right\rangle.$$
(A.51)

The eleventh term in Equation (A.40) also requires nothing new and is given by,

$$\left\langle N^{2} \nabla \rho G[\rho] \nabla \left(\eta_{i} \sum_{l} c_{l} \eta_{l} \right) \left[\sum_{l} c_{l} \nabla^{2} \eta_{j} \eta_{l} + \sum_{k} c_{k} \nabla^{2} \eta_{k} \eta_{j} + 2 \sum_{k} c_{k} \sum_{\substack{m=1\\(A.52)}}^{3} \frac{\partial \eta_{j}}{\partial x_{m}} \frac{\partial \eta_{k}}{\partial x_{m}} \right] \right\rangle$$

The twelfth term of in Equation (A.40) further requires,

$$\frac{\partial^2 f}{\partial \nabla^2 \rho^2} = \frac{1}{2} \frac{\partial}{\partial \nabla^2 \rho} \left[\frac{\partial f}{\partial \nabla^2 \rho_\alpha} + \frac{\partial f}{\partial \nabla^2 \rho_\alpha} \right]
= \frac{1}{4} \left(\frac{\partial^2 f}{\partial \nabla^2 \rho_\alpha^2} + \frac{\partial^2 f}{\partial \nabla^2 \rho_\beta^2} + 2 \frac{\partial^2 f}{\partial \nabla^2 \rho_\alpha \partial \nabla^2 \rho_\beta} \right)$$

$$:= H[\rho].$$
(A.53)

Therefore, the twelfth term in Equation (A.40) is

$$\left\langle N^2 H[\rho] \left[\sum_l c_l \nabla^2 \eta_j \eta_l + \sum_k c_k \nabla^2 \eta_k \eta_j + 2 \sum_k c_k \sum_{m=1}^3 \frac{\partial \eta_j}{\partial x_m} \frac{\partial \eta_k}{\partial x_m} \right] [i \leftrightarrow j] \right\rangle.$$
(A.54)

The notation $[i \leftrightarrow j]$ stands for an expression which has the form to the expression on its immediate left but with the *i* and *j* labels interchanged.

B. Equations for the spin-polarised potential used in this work

B.1. Derivation

For spin-polarised DFT we can consider a functional of the following form

$$F[\rho] = \int f(\rho_{\alpha}, \rho_{\beta}, \boldsymbol{\nabla}\rho_{\alpha}, \boldsymbol{\nabla}\rho_{\beta}) d\mathbf{r}$$
(B.1)

The functional derivative with respect to the total density can decomposed into variations with respect to the α and β components of the density through,

$$\frac{\delta F}{\delta \rho} = \frac{\delta F}{\delta \rho_{\alpha}} + \frac{\delta F}{\delta \rho_{\beta}}.$$
 (B.2)

From the calculus of variations we know that

$$\frac{\delta F}{\delta \rho_{\alpha}} = \frac{\partial f}{\partial \rho_{\alpha}} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla} \rho_{\alpha}}.$$
(B.3)

Here we should note that XCFun expects a functional with the following form

$$F[\rho] = \int f(\rho_{\alpha}, \rho_{\beta}, G_{\alpha\alpha}, G_{\beta\beta}, G_{\alpha\beta}) d\mathbf{r}$$
(B.4)

where

$$G_{\sigma\sigma'} = \boldsymbol{\nabla} \rho_{\sigma} \cdot \boldsymbol{\nabla} \rho^{\sigma'}.$$

Therefore we can write Eq. (B.3) as

$$\frac{\delta F}{\delta \rho_{\alpha}} = \frac{\partial f}{\partial \rho_{\alpha}} - \boldsymbol{\nabla} \cdot \left[2 \boldsymbol{\nabla} \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha \alpha}} + \boldsymbol{\nabla} \rho_{\beta} \frac{\partial f}{\partial G_{\alpha \beta}} \right]. \tag{B.5}$$

Expanding the dot product and use of the chain rule leads to ,

$$\frac{\delta F}{\delta \rho_{\alpha}} = \frac{\partial f}{\partial \rho_{\alpha}}
- 2 \left[\nabla^{2} \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha\alpha}} + \nabla \rho_{\alpha} \cdot \left\{ \nabla \rho_{\alpha} \frac{\partial^{2} f}{\partial \rho_{\alpha} \partial G_{\alpha\alpha}} + \nabla \rho_{\beta} \frac{\partial^{2} f}{\partial \rho_{\beta} \partial G_{\alpha\alpha}} + \sum_{\sigma' \ge \sigma} \nabla G_{\sigma\sigma'} \frac{\partial^{2} f}{\partial G_{\sigma\sigma'} \partial G_{\alpha\alpha}} \right\} \right]
- \left[\nabla^{2} \rho_{\beta} \frac{\partial f}{\partial G_{\alpha\beta}} + \nabla \rho_{\beta} \cdot \left\{ \nabla \rho_{\alpha} \frac{\partial^{2} f}{\partial \rho_{\alpha} \partial G_{\beta\alpha}} + \nabla \rho_{\beta} \frac{\partial^{2} f}{\partial \rho_{\beta} \partial G_{\beta\alpha}} + \sum_{\sigma' \ge \sigma} \nabla G_{\sigma\sigma'} \frac{\partial^{2} f}{\partial G_{\sigma\sigma'} \partial G_{\beta\alpha}} \right\} \right].$$
(B.6)

So therefore,

$$\begin{split} \frac{\delta F}{\delta \rho} &= \frac{\partial f}{\partial \rho_{\alpha}} + \frac{\partial f}{\partial \rho_{\beta}} \\ &- 2 \bigg[\nabla^2 \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha \alpha}} + \nabla \rho_{\alpha} \cdot \bigg\{ \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\alpha \alpha}} + \nabla \rho_{\beta} \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\alpha \alpha}} + \sum_{\sigma' \geq \sigma} \nabla G_{\sigma \sigma'} \frac{\partial^2 f}{\partial G_{\sigma \sigma'} \partial G_{\alpha \alpha}} \bigg\} \bigg] \\ &- \bigg[\nabla^2 \rho_{\beta} \frac{\partial f}{\partial G_{\alpha \beta}} + \nabla \rho_{\beta} \cdot \bigg\{ \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta \alpha}} + \nabla \rho_{\beta} \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta \alpha}} + \sum_{\sigma' \geq \sigma} \nabla G_{\sigma \sigma'} \frac{\partial^2 f}{\partial G_{\sigma \sigma'} \partial G_{\beta \alpha}} \bigg\} \bigg] \\ &- 2 \bigg[\nabla^2 \rho_{\beta} \frac{\partial f}{\partial G_{\beta \beta}} + \nabla \rho_{\beta} \cdot \bigg\{ \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta \beta}} + \nabla \rho_{\beta} \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta \beta}} + \sum_{\sigma' \geq \sigma} \nabla G_{\sigma \sigma'} \frac{\partial^2 f}{\partial G_{\sigma \sigma'} \partial G_{\beta \beta}} \bigg\} \bigg] \\ &- \bigg[\nabla^2 \rho_{\alpha} \frac{\partial f}{\partial G_{\alpha \beta}} + \nabla \rho_{\alpha} \cdot \bigg\{ \nabla \rho_{\alpha} \frac{\partial^2 f}{\partial \rho_{\alpha} \partial G_{\beta \alpha}} + \nabla \rho_{\beta} \frac{\partial^2 f}{\partial \rho_{\beta} \partial G_{\beta \alpha}} + \sum_{\sigma' \geq \sigma} \nabla G_{\sigma \sigma'} \frac{\partial^2 f}{\partial G_{\sigma \sigma'} \partial G_{\beta \beta}} \bigg\} \bigg] . \end{split}$$
(B.7)

To provide the explicit expressions which were coded into QUEST the only remaining terms which require manipulation are of the following form

$$\boldsymbol{\nabla} \rho_{\alpha} \cdot \left\{ \sum_{\sigma' \ge \sigma} \boldsymbol{\nabla} G_{\sigma\sigma'} \frac{\partial^2 f}{\partial G_{\sigma\sigma'} \partial G_{\alpha\alpha}} \right\}.$$
(B.8)

Firstly,

$$G_{\sigma\sigma'} = \frac{\partial\rho^{\sigma}}{\partial x}\frac{\partial\rho^{\sigma'}}{\partial x} + \frac{\partial\rho^{\sigma}}{\partial y}\frac{\partial\rho^{\sigma'}}{\partial y} + \frac{\partial\rho^{\sigma}}{\partial z}\frac{\partial\rho^{\sigma'}}{\partial z}.$$
 (B.9)

Then the components of the spatial gradient are given by,

$$(\boldsymbol{\nabla}G_{\sigma\sigma'})_{i} = \frac{\partial^{2}\rho^{\sigma}}{\partial x\partial x_{i}}\frac{\partial\rho^{\sigma'}}{\partial x} + \frac{\partial\rho^{\sigma}}{\partial x}\frac{\partial^{2}\rho^{\sigma'}}{\partial x\partial x_{i}} + \frac{\partial^{2}\rho^{\sigma}}{\partial y\partial x_{i}}\frac{\partial\rho^{\sigma'}}{\partial y} + \frac{\partial\rho^{\sigma}}{\partial y}\frac{\partial^{2}\rho^{\sigma'}}{\partial y\partial x_{i}} + \frac{\partial^{2}\rho^{\sigma}}{\partial z\partial x_{i}}\frac{\partial\rho^{\sigma'}}{\partial z} + \frac{\partial\rho^{\sigma}}{\partial z}\frac{\partial^{2}\rho^{\sigma'}}{\partial z\partial x_{i}}$$
$$:= \rho_{xx_{i}}^{\sigma}\rho_{x}^{\sigma'} + \rho_{x}^{\sigma}\rho_{xx_{i}}^{\sigma'} + \rho_{yx_{i}}^{\sigma}\rho_{y}^{\sigma'} + \rho_{y}^{\sigma}\rho_{yx_{i}}^{\sigma'} + \rho_{zx_{i}}^{\sigma}\rho_{z}^{\sigma'} + \rho_{z}^{\sigma}\rho_{zx_{i}}^{\sigma'}.$$

$$(B.10)$$

Variational Solutions in OF-DFT M. S. Ryley

198

The dot product can therefore be expressed as,

$$\boldsymbol{\nabla}\rho^{\alpha} \cdot \boldsymbol{G}_{\sigma\sigma'} = \sum_{i=1}^{3} \rho_{x_i}^{\alpha} \Big(\rho_{xx_i}^{\sigma} \rho_x^{\sigma'} + \rho_x^{\sigma} \rho_{xx_i}^{\sigma'} + \rho_{yx_i}^{\sigma} \rho_y^{\sigma'} + \rho_y^{\sigma} \rho_{yx_i}^{\sigma'} + \rho_{zx_i}^{\sigma} \rho_z^{\sigma'} + \rho_z^{\sigma} \rho_{zx_i}^{\sigma'} \Big).$$
(B.11)

We have 3 unique pairs of spin in Eq. (B.8) to sum over; $(\alpha, \alpha), (\beta, \beta)$ and (α, β) . The (α, α) part is,

$$\boldsymbol{\nabla}\rho^{\alpha} \cdot \boldsymbol{G}_{\alpha\alpha} = 2\sum_{i=1}^{3} \rho_{x_i}^{\alpha} \left(\rho_{xx_i}^{\alpha} \rho_x^{\alpha} + \rho_{yx_i}^{\alpha} \rho_y^{\alpha} + \rho_{zx_i}^{\alpha} \rho_z^{\alpha} \right).$$
(B.12)

Similarly for the (β, β) term is,

$$\boldsymbol{\nabla}\rho^{\alpha} \cdot G_{\beta\beta} = 2\sum_{i=1}^{3} \rho_{x_i}^{\alpha} \left(\rho_{xx_i}^{\beta} \rho_x^{\beta} + \rho_{yx_i}^{\beta} \rho_y^{\beta} + \rho_{zx_i}^{\beta} \rho_z^{\beta} \right).$$
(B.13)

The (α, β) term does not have any symmetry and so just evaluates to

$$\boldsymbol{\nabla}\rho^{\alpha} \cdot \boldsymbol{G}_{\alpha\beta} = \sum_{i=1}^{3} \rho_{x_{i}}^{\alpha} \left(\rho_{xx_{i}}^{\alpha} \rho_{x}^{\beta} + \rho_{x}^{\alpha} \rho_{xx_{i}}^{\beta} + \rho_{yx_{i}}^{\alpha} \rho_{y}^{\beta} + \rho_{y}^{\alpha} \rho_{yx_{i}}^{\beta} + \rho_{zx_{i}}^{\alpha} \rho_{z}^{\beta} + \rho_{z}^{\alpha} \rho_{zx_{i}}^{\beta} \right).$$
(B.14)

Therefore Eq. (B.8) is equal to

$$2\sum_{i=1}^{3} \rho_{x_{i}}^{\alpha} \left(\rho_{xx_{i}}^{\alpha}\rho_{x}^{\alpha} + \rho_{yx_{i}}^{\alpha}\rho_{y}^{\alpha} + \rho_{zx_{i}}^{\alpha}\rho_{z}^{\alpha}\right) \frac{\partial^{2} f}{\partial G_{\alpha\alpha}^{2}} + 2\sum_{i=1}^{3} \rho_{x_{i}}^{\alpha} \left(\rho_{xx_{i}}^{\beta}\rho_{x}^{\beta} + \rho_{yx_{i}}^{\beta}\rho_{y}^{\beta} + \rho_{zx_{i}}^{\beta}\rho_{z}^{\beta}\right) \frac{\partial^{2} f}{\partial G_{\alpha\alpha} \partial G_{\beta\beta}} + \sum_{i=1}^{3} \rho_{x_{i}}^{\alpha} \left(\rho_{xx_{i}}^{\alpha}\rho_{x}^{\beta} + \rho_{x}^{\alpha}\rho_{xx_{i}}^{\beta} + \rho_{yx_{i}}^{\alpha}\rho_{y}^{\beta} + \rho_{yx_{i}}^{\alpha}\rho_{y}^{\beta} + \rho_{yx_{i}}^{\alpha}\rho_{y}^{\beta} + \rho_{zx_{i}}^{\alpha}\rho_{zx_{i}}^{\beta} + \rho_{z}^{\alpha}\rho_{zx_{i}}^{\beta}\right) \frac{\partial^{2} f}{\partial G_{\alpha\alpha} \partial G_{\alpha\beta}}$$
(B.15)