Theoretical prediction of properties of atomistic systems

Density functional theory and machine learning

Alexander Lindmaa

Linköping University
Department of Physics, Chemistry and Biology
Theory and modeling
SE-581 83 Linköping, Sweden

Linköping 2017

© Alexander Lindmaa
ISBN 978-91-7685-486-0
ISSN 0345-7524
Typeset using L^AT_EX

Printed by LiU-Tryck, Linköping 2017

Till min käre far

Abstract

The prediction of ground state properties of atomistic systems is of vital importance in technological advances as well as in the physical sciences. Fundamentally, these predictions are based on a quantum-mechanical description of many-electron systems. One of the hitherto most prominent theories for the treatment of such systems is density functional theory (DFT). The main reason for its success is due to its balance of acceptable accuracy with computational efficiency. By now, DFT is applied routinely to compute the properties of atomic, molecular, and solid state systems.

The general approach to solve the DFT equations is to use a density-functional approximation (DFA). In Kohn-Sham (KS) DFT, DFAs are applied to the unknown exchange-correlation (xc) energy. In orbital-free DFT on the other hand, where the total energy is minimized directly with respect to the electron density, a DFA applied to the non-interacting kinetic energy is also required. Unfortunately, central DFAs in DFT fail to qualitatively capture many important aspects of electronic systems. Two prime examples are the description of localized electrons, and the description of systems where electronic edges are present.

In this thesis, I use a model system approach to construct a DFA for the electron localization function (ELF). The very same approach is also taken to study the non-interacting kinetic energy density (KED) in the slowly varying limit of inhomogeneous electron densities, where the effect of electronic edges are effectively included. Apart from the work on model systems, extensions of an exchange energy functional with an improved KS orbital description are presented: a scheme for improving its description of energetics of solids, and a comparison of its description of an essential exact exchange feature known as the derivative discontinuity with numerical data for exact exchange.

An emerging alternative route towards the prediction of the properties of atomistic systems is machine learning (ML). I present a number of ML methods for the prediction of solid formation energies, with an accuracy that is on par with KS DFT calculations, and with orders-of-magnitude lower computational cost.

Svensk sammanfattning

Att kunna förutsäga egenskaper hos atomistiska system utgör en viktig del av vår teknologiska utveckling, samt spelar en betydande roll i de fysikaliska vetenskaperna. Sådana förutsägelser bygger på en kvantmekanisk beskrivning av mångelektronssystem. En av de mest framstående teorierna för att behandla den här typen av system är täthetsfunktionalteorin (DFT). Den främsta orsaken till dess framgång är att den lyckas kombinera skaplig noggrannhet med en bra beräkningseffektivitet. DFT används numera rutinmässigt för att beräkna storheter hos atomer, molekyler, och fasta kroppar.

Generellt sett löses ekvationerna inom DFT genom att man inför en täthetsfunktionalapproximation (DFA). I Kohn-Sham (KS) DFT, används DFAer för att approximera utbytes-korrelationsenergin. Inom orbitalfri DFT, där målet är att direkt minimera den totala energin med avseende på elektrontätheten, så approximerar man också den icke-interagerande rörelseenergin hos elektronerna. Dessvärre så fallerar många centrala DFAer att kvalitativt beskriva många viktiga aspekter hos elektronssystem. Två viktiga exempel är beskrivningen av lokaliserade elektroner, samt beskrivningen av system där det förekommer elektronytor.

I denna avhandling använder jag modellsystem för att konstruera en DFA för elektronlokaliseringsfunktionen (ELF). Samma tillvägagångssätt appliceras sedan för att studera den kinetiska energitätheten i gränsen av långsamt varierande elektrontätheter, där effekten av elektronytor effektivt inkluderas. Förutom arbetet som berör modellsystem, så presenteras en utökad variant av en utbytes-energifunktional med en förbättrad KS orbitalbeskrivning; ett schema för att förbättra dess energiegenskaper för solida material, samt en jämförelse av dess beskrivning av en viktig egenskap hos den exakta utbytesenergin, vilket utgörs av diskontinuiteter i dess derivata.

Ett mera nyligen uppkommet samt alternativt sätt att kunna förutsäga egenskaper hos atomistiska system utgörs av maskinlärning (ML). Jag presenterar ett antal ML-modeller för att kunna förutsäga formeringsenergier hos fasta material med en noggrannhet som är i linje med resultat som uppnås av beräkningar med hjälp av KS DFT, och med en beräkningseffektivitet som är flera storleksordningar snabbare.

Preface

This thesis is a summary of some of the work I was part of between 2012 and 2017 in the theoretical physics group at Linköping University. I am indebted to a large number of people, without whom this work would not have seen the light of day.

First of all, I would like to thank my supervisor, Rickard Armiento, who successfully, and with great enthusiasm and perseverance, has guided me during these past years. Working with you has truly been a pleasure, and a privilege. I also wish to thank my co-supervisor, Igor Abrikosov, for his encouragement and support, and for making all of this happen in the first place.

Ann Mattsson is greatly acknowledged for nice and insightful conversations. I am grateful for our current and future collaboration. In addition, I want to thank John Wills for his encouraging and nice ways. I have also been able to enjoy the warm hospitality of Stephan Kümmel at the University of Bayreuth and Olle Eriksson at Uppsala University, when visiting their groups.

I want to acknowledge everyone in the theoretical physics group, from the weathered members of the old breed to the young and vigilant brothers in arms. Also, the support that I have received from my family and friends has been invaluable. Finally, I want to thank B, for her enduring support during the writing of this thesis, and for reminding me to continue to *get after it*.

Alex Lindmaa
Linköping, August 2017

CONTENTS

Acronyms xiii

1	Introduction	3	
1.1	Outline of thesis		4
2	Density functional theory	5	
2.1	The Hohenberg-Kohn variational principle		5
2.2	The Kohn-Sham non-interacting system		8
2.3	Extension to open systems		11
2.3.1	Derivative discontinuity		11
2.4	Some exact conditions		13
2.4.1	General relations		13
2.4.2	Scaling relations		13
2.4.3	Exchange-correlation hole		14
2.5	Thomas-Fermi theory and orbital-free DFT		16
2.6	Functional approximations		17
2.6.1	The local density approximation		17
2.6.2	Generalized gradient approximations		17
2.6.3	The Perdew-Burke-Ernzerhof functional		18
2.6.4	Extensions to Thomas-Fermi theory		18
2.7	Optimized effective potential		19
2.8	The electron localization function		21
3	Model system approach to kinetic energy functionals	23	
3.1	Edge electron gas		23
3.1.1	The Airy gas model system		24
3.2	Quantum oscillations in the kinetic energy density		26
3.2.1	The slowly varying limit of the Airy gas		26
3.2.2	Bohr atoms and the isotropic harmonic oscillator		28
3.2.3	The Mathieu gas		30
3.2.4	Characterization of model systems in terms of eigenvalues		33
3.2.5	Development of a kinetic energy functional		33
3.3	Parameterization of the ELF in the harmonic oscillator model system		36
4	Properties of the AK13 exchange functional	39	
4.1	The Becke-Johnson model potential		39
4.2	The Armiento-Kümmel exchange functional		41
4.3	Dealing with inaccurate energetics		43
4.3.1	Self-consistent scheme for volume relaxations of solids		43
4.4	Modeling of the exact-exchange derivative discontinuity		45
5	Machine learning	47	
5.1	Kernel ridge regression		47
5.1.1	Generalization error		48
5.2	Feature vector representations		49
5.2.1	Principal component analysis		49

5.3	Coordinate-based representations of extended systems	50
5.3.1	The Coulomb matrix representation for molecules	50
5.3.2	The Ewald-sum vector representation	50
5.3.3	General results on datasets	51
5.4	Symmetry-restricted vector representations	52
6	Conclusions	53
7	Appendix	55
7.1	Positive definite KED of closed, radially symmetric systems	55
7.1.1	The Bohr atom	55
7.2	Properties of the Mathieu gas	57
7.3	Positive definite KED of the Hermite gas	58
	Bibliography	61
	List of figures	66

Papers 67

List of publications and my contribution 68

Paper I 69

Quantum oscillations in the kinetic energy density: Gradient corrections from the Airy gas

Paper II 87

Parameterization of the Electron Localization Function in Semi-Local Density Quantities

Paper III 95

Energetics of the AK13 semilocal Kohn-Sham exchange energy functional

Paper IV 105

Modeling of the discontinuous shift of the exchange potential in semi-local density functional theory

Paper V 115

Crystal structure representations for machine learning models of formation energies

Paper VI 125

Machine learning energies of 2 million Elpasolite (ABC_2D_6) crystals

ACRONYMS

AG	Airy gas
AK13	Armiento-Kümmel functional
BJ	Becke-Johnson
DD	derivative discontinuity
DFT	density functional theory
EG	edge gas
ELF	electron localization function
ETF	extended Thomas-Fermi
EXX	exact exchange
GE	gradient expansion
GGA	generalized gradient approximation
HK	Hohenberg-Kohn
KED	kinetic energy density
KRR	kernel ridge regression
KS	Kohn-Sham
LDA	local density approximation
MG	Mathieu gas
ML	machine learning
OEP	optimized effective potential
PBE	Perdew-Burke-Ernzerhof
PCA	principal component analysis
TF	Thomas-Fermi
xc	exchange-correlation

There are known knowns; these are
things we know that we know.
There are also known unknowns;
that is to say, there are things that
we know we don't know. But there
are also unknown unknowns.
These are the things we don't know
we don't know.

— Donald Rumsfeld

INTRODUCTION

The past few decades have seen an enormous progress in the field of theoretical materials science. The reason for this is at least three-fold. First, the steady improvement in computer hardware has made it possible to simulate large and complex electronic systems, both statically and dynamically, with the aid of modern computers. Second, the development of new and improved theoretical methods has been of essential importance. The fundamental problem that underlies any such method is the description of systems of many interacting electrons and nuclei, according to the theory of quantum mechanics. Finally, many achievements have been made in software development, which has resulted in a vast number of various kinds of efficient and highly accurate numerical solvers.

Since the birth of quantum theory, up to the modern day era, many sophisticated approaches to the many electron problem have been suggested. Traditional wave function based methods that rely on different kinds of variational principles have been successfully applied to finite systems, such as atoms and molecules. However, for extended systems these methods suffer from the well known exponential-wall problem in the number of particles described.²⁹ Another category of methods rests on perturbation theory, such as the Green's function technique.¹⁸ While this technique can be applied to infinite systems, such as quantum liquids, it typically requires vast computational resources, which renders it ineffective when it comes to the computation of properties of inhomogeneous electron systems, such as solids.

Initial steps for solving these problems was initiated roughly 50 years ago, when a formally exact way of dealing with the interacting many-body was proposed: the density functional method.^{10,47,11} Founded by Hohenberg and Kohn²⁴ and Kohn and Sham³², density functional theory (DFT) generally requires less computational effort, and can be applied to both finite and infinite systems. The gist of DFT is that the interacting many-electron problem can be mapped to an auxiliary, non-interacting one, via an effective potential, the most commonly used one being the one suggested by Kohn and Sham themselves in 1965. The accuracy of a typical Kohn-Sham (KS) density-functional calculation then depends on approximating the unknown exchange-correlation (xc) energy term, $E_{xc}[n]$. By now, numerous successful approximations exist and can be used in virtually any context where the total energy of the system is sought as function of the positions of the nuclei.²⁵

While the most commonly used approximations in KS DFT are generally accurate when it comes to the prediction of total-energy related properties, they exhibit some serious systematic errors in a number of important situations. One example is the lack of a derivative discontinuity (DD) at integer electron numbers⁴⁹, which may result in e.g. the underestimation of fundamental gaps when relying on the KS eigenvalues^{55,63,62,8}, a qualitative failure of the dissociation limit of a stretched molecule⁴⁹, and as a possible result, the failure of an accurate description of charge transfer in molecules and solids.⁶¹ In this thesis, I discuss some new directions where the aim is to include the effect of a DD in a semi-local level of theory, in combination with a more accurate description the energetic properties of solids.

A related issue is the problem of determining the degree of electron localization present in electronic systems. At a semi-local level, this can be provided by the electron localization function (ELF)^{4,53}, which depends on the kinetic energy of the KS single-particle orbitals. In this thesis, I use a model system approach to derive a parameterization of the ELF entirely in terms of the density and its second-order derivatives, and apply it to a number of physically relevant cases.

One possibility of speeding up density-functional calculations even further is to directly minimize the total energy with respect to the electron density, i.e., by leaving out the KS mapping to the auxiliary, non-interacting system in its entirety. This procedure is commonly known as orbital-free, or pure, DFT. However, existing orbital-free approximations are currently far from as accurate as KS approximations. An especially important feature of electronic systems is the presence of electronic surfaces.³³ In this thesis, I present theoretical and numerical results related to the positive definite kinetic energy density (KED) in the slowly varying limit of the inhomogeneous gas, where the effect of electronic surfaces is effectively included.

Another way to handle the interacting many-body problem is to circumvent it altogether. One way of doing this is to introduce various kinds of model Hamiltonians, such as the Hubbard and Heisenberg Hamiltonians.¹⁵ Such constructs are limited in the sense that they can only handle the physics they have been designed for. The pursuit of an ever increased computational efficiency has, however, very recently inspired the use of methods based on statistical regression models. One example of this is machine learning.²² ML consists of various algorithms which via induction are able to recognize patterns in high-dimensional spaces. In this thesis, I develop ML methods capable of computing solid formation energies with an accuracy which is on par with DFT, but with an orders-of-magnitude increase in computational efficiency.

1.1 *Outline of thesis*

In chapter 2, I give a brief summary of ground state density-functional theory. In chapter 3, I use a model system approach to study slowly varying properties of the positive definite kinetic energy density for systems with electronic edges. I also develop a theoretical description of the electron localization function in terms semi-local density-based quantities. Chapter 4 is devoted to the incorporation of the features of exact-exchange into semi-local density functional theory, with emphasis on the derivative discontinuity. In chapter 5 I develop machine learning models aimed at the accurate prediction of solid formation energies. Finally, I give my concluding remarks in chapter 6.

DENSITY FUNCTIONAL THEORY

In this chapter I discuss the main aspects of ground state DFT that are relevant to this work. I start with the original variational principle of Hohenberg and Kohn, and go on to describe the single-particle formulation of Kohn and Sham. I also discuss the generalization of the theory to open systems, the theory of exact exchange within the optimized effective potential method, as well as pure DFT and some of its connections to Thomas-Fermi (TF) theory.

The aim of DFT is to solve the Schrödinger equation solely in terms of the electron density. While this can be done exactly in principle, we are still left with the problem of having to approximate the exchange-correlation energy in some way or another.

2.1 The Hohenberg-Kohn variational principle

Consider a non-relativistic, interacting system of N electrons and M nuclei. I shall assume the validity of the Born-Oppenheimer approximation⁶, so that the quantum mechanics of the electrons can be treated separately. The Hamiltonian of the electronic degrees of freedom is (throughout this thesis, I use Hartree atomic units)

$$\begin{aligned}\hat{H} &= -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N v(\mathbf{r}_i) + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &\equiv \hat{T} + \hat{V} + \hat{V}_{ee}.\end{aligned}\quad (2.1)$$

Here, $\hat{T}$ is the kinetic energy of the electrons, $\hat{V}$ electron-nuclear interaction, and $\hat{V}_{ee}$ is the Coulomb interaction between the electrons. The external potential $v(\mathbf{r})$ is given by

$$v(\mathbf{r}) = - \sum_{I=1}^M \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}, \quad (2.2)$$

i.e., the Coulomb potential from M classical nuclei, with atomic numbers $\{Z_I\}$, situated at positions $\{\mathbf{R}_I\}$. I will not be concerned with the presence of any external electric or magnetic fields.

The Rayleigh-Ritz principle states that for any antisymmetric wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \in L^2(\mathbb{R}^N)$, the ground-state energy E_0 of the system satisfies the relation

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0, \quad (2.3)$$

where equality holds for the ground state Ψ_0 . Given such a Ψ_0 , equation 2.3 yields

$$\begin{aligned}\frac{\delta E}{\delta \Psi} &= \frac{1}{\langle \Psi | \Psi \rangle} \left[\frac{\delta \langle \Psi | \hat{H} | \Psi \rangle}{\delta \Psi} - E[\Psi] \frac{\delta \langle \Psi | \Psi \rangle}{\delta \Psi} \right] \\ &= \frac{1}{\langle \Psi | \Psi \rangle} [\hat{H} \Psi^* - E \Psi^*] = 0,\end{aligned}\quad (2.4)$$

For systems of interest, where N may very well be of the order of Avogadro's number, equation 2.4 becomes enormously complex. In order to make progress, we must resort to alternative formulations of some sort. First, to reduce the complexity associated with having to deal directly with the wave function Ψ , we instead focus on a related quantity, namely the electron density

$$\begin{aligned} n(\mathbf{r}) &= \langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle = \left\langle \Psi \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right| \Psi \right\rangle \\ &= N \int d^3 r_2 \dots \int d^3 r_N |\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2, \end{aligned} \quad (2.5)$$

where $\hat{n}$ is the electron number operator. The density $n(\mathbf{r})$ of equation 2.5 measures the probability of finding any electron situated at $\mathbf{r}$. It is evident that n can be handled much more easily than Ψ , as it contains only three dimensions, regardless of the value of N .

The next step is to formulate equation 2.4 as a variational principle in terms of n . This we can achieve by invoking the Hohenberg-Kohn (HK) theorem²⁴, which can be summarized by the following two statements (stated here without proof):

1. The map $\mathcal{V} : v \mapsto \Psi_0[v]$, as defined by the relation

$$(\hat{H} - E)\Psi = 0,$$

where each Ψ_0 is a non-degenerate ground state, is invertible.

2. The map $\mathcal{D} : \Psi_0 \mapsto n_0[\Psi]$, as defined by equation 2.5, and where n_0 is the ground state density, is also invertible.

Hence, the joint map $\mathcal{VD}$ is invertible, and v is uniquely determined by n_0 (up to an additive constant), i.e.,

$$n_0 \rightarrow v + \mathcal{C}, \quad \mathcal{C} \in \mathbb{R}. \quad (2.6)$$

In turn, all observable ground-state properties of a system characterized by the Hamiltonian $\hat{H}$ of equation 2.1 are unique functionals of n_0 . HK also discussed the existence of a variational principle in the density for determining the ground state energy. Here, we proceed with this discussion along the direction of Levy and Lieb.^{36,39} The functional for the total energy is defined as

$$E[n, v] = \underbrace{\min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{V}_{\text{ee}} | \Psi \rangle}_{F[n]} + \underbrace{\int d^3 r n(\mathbf{r}) v(\mathbf{r})}_{V[n]}, \quad (2.7)$$

where $F[n]$ is the universal Levy-Lieb functional, which is defined by performing a constrained search over all anti-symmetric wave functions Ψ yielding the density n , and where we have used the fact that

$$\langle V \rangle = \left\langle \Psi \left| \sum_{i=1}^N v(\mathbf{r}_i) \right| \Psi \right\rangle = \int d^3 r n(\mathbf{r}) v(\mathbf{r}), \quad \forall \Psi \rightarrow n. \quad (2.8)$$

For a given external potential v , the ground-state energy and density becomes accessible by performing the two-step minimization procedure

$$E_0 = \min_n E[n, v] = \min_n \{F[n] + V[n]\}. \quad (2.9)$$

The minimizing density is the one for which the Lagrangian

$$\mathcal{L}[n, v] = F[n] + V[n] - \mu \left(\int d^3r n(\mathbf{r}) - N \right) \quad (2.10)$$

becomes stationary, where we have introduced the Lagrange multiplier μ to ensure particle-number conservation throughout the search. This condition leads to the Euler-Lagrange equation for the interacting system:

$$\frac{\delta F}{\delta n} + v(\mathbf{r}) = \mu. \quad (2.11)$$

All of the quantum many-body effects inherent to the Schrödinger equation have been included in the universal functional $F[n]$, in accordance with equation 2.7. Unfortunately, the exact functional form of $F[n]$ is unknown, and must hence in general be approximated. A common approach that allows the largest contributions to $F[n]$ to be extracted exactly is the scheme by Kohn and Sham.³²

2.2 The Kohn-Sham non-interacting system

The HK theorem discussed in the previous section is valid for any reasonable electron-electron interaction $\hat{V}_{ee}$. Of particular interest to the present discussion are systems of non-interacting electrons, i.e., systems where $\hat{V}_{ee} \equiv 0$. The resulting single-particle Hamiltonian is thus

$$\hat{H}_s = \hat{T} + \hat{V}_s \equiv -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N v(\mathbf{r}_i). \quad (2.12)$$

Applying the HK theorem results in the existence of a unique energy functional

$$E[n_s] = T_s[n_s] + \int d^3r n_s(\mathbf{r}) v_s(\mathbf{r}), \quad (2.13)$$

with the associated Euler-Lagrange equation

$$\frac{\delta T_s}{\delta n} + v_s = \mu_s, \quad (2.14)$$

and where the universal functional is given by

$$T_s[n_s] = \min_{\Phi_s \rightarrow n_s} \langle \Phi_s | \hat{T} | \Phi_s \rangle. \quad (2.15)$$

The search is now performed over all anti-symmetric non-interacting wave functions Φ_s . Once again, we assume these to be non-degenerate, which means that they will take on the form of single Slater determinants

$$\Phi_s(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \det\{\phi_i(\mathbf{r}_k)\}, \quad \langle \phi_i | \phi_j \rangle = \delta_{ij}, \quad \forall i, j, \quad (2.16)$$

where the single-particle orbitals ϕ_i satisfy

$$\left(-\frac{\nabla^2}{2} + v_s \right) \phi_i = \varepsilon_i \phi_i, \quad \varepsilon_1 \leq \varepsilon_2 \leq \dots, \quad (2.17)$$

and each ε_i is a KS eigenvalue. The density and kinetic energy are now readily obtained as

$$n_s = \sum_{i=1}^N \phi_i^* \phi_i, \quad (2.18)$$

and

$$T_s[n_s] = -\frac{1}{2} \sum_{i=1}^N \langle \phi_i | \nabla^2 | \phi_i \rangle = \frac{1}{2} \int d^3r \sum_{i=1}^N \nabla \phi_i^* \nabla \phi_i. \quad (2.19)$$

The fundamental assumption made in the original work of KS is that for each interacting ground state density n_0 , there is a one-body potential v_s , called the KS potential, with a corresponding non-interacting ground-state density $n_{s,0} = n_0$. With this assumption in mind, we can decompose $F[n]$ of equation 2.7 in the following way, and for any n :

$$F[n] = T_s[n] + E_H[n] + E_{xc}[n], \quad (2.20)$$

where $T_s[n]$ is defined by equation 2.15, $E_H[n]$ is the classical electrostatic self-energy (Hartree energy) of the electrons, i.e.,

$$E_H[n] = \frac{1}{2} \int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (2.21)$$

and $E_{xc}[n]$ is the xc energy, as defined by equation 2.20.

In order to determine the one-body potential $v_s[n_0](\mathbf{r})$, we let n_0 be the stationary state. Then, for a given external potential v , this requirement leads to

$$\begin{aligned} \delta E[n_0, v] &= E[n_0 + \delta n, v] - E[n_0, v] \\ &= \delta T_s + \int d^3r \delta n(\mathbf{r}) [v(\mathbf{r}) + v_H[n_0](\mathbf{r}) + v_{xc}[n_0](\mathbf{r})] \\ &= 0, \end{aligned} \quad (2.22)$$

for interacting v -representable, non-particle conserving density variations $n_0 + \delta n$. Here, $v_{xc}[n_0](\mathbf{r})$ is the xc potential, which is defined as

$$v_{xc}[n_0](\mathbf{r}) = \left. \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \right|_{n=n_0}, \quad (2.23)$$

and

$$v_H[n_0](\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \Big|_{n=n_0} \quad (2.24)$$

is the Hartree potential. In order to handle the variation δT_s appearing in 2.22, we first note that variations of the form $n_0 + \delta n$ are also non-interacting v -representable w.r.t the potential v_s . This means that n_0 is uniquely determined by the orbitals $\phi_{i,0} + \delta\phi_i$. Thus, we get

$$\begin{aligned} \delta T_s[n] &= \delta \sum_{i=1}^N \left\langle \phi_i \left| -\frac{\nabla^2}{2} \right| \phi_i \right\rangle \\ &= \sum_{i=1}^N \int d^3r \left[\delta\phi_i^* \left(-\frac{\nabla^2}{2} \right) \phi_{i,0} + \delta\phi_i \left(-\frac{\nabla^2}{2} \right) \phi_{i,0}^* \right] \\ &= \left\langle -\frac{\nabla^2}{2} \phi_i = (\varepsilon_i - v_s) \phi_i, \quad \forall i \right\rangle = \\ &= \sum_{i=1}^N \int d^3r \left[\delta\phi_i^* (\varepsilon_i - v_s) \phi_{i,0} + \delta\phi_i (\varepsilon_i - v_s) \phi_{i,0}^* \right] \\ &= \sum_{i=1}^N \varepsilon_i \int d^3r \delta|\phi_i|^2 - \sum_{i=1}^N \int d^3r v_s \delta|\phi_i|^2. \end{aligned}$$

Using the fact that, to first order

$$\begin{aligned} \delta|\phi_i|^2 &= |\phi_{i,0} + \delta\phi_i|^2 - |\phi_{i,0}|^2 \\ &= \phi_{i,0}^* \delta\phi_i + \phi_{i,0} \delta\phi_i^*, \end{aligned}$$

and the variation 2.22, we obtain

$$\delta T_s[n] = - \int d^3r v_s(\mathbf{r}) n(\mathbf{r}). \quad (2.25)$$

Thus, by using equation 2.22, the KS potential is given by

$$v_s(\mathbf{r}) = v(\mathbf{r}) + v_H[n_0](\mathbf{r}) + v_{xc}[n_0](\mathbf{r}). \quad (2.26)$$

Solving equations 2.17, 2.18, and 2.26 self-consistently constitutes the method of KS.³² For particle-number conserving variations, the Euler-Lagrange equation for the interacting system, equation 2.11, together with the definition 2.20, at stationary density becomes

$$\begin{aligned} 0 &= \frac{\delta}{\delta n} \left[E[n, v] - \mu \int n(\mathbf{r}) d^3r \right] \\ &= \frac{\delta T_s}{\delta n} + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' + \frac{\delta E_{xc}}{\delta n} - \mu. \end{aligned} \quad (2.27)$$

The fact that equation 2.14, i.e., Euler-Lagrange equation for the non-interacting system is solved for the same density, again results in the equation 2.26, up to a constant.

The advantage of the KS method is that instead of having to approximate the universal functional $F[n]$, one only needs to approximate a much smaller quantity, namely the xc energy of equation 2.20.

2.3 Extension to open systems

In sections 2.1, and 2.2, the discussion involved ordinary single, pure systems. However, in order to handle general situations where the system is allowed to exchange particles with another system, it is necessary to extend the formalism somewhat. The starting point for the treatment of open systems within a density-functional framework is a zero-temperature ensemble Γ of pure states $|\Psi_M\rangle$:^{39,65}

$$\hat{\Gamma} = \alpha_M |\Psi_M\rangle \langle \Psi_M| + \alpha_{M+1} |\Psi_{M+1}\rangle \langle \Psi_{M+1}|, \quad \alpha_M + \alpha_{M+1} = 1. \quad (2.28)$$

In general, for an arbitrary operator $\hat{O}$, its expectation value O is given by

$$O = \text{Tr} \{ \hat{\Gamma} \hat{O} \} = \sum_k \langle \Psi_k | \hat{\Gamma} \hat{O} | \Psi_k \rangle, \quad (2.29)$$

where Tr is the trace. In particular, the density is given by

$$n(\mathbf{r}) = \text{Tr} \{ \hat{\Gamma} \hat{n} \}, \quad (2.30)$$

so that

$$\int n(\mathbf{r}) d^3r = M + \eta, \quad M \in \mathbb{N}_0, \quad \eta \in [0, 1]. \quad (2.31)$$

The total energy functional for fractional particle numbers is^{39,10}

$$E_f[n, v] = F_f[n] + V[n], \quad (2.32)$$

where

$$F_f[n] = \min_{\hat{\Gamma} \rightarrow n} \text{Tr} \{ \hat{\Gamma} (\hat{T} + \hat{V}_{\text{ee}}) \}. \quad (2.33)$$

An important consequence that follows from the extended formalism to open systems is that it can be used to define the fundamental gap in KS DFT. Its definition is given by⁴⁹

$$I(M) - A(M) = \Delta_{\text{KS}} + \Delta_{\text{xc}}, \quad (2.34)$$

where $\Delta_{\text{KS}} = \varepsilon_{M+1}(M) - \varepsilon_M(M)$ is the single particle KS gap and Δ_{xc} is the derivative discontinuity in the xc potential of equation 2.23. It is computed via

$$\Delta_{\text{xc}}[n] = \lim_{\eta \rightarrow 0^+} \left\{ \left. \frac{\delta E_{\text{xc}}}{\delta n} \right|_{M+\eta} - \left. \frac{\delta E_{\text{xc}}}{\delta n} \right|_{M-\eta} \right\}, \quad (2.35)$$

where the functional derivatives are to be evaluated at the respective $n_{M \pm \eta}$ densities.

2.3.1 DERIVATIVE DISCONTINUITY

In figure 2.1, an important consequence of the extension of the DFT formalism to systems with fractional particle numbers is illustrated, namely the general behavior of the total energy E as a function of the time-averaged particle number N .⁴⁹ When the number of particles increases through an integer value M ,

there is an abrupt change in the slope of the curve, which corresponds to a new KS orbital getting occupied. At these points there is hence a DD in the total energy: This manifests itself as a discontinuous jump in the KS potential, as the fraction η tends to zero. The full contribution to this jump is given by equation 2.34, i.e., there is also a contribution to the jump that arises from the non-interacting kinetic energy.

FIGURE 2.1: Total energy of system as a function of the time-averaged particle number N . When $N = M$ is an integer value, the slope of the curve changes abruptly. For these values, the KS potential may jump, with a magnitude equal to the derivative discontinuity Δ_{xc} , which is given by equation 2.35.

2.4 Some exact conditions

While the universal functional $F[n]$ of equation 2.20 is unknown, it is still possible to derive various sets of exact conditions for its constituents, the non-interacting kinetic energy $T_s[n]$, and the xc-correlation energy $E_{xc}[n]$. Exact conditions are important tools for the construction of functional approximations. There are many exact conditions, and I shall only be considering the ones that are most relevant to this work.

2.4.1 GENERAL RELATIONS

In section 2.2, the xc energy was defined via equation 2.20, as

$$E_{xc}[n] = F[n] - T_s[n] - E_H[n]. \quad (2.36)$$

It is common to further decompose this term according to

$$E_{xc}[n] = E_x[n] + E_c[n], \quad (2.37)$$

where

$$E_x[n] = \langle \Phi_s | \hat{V}_{ee} | \Phi_s \rangle - E_H[n], \quad (2.38)$$

and Φ_s is the single Slater determinant of equation 2.16. It follows from equation 2.36 and equation 2.38 that the correlation energy is

$$E_c[n] = F[n] - \{T_s[n] + E_H[n] + E_x[n]\}. \quad (2.39)$$

2.4.2 SCALING RELATIONS

A useful set of exact conditions follow from uniform coordinate scaling.³⁷ For a given ground state $\Psi \in L^2(\mathbb{R}^N)$, we consider the scaling of all its coordinates by some positive number γ , which is such that the normalization to N particles is conserved, i.e.,

$$\Psi_\gamma(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \gamma^{\frac{3N}{2}} \Psi(\gamma\mathbf{r}_1, \gamma\mathbf{r}_2, \dots, \gamma\mathbf{r}_N), \quad \gamma > 0. \quad (2.40)$$

The corresponding scaled density then becomes

$$n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r}). \quad (2.41)$$

For the non-interacting kinetic energy and the exchange energy this results in³⁷

$$\begin{aligned} T_s[n_\gamma] &= \gamma^2 T_s[n] \\ E_x[n_\gamma] &= \gamma E_x[n]. \end{aligned} \quad (2.42)$$

For $E_c[n]$, the correlation energy, only scaling inequalities can be derived. The reason for this is that the scaled wave function Ψ_γ is not the ground state of the scaled density n_γ .¹⁴

In order to transform density functionals into their spin-polarized equivalents, we use the spin-scaling relations. For the spin-densities $n^\uparrow$ and $n^\downarrow$, the following relation holds:⁴⁵

$$T_s[n^\uparrow, n^\downarrow] = \frac{1}{2} T_s[2n^\uparrow] + \frac{1}{2} T_s[2n^\downarrow], \quad (2.43)$$

and where the exact same relation also holds for $E_{\text{x}}[n]$.

The uniform scaling relations 2.42 can, e.g., be used to derive the precise form of a local density approximation, which is generally written as

$$\mathcal{F}[n] = \int d^3r f(n(\mathbf{r})). \quad (2.44)$$

Inserting the scaled density n_{γ} into $\mathcal{F}$, we now seek an $f(n(\mathbf{r}))$ for which it holds that $\mathcal{F}[n_{\gamma}] = \gamma^p \mathcal{F}[n]$, which leads to¹⁴

$$f(\gamma^3 n) = \gamma^{p+3} f(n), \quad (2.45)$$

from which it follows that $f(n) = n^{1+p/3}$. Now, for the non-interacting kinetic energy (first relation of 2.42), we see that $p = 2$, which gives the form of the TF approximation

$$T_{\text{s}}^{\text{TF}}[n] = A_{\text{TF}} \int d^3r n^{5/3}(\mathbf{r}). \quad (2.46)$$

In section 2.5, I discuss TF theory in more detail.

2.4.3 EXCHANGE-CORRELATION HOLE

A common starting point for working with the xc energy is to express it as

$$E_{\text{xc}}[n] = \int n(\mathbf{r}) \epsilon_{\text{xc}}[n](\mathbf{r}) d^3r, \quad (2.47)$$

where $\epsilon_{\text{xc}}[n](\mathbf{r})$ is the xc energy per particle. Writing it this way suggests that $\epsilon_{\text{xc}}[n](\mathbf{r})$ behaves as essentially a semi-local quantity. A large class of functionals are based on the way equation 2.47 is expressed.¹⁴

One way to obtain a canonical definition of $\epsilon_{\text{xc}}[n](\mathbf{r})$ is to go via the so-called adiabatic connection formula.²¹ We define the scaled Hamiltonian

$$\hat{H}_{\lambda} = \hat{T} + \lambda \hat{V}_{\text{ee}} + \hat{V}_{\lambda}, \quad \lambda \in [0, 1], \quad (2.48)$$

where λ is a coupling constant. We see that $\lambda = 0$ means no interaction, and $\lambda = 1$ gives the full interacting system. For each λ the universal functional thus takes the form

$$F_{\lambda}[n] = \min_{\Psi^{\lambda} \rightarrow n} \langle \Psi^{\lambda} | \lambda \hat{V}_{\text{ee}} + \hat{V}_{\lambda} | \Psi^{\lambda} \rangle. \quad (2.49)$$

Hence, we may express the xc energy of equation 2.36 as

$$E_{\text{xc}}[n] = \int_0^1 d\lambda \frac{d}{d\lambda} F_{\lambda}[n] - E_{\text{H}}[n]. \quad (2.50)$$

Application of the Hellman-Feynman theorem to F_{λ} means only the expectation value $\langle \hat{V}_{\text{ee}} \rangle$ of equation 2.49 comes out alive, which results in the so-called adiabatic connection formula

$$E_{\text{xc}}[n] = \int_0^1 d\lambda \langle \Psi^{\lambda}[n] | \hat{V}_{\text{ee}} | \Psi^{\lambda}[n] \rangle - E_{\text{H}}[n]. \quad (2.51)$$

In order to handle the fact that equation 2.51 contains the expectation value of the two-body operator $\hat{V}_{ee}$, we make use of the reduced density matrix

$$\rho(\mathbf{r}, \mathbf{r}') = N(N-1) \int d^3 r_3 \dots d^3 r_N |\Psi(\mathbf{r}', \mathbf{r}, \dots, \mathbf{r}_N)|^2 \quad (2.52)$$

in order to calculate it. By writing

$$\rho(\mathbf{r}, \mathbf{r}') = n(\mathbf{r})n_2(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}) [n(\mathbf{r}') + n_h(\mathbf{r}, \mathbf{r}')], \quad (2.53)$$

where $n_h(\mathbf{r}, \mathbf{r}')$ is the hole density, the xc-hole is then obtained as the coupling-constant average

$$n_{xc}(\mathbf{r}, \mathbf{r}') = \int_0^1 d\lambda n_h^\lambda(\mathbf{r}, \mathbf{r}'). \quad (2.54)$$

The adiabatic connection formula equation 2.51 makes it possible to represent the xc energy as an integral over the xc hole, i.e.,

$$E_{xc}[n] = \frac{1}{2} \int d^3 r d^3 r' \frac{n(\mathbf{r}) n_{xc}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.55)$$

The xc-hole satisfies the sum rule

$$\int d^3 r' n_{xc}(\mathbf{r}, \mathbf{r}') = -1. \quad (2.56)$$

An interpretation of the above result is that the xc energy can be thought of as an electrostatic interaction between the density n and xc hole n_{xc} .

Kohn and Mattsson³³ used the definition of the first-order density matrix γ of equation 2.52, in terms of single-particle states ϕ_i ,

$$\rho^0(\mathbf{r}, \mathbf{r}') = 2 \sum_i \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}'), \quad (2.57)$$

in order to define the exchange-hole density as

$$n_x(\mathbf{r}, \mathbf{r}') = -\frac{1}{2} \frac{|\rho^0(\mathbf{r}, \mathbf{r}')|^2}{n(\mathbf{r})}. \quad (2.58)$$

From the definition 2.58, the exchange energy can then be written in the form

$$E_x[n] = -\frac{1}{2} \int n(\mathbf{r}) R_x^{-1}(\mathbf{r}) d^3 r, \quad (2.59)$$

where $R_x^{-1}(\mathbf{r})$ is the inverse radius of the exchange hole. It takes the form

$$R_x^{-1}(\mathbf{r}) = - \int \frac{n_x(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r'. \quad (2.60)$$

2.5 Thomas-Fermi theory and orbital-free DFT

The KS method discussed in section 2.2 is significantly more computationally efficient than traditional wave function based methods. The reason for this is the effective decoupling of the electron-electron repulsion expressed in the multiplicativity of the KS potential. Nonetheless, relatively speaking, KS DFT still requires a non-trivial computational effort, mainly due to orthonormalization of the orbitals and Brillouin zone sampling. However, it is possible to circumvent the use of orbitals by direct minimization of the total energy, as prescribed in equations 2.7 and 2.11. This approach is called orbital-free, or orbital-free DFT. The fundamental problem of pure DFT is how to construct viable functional approximations of the kinetic energy. The promise, however, is a description of large and complex systems consisting of millions of atoms.

One of the earliest theories for calculating electron energies by using the density as a basic variable was formulated by Thomas and Fermi.^{60,13} They derived their theory using only a few basic assumptions based on the properties of large atoms and molecules with a large number of electrons.^{41,38,58} The energy functional of TF theory reads

$$E_{\text{TF}}[n, v] = A_{\text{TF}} \int n^{5/3}(\mathbf{r}) d^3r + \int n(\mathbf{r}) v(\mathbf{r}) d^3r + E_{\text{H}}[n], \quad (2.61)$$

where $A_{\text{TF}} = (3/10)(3\pi^2)^{2/3}$, and $E_{\text{H}}[n]$ is the Hartree energy of equation. 2.21. The kinetic energy of TF functional, i.e., the first term of equation 2.61 is that of a uniform electron gas.¹⁴ The ground state density can be computed from the Euler-Lagrange equation

$$\frac{5}{3} A_{\text{TF}} n^{2/3}(\mathbf{r}) + v(\mathbf{r}) + v_{\text{H}}[n](\mathbf{r}) - \mu = 0, \quad (2.62)$$

where μ is the chemical potential and v_{H} can be obtained by solving the Poisson equation

$$-\nabla^2 v_{\text{H}}[n](\mathbf{r}) = 4\pi n(\mathbf{r}). \quad (2.63)$$

The equations 2.62 and 2.63 are then solved so that the ground state energy and density is found. While TF theory is generally too crude of an approximation, it nevertheless forms parts of the conceptual basis of modern DFT theories in general, and pure DFT in particular. One important result is the fact that it becomes asymptotically exact in the limit of large atomic numbers Z , and electrons N .⁴⁰ A serious drawback of TF theory is however that the energetics of equation 2.61 predicts molecules to be unstable.⁵⁹

2.6 Functional approximations

In all present implementations of the KS scheme, the xc functional is approximated. The fact that there is no accepted systematic way of constructing improved approximations has led to a wealth of different functionals.

2.6.1 THE LOCAL DENSITY APPROXIMATION

The simplest possible approximation to the xc energy is the local density approximation (LDA), which was suggested by HK.³² It has been widely applied across many kinds of systems with surprisingly good results. It is defined as

$$E_{xc}^{LDA}[n] = \int n(\mathbf{r}) \epsilon_{xc}^{unif.}(n(\mathbf{r})) d^3r, \quad (2.64)$$

where $\epsilon_{xc}^{unif.}$ is the xc energy per particle of a uniform density n . For the exchange part ϵ_x , an explicit expression exists¹⁴, and ϵ_c can be calculated using quantum Monte Carlo techniques.^{7,50}

2.6.2 GENERALIZED GRADIENT APPROXIMATIONS

The major drawback of the LDA is that real systems rarely can be described by uniform densities. It has been proposed that functionals improving on the quantitative results of the LDA can be produced using the generalized gradient approximation (GGA) form for the xc energy.¹⁴ It is written as

$$E_{xc}^{GGA}[n] = \int f(n(\mathbf{r}), |\nabla n|, |\nabla n|^2, \dots) d^3r, \quad (2.65)$$

so that it also includes information about local deviations from the local density, in terms of its gradients. By introducing the so-called refinement function $F_x(s)$, a GGA can be written in the form

$$E_x^{GGA}[n] = -A_x \int n^{4/3}(\mathbf{r}) F_x(s) d^3r, \quad (2.66)$$

where

$$A_x = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{\frac{1}{3}}. \quad (2.67)$$

In the limit of uniform electron gas, i.e., the limit of vanishing density gradients to all orders, the LDA should be recovered. This is equivalent to the requirement that

$$\lim_{s \rightarrow 0} F_x(s) = 1 \quad (2.68)$$

Here, it is useful to introduce the dimensionless, reduced gradient s , and Laplacian q so that

$$\begin{aligned} s &= \frac{|\nabla n|}{2k_F n} = \frac{|\nabla n|}{2(3\pi^2)^{\frac{2}{3}} n^{\frac{4}{3}}} \\ q &= \frac{\nabla^2 n}{(2k_F)^2 n} = \frac{\nabla^2 n}{4(3\pi^2)^{\frac{4}{3}} n^{\frac{5}{3}}} \end{aligned} \quad (2.69)$$

where k_F is the Fermi wave vector, defined by

$$n = \frac{k_F^3}{3\pi^2} = \frac{3}{4\pi r_s^3}, \quad (2.70)$$

and r_s is the average distance between electrons. The values of s and q measures the rate at which the density changes on the scale of the local Fermi wavelength $\lambda_F = 2\pi/k_F$.

Similarly, for the non-interacting kinetic energy,

$$T_s^{\text{GGA}}[n] = A_\tau \int n^{\frac{5}{3}}(\mathbf{r}) F_\tau(s) d^3r, \quad (2.71)$$

with

$$A_\tau = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} \quad (2.72)$$

Again the TF result is obtained in the limit $s \rightarrow 0$. Thus, a functional approximation of the positive KED is written in the form

$$\tau[n](\mathbf{r}) = \tau_{\text{TF}}[n] F_\tau(s, q, \dots), \quad (2.73)$$

where $\tau[n](\mathbf{r})$ is the KED that was implicitly defined in equation 2.19. Hence, the total non-interacting kinetic energy of a system may be written as

$$T_s[n] = \int_\Omega d^3r \tau[n](\mathbf{r}), \quad (2.74)$$

integration is performed over the entire size Ω of the system.

2.6.3 THE PERDEW-BURKE-ERNZERHOF FUNCTIONAL

One of the most commonly used GGAs is the Perdew-Burke-Ernzerhof (PBE) functional.⁴⁸ Its derivation is based on the idea that certain exact constraints be satisfied, and its parameters are given as known fundamental constants. The refinement factor of PBE is given by

$$F_x^{\text{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa}, \quad (2.75)$$

where $\kappa = 0.804$.

2.6.4 EXTENSIONS TO THOMAS-FERMI THEORY

TF theory can be extended in a number of ways. Early on, a gradient term was by von Weizsäcker.⁶⁷ and Dirac included a term to account for exchange.⁹ In this thesis, I shall mainly be concerned with extensions to the kinetic energy term. To second order, the gradient corrections in the limit of slowly varying density in extended Thomas-Fermi (ETF) reads

$$F_\tau^{\text{ETF}}(s, q) = 1 + \frac{5}{27}s^2 + \frac{20}{9}q, \quad (2.76)$$

where s and q are defined by the relations 2.69. In chapter 3, I will discuss the properties of the positive definite KED in the limit of slowly varying densities for systems with electronic edges, using a model system approach.

2.7 Optimized effective potential

Semi-local approximations of the xc energy are computationally more effective, but yields inadequate results in many situations. Since the KS orbitals, by virtue of the HK theorem, formally are functionals of the density, it is possible in principle to construct orbital-dependent approximations of the xc energy.³⁴ In this thesis, I will primarily be dealing with a full orbital treatment of the exchange energy of equation 2.38. The most natural way to define the exchange energy in terms of KS orbitals is to use the Hartree term from many-body physics, which results in the theory of exact exchange (EXX).³⁴ Thus, the full exchange energy term, as given by equation 2.38, reads

$$E_x^{\text{EXX}}[n] = -\frac{1}{2} \sum_{\sigma=\uparrow,\downarrow} \sum_{j,k=1}^{N_\sigma} \int d^3r \int d^3r' \frac{\phi_{j\sigma}^*(\mathbf{r}) \phi_{k\sigma}^*(\mathbf{r}') \phi_{k\sigma}(\mathbf{r}) \phi_{j\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (2.77)$$

where $\phi_{i\sigma}$ is a KS orbital, and σ refers to the spin degree of freedom.

Since the EXX functional is expressed explicitly in the KS orbitals, it is not immediately clear how the functional derivative w.r.t the density is to be performed. We continue via the chain rule of functional derivatives, which yields

$$\begin{aligned} v_{x,\sigma}[n](\mathbf{r}) &= \frac{\delta E_x[\{\phi_{j\tau}\}]}{\delta n_\sigma(\mathbf{r})} \\ &= \sum_{\alpha,\beta=\uparrow,\downarrow} \sum_{i=1}^{N_\alpha} \int \frac{\delta E_x[\{\phi_{j\tau}\}]}{\delta \phi_{i\alpha}(\mathbf{r}')} \frac{\delta \phi_{i\alpha}(\mathbf{r}')}{\delta v_{\text{KS},\beta}(\mathbf{r}'')} \\ &\quad \times \frac{\delta v_{\text{KS},\beta}(\mathbf{r}'')}{\delta n_\sigma(\mathbf{r})} d^3r' d^3r'' + \text{c.c.}, \end{aligned} \quad (2.78)$$

where c.c. is the complex conjugate of the terms appearing in Eq. (2.78). The third factor of the integrand is the inverse of the non-interacting KS response function

$$\chi_{\text{KS},\sigma\beta}(\mathbf{r}, \mathbf{r}') = \delta_{\sigma\beta} \frac{\delta n_\sigma(\mathbf{r})}{\delta v_{\text{KS},\beta}(\mathbf{r}')}, \quad (2.79)$$

which is here expressed on its diagonal form w.r.t σ . From first-order perturbation theory applied to the second term of the integrand, as well as to the response function allows for a direct functional differentiation of $E_x[\{\phi_{j\sigma}\}]$ w.r.t to the orbitals $\{\phi_{j\sigma}\}$. Multiplying Eq. (2.78) with $\chi_{\text{KS},\sigma\beta}(\mathbf{r}, \mathbf{r}')$ from the right, followed by integration over $\mathbf{r}$ yields after some manipulation the OEP equation

$$\begin{aligned} &\sum_{i=1}^{N_\sigma} \int \phi_{i\sigma}^*(\mathbf{r}') [v_{x,\sigma}(\mathbf{r}') - u_{x,i\sigma}(\mathbf{r}')] \\ &\quad \times G_{\text{KS}i,\sigma}(\mathbf{r}', \mathbf{r}) \phi_{i\sigma}(\mathbf{r}) d^3r' + \text{c.c.} = 0. \end{aligned} \quad (2.80)$$

Here, the $G_{\text{KS}i,\sigma}(\mathbf{r}', \mathbf{r})$ is the KS Green's function

$$G_{\text{KS}i,\sigma}(\mathbf{r}', \mathbf{r}) = \sum_{i \neq j} \frac{\phi_{j\sigma}(\mathbf{r}') \phi_{j\sigma}^*(\mathbf{r})}{\varepsilon_{i\sigma} - \varepsilon_{j\sigma}}, \quad (2.81)$$

and

$$u_{\times}(\mathbf{r}') = \frac{1}{\phi_{i\sigma}^*(\mathbf{r})} \frac{\delta E_{\times}[\{\phi_{j\tau}\}]}{\delta \phi_{i\sigma}(\mathbf{r})}. \quad (2.82)$$

The EXX-optimized effective potential (OEP) potential exhibits a discontinuous shift that it is related to the DD that was discussed in section 2.3.1.

2.8 The electron localization function

One of the goals of this thesis is to construct semi-local methods that are able to describe systems where the electrons are localized. This information is not easily extracted from the electron density, as it is not trivial to infer from its appearance, the contribution to it from each and every KS orbital. However, the ELF^{4,53} provides a natural way to quantify the degree of electron localization in electronic systems. It has also been applied to the study of chemical bonds in molecular and solid state systems.^{54,16} The ELF is defined by

$$\text{ELF} = \frac{1}{1 + (D/D_h)^2}, \quad (2.83)$$

where D/D_h is the dimensionless quantity

$$\frac{D}{D_h} = \frac{\tau(\mathbf{r}) - \tau_W(\mathbf{r})}{\tau_{\text{TF}}(\mathbf{r})}. \quad (2.84)$$

Here, $\tau(\mathbf{r})$ is the positive definite KED, τ_W is the bosonic, or von Weiszäcker KED, given by

$$\tau_W = \frac{1}{8} \frac{|\nabla n|^2}{n}, \quad (2.85)$$

and τ_{TF} is the TF KED, defined by equation 2.61. It follows from equation 2.84 that D/D_h take on the values $(0, \infty)$, which are mapped onto $0 \leq \text{ELF} \leq 1$. In section 3.3, the harmonic oscillator model system is used to describe the ELF.

MODEL SYSTEM APPROACH TO KINETIC ENERGY FUNCTIONALS

In this chapter, I shall be concerned with two properties that are related to the non-interacting kinetic energy in KS DFT: The positive definite KED discussed in section 2.6, and the ELF that was mentioned in section 2.8. The discussion will be based on the model system approach, which was introduced by Kohn and Mattsson.³³ This approach is centred around the principle of nearsightedness that was proposed by Kohn.²⁸ In essence, it states that quantities that can be derived from the first order non-interacting density matrix, such as the electron density are local in nature. This means that their behavior near points $\mathbf{r}$ are largely determined by that at points $\mathbf{r}'$ in the vicinity of $\mathbf{r}$.

3.1 Edge electron gas

Traditionally, the uniform electron gas has formed a theoretical and conceptual basis for the development of semi-local density-based approximations of inhomogeneous electronic systems. In the limit of uniform densities, TF theory discussed in 2.5 becomes exact by definition. In real physical systems, however, there are typically electronic edges present. These are evanescent regions, which affect the spatial variations in the density throughout the entire system. This occurs, e.g., outside the nucleus of an atom, or when a vacancy is formed in a solid.³⁰

In order to accurately describe inhomogeneous systems where electronic edges are present, Kohn and Mattsson developed the concept of the edge electron gas³³, or edge gas (EG) for short. Here, I will closely follow their presentation. However, whereas the original work focused on the xc energy, I will also insert results that are relevant for the kinetic energy. The starting point for the theoretical considerations of the EG is the edge surface of a system, which may be defined by

$$v_{\text{KS}}(\mathbf{r}) - \mu = 0, \quad (3.1)$$

where $v_{\text{KS}}(\mathbf{r})$ is the exact, self-consistent KS potential, and μ is the chemical potential, situated in the middle of the non-interacting KS gap.

If the effective potential does not depend on x and y , but varies in the z -direction, the orbitals take on the form

$$\phi_\nu(\mathbf{r}) = \frac{1}{\sqrt{\Omega_A}} e^{i(k_1 x + k_2 y)} \varphi_\eta(z), \quad \langle \varphi_j | \varphi_{j'} \rangle = \delta_{\eta\eta'}, \quad (3.2)$$

where $\nu = (k_1, k_2, \eta)$, $\Omega_A = L_1 L_2$ defines the cross-sectional area of the surface, and where the wave numbers k_i are given by

$$k_i L_i = 2\pi n_i, \quad n_i \in \mathbb{Z}. \quad (3.3)$$

The single particle eigenvalues satisfy

$$\varepsilon_\nu = \frac{1}{2} (k_1^2 + k_2^2) + \varepsilon_\eta \leq \mu, \quad \forall \eta. \quad (3.4)$$

The eigenfunctions $\{\varphi_\eta\}$ and corresponding eigenvalues $\{\varepsilon_\eta\}$ may or may not be countable in η . Using equation 3.4 and letting both L_i tend to infinity, the sum to integral transition

$$\sum_{k_1, k_2} \rightarrow \frac{\Omega_A}{(2\pi)^2} \int dk_1 dk_2 \quad (3.5)$$

leads to the following expression for the density of the edge gas:

$$n_{\text{EG}}(\mathbf{r}) = \frac{1}{2\pi} \sum_{\varepsilon_\eta \leq \mu} |\mu - \varepsilon_\eta| \varphi_\eta^2(z). \quad (3.6)$$

Similarly, the positive definite KED becomes:

$$\tau_{\text{EG}}(\mathbf{r}) = \frac{1}{4\pi} \sum_{\varepsilon_\eta \leq \mu} \left\{ |\mu - \varepsilon_\eta|^2 \varphi_\eta^2(z) + |\mu - \varepsilon_\eta| \varphi_\eta'^2(z) \right\}. \quad (3.7)$$

FIGURE 3.1: Illustration of model potentials in one dimension. Panel (a) shows an atomic potential with two classical turning points. Panel (b) shows a model system where the chemical potential lies above the potential, and thus no classical turning points are present.

3.1.1 THE AIRY GAS MODEL SYSTEM

The simplest realization of the EG is the Airy gas (AG). To first order, the electronic surface behaves linearly in z , i.e., the effective potential is expected to behave as

$$v_{\text{KS}}^{\text{AG}}(z) = \begin{cases} Fz & z \geq -L \\ \infty & \text{otherwise,} \end{cases} \quad (3.8)$$

where F is the slope in the potential. In order to obtain the orbitals $\varphi_j(z)$, we solve

$$\left(-\frac{1}{2} \frac{d^2}{dz^2} + Fz \right) \varphi_\eta(z) = \varepsilon_\eta \varphi_\eta(z), \quad (3.9)$$

which results in

$$\varphi_\eta(z) = \frac{\sqrt{\pi}}{\sqrt[4]{Ll}} \text{Ai} \left(\frac{z}{l} + \frac{\varepsilon_\eta}{\varepsilon} \right), \quad (3.10)$$

and

$$\varepsilon_\eta = -\eta \sqrt{\frac{l}{L}} \pi \tilde{\varepsilon}, \text{ with } \tilde{\varepsilon} \equiv \sqrt[3]{\frac{F^2}{2}}. \quad (3.11)$$

By introducing the dimensionless quantities

$$\zeta \equiv z/l, \quad \varepsilon \equiv \varepsilon_\eta/\tilde{\varepsilon}, \quad (3.12)$$

and taking the limit $L \rightarrow \infty$, the density and the positive definite KED of the AG becomes

$$n_0^{\text{AG}}(\zeta) = \frac{1}{6\pi} \left[2\zeta^2 \text{Ai}^2(\zeta) - \text{Ai}(\zeta) \text{Ai}'(\zeta) - 2\zeta \text{Ai}'^2(\zeta) \right] \quad (3.13)$$

$$\tau_0^{\text{AG}}(\zeta) = \frac{1}{20\pi} \left[2(1 - \zeta^3) \text{Ai}^2(\zeta) + \zeta \text{Ai}(\zeta) \text{Ai}'(\zeta) + 2\zeta^2 \text{Ai}'(\zeta) \right]. \quad (3.14)$$

Here, $n^{\text{AG}}(z/l) = (1/l^3)n_0^{\text{AG}}(\zeta)$ and $\tau^{\text{AG}}(z/l) = (1/l^5)\tau_0^{\text{AG}}(\zeta)$, which follow from the scaling relations discussed in section 2.4.2. For the scaled gradient and Laplacian we have

$$s^{\text{AG}}(\zeta) = \frac{1}{2\pi} \frac{\text{Ai}'^2(\zeta) - \zeta \text{Ai}^2(\zeta)}{2(3\pi^2)^{\frac{1}{3}} [n_0^{\text{AG}}(\zeta)]^{\frac{4}{3}}}, \quad (3.15)$$

$$q^{\text{AG}}(\zeta) = \frac{1}{2\pi} \frac{\text{Ai}^2(\zeta)}{4(3\pi^2)^{\frac{2}{3}} [n_0^{\text{AG}}(\zeta)]^{\frac{5}{3}}}. \quad (3.16)$$

3.2 Quantum oscillations in the kinetic energy density

For a system with electronic edges, the gradient expansion of the ETF is no longer valid, and further corrections of the same order, in order to describe the quantum oscillations, needs to be added to the expansion. In an early work of Kohn and Sham³¹, quantum oscillations in the density was thoroughly investigated. I will now proceed to discuss the slowly varying properties of the positive definite KED for the AG model system. Outside the linear potential of equation 3.8, the density decays to zero. This creates quantum oscillations in the KED, which propagate through the entire system, all the way to its far inner region, where the limit of a slowly varying density is reached.

I shall derive an exact gradient expansion (GE), to second order in ∇n , which describe the quantum corrections in the AG, and apply it to other model systems.

3.2.1 THE SLOWLY VARYING LIMIT OF THE AIRY GAS

The properties of a solid, or a sufficiently large, finite system is generally well described by the properties of a density which varies slowly over space. This occurs on a scale where the local Fermi wavelength $\lambda_F = 2\pi/k_F(\mathbf{r})$, and the LDA of $T_s[n]$ and $E_{xc}[n]$ become asymptotically exact. In the limit of slowly varying density, the reduced density gradient and Laplacian are such that

$$s, q \ll 1, \quad (3.17)$$

which is equivalent to $s \rightarrow 0$, and $|q/s| \rightarrow 0$ everywhere.³²

For the AG, the slowly varying limit is obtained if we take the limit $\zeta \rightarrow -\infty$. In this limit, both s^{AG} and q^{AG} approach zero. In order to derive the quantum corrections in the slowly varying limit, I start by considering the refinement factor of the AG, which is given by

$$F_\tau^{\text{AG}}(\zeta) = \frac{\tau_0^{\text{AG}}(\zeta)}{\tau^{\text{TF}}[n_0^{\text{AG}}](\zeta)}, \quad (3.18)$$

where $\tau_0^{\text{AG}}(\zeta)$ is given by equation 3.14, $n_0^{\text{AG}}(\zeta)$ is given by equation 3.13, and τ^{TF} is the TF KED. A series expansion around $\zeta = -\infty$ gives me

$$\begin{aligned} F^{\text{AG}}(\zeta) = 1 + \frac{5 \left(5 + 6 \sin \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right) \right)}{48 \zeta^3} \\ - \frac{55}{192} \left(\frac{1}{\zeta} \right)^{\frac{9}{2}} \cos \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right) + \mathcal{O} \left(\frac{1}{\zeta^5} \right). \end{aligned} \quad (3.19)$$

In a similar fashion, expanding the quantities $s^{\text{AG}}(\zeta)$ and $q^{\text{AG}}(\zeta)$ yields

$$s^{\text{AG}}(\zeta) = \frac{3}{4} \left(\frac{1}{\zeta} \right)^{\frac{1}{3}} - \frac{3 \cos \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right)}{16 \zeta^3} + \mathcal{O} \left(\frac{1}{\zeta^{\frac{1}{9}}} \right), \quad (3.20)$$

and

$$q^{\text{AG}}(\zeta) = \frac{3 \left(1 + \sin \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right) \right)}{16 \zeta^3} - \frac{5}{128} \left(\frac{1}{\zeta} \right)^{\frac{9}{2}} \cos \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right) + \mathcal{O} \left(\frac{1}{\zeta^6} \right). \quad (3.21)$$

We note that, to leading order, both equation 3.19 and equation 3.21 contain the terms that are proportional to the single term

$$\frac{1}{\zeta^3} \sin \left(\frac{4}{3} \zeta^{\frac{3}{2}} \right). \quad (3.22)$$

Hence, the coefficient of a term that is proportional to the Laplacian $q^{\text{AG}}(\zeta)$ can be identified. This leaves me with, to leading order, the expression for the refinement factor in the slowly varying limit:

$$F_{\tau}^{\text{AG}}(\zeta) = 1 + \frac{10}{3} q^{\text{AG}}(\zeta) - \frac{5}{48} \frac{1}{\zeta^3}. \quad (3.23)$$

Since $s^{\text{AG}}(\zeta)$ is non-oscillatory to leading order in ζ , it can be inverted to give

$$\zeta^{\text{AG}}(s) = \frac{1}{2} \left(\frac{9}{2s^2} \right)^{\frac{1}{3}}. \quad (3.24)$$

Inserting equation 3.24 into the equation 3.23 results in an expression, which to second order in $|\nabla n|$ becomes

$$F_{\tau}^{\text{AG}}(s, q) = 1 - \frac{5}{27} s^2 + \frac{10}{3} q. \quad (3.25)$$

Figure 3.2 shows the oscillations in the far inner region of the AG, where $s, q \rightarrow 0$, as $\zeta \rightarrow -\infty$. It is seen that the ETF-GE fails to accurately capture the quantum oscillations, whereas they are exactly reproduced by equation 3.25. Figure 3.3 shows the surface region of the AG, where the vertical line indicates the position of the classical turning point. Here, both of the expansions are expected to fail, because of the presence of the surface. It is seen that the oscillations of equation 3.25 deviates more in the vicinity of the surface, when compared to the ETG-GE. This is a general feature, which I discuss at more length in Paper I.

At this point, it is interesting to ask to what extent the AG-GE of equation 3.25 is capable of also describing the quantum oscillations that appear in other kinds of model systems. Is it valid for all types of systems with electronic edges? If not, is there a way of characterizing the system, which allows us to say anything with regards to its validity? In order to address this question, I will consider a few key model systems. The way these model systems differ, is the nature of their respective eigenvalue spectrums.

FIGURE 3.2: The ETF GE and the AG-GE in Eq. (3.25) compared to the exact KED for the AG, in the slowly varying limit.

FIGURE 3.3: The ETF GE and the AG-GE in equation 3.25 compared to the exact KED for the AG, across the surface region.

3.2.2 BOHR ATOMS AND THE ISOTROPIC HARMONIC OSCILLATOR

I start by considering two finite systems. The first one, the large Bohr atom is formed by letting the non-interacting particles settle in a hydrogen-like potential²³, which is given by

$$v_{\text{KS}}^{\text{Bohr}}(\mathbf{r}) = -\frac{Z}{|\mathbf{r}|}, \quad \text{with } \varepsilon_{\eta} = -\frac{Z^2}{2} \frac{1}{\eta^2}, \quad (3.26)$$

and where $\eta = 1, 2, \dots$ is the principal quantum number and there is an η^2 -fold degeneracy in quantum numbers l and m . I let the system be filled with particles up to $\eta = N_{\text{shell}}$, so that it contains $N = 2 \sum_{\eta=1}^{N_{\text{shell}}} \eta^2 = (1/3)N_{\text{shell}}(N_{\text{shell}} + 1)(2N_{\text{shell}} + 1)$ particles (including the spin degree of freedom). Figure 3.4 shows the oscillations for an atom which is filled up to the 30th shell. In the region where the electron density is slowly varying, the oscillatory behavior is well described by the AG-GE in equation 3.25, whereas the ETF give oscillations with too small amplitude.

FIGURE 3.4: The performance of the ETF GE and the AG-GE of equation 3.25 in the slowly varying limit for a large Bohr atom.

A similar finite system is the closed, isotropic harmonic oscillator. The potential and the corresponding eigenvalues are given by

$$v_{\text{KS}}^{\text{HO}}(\mathbf{r}) = \frac{1}{2}\omega^2 r^2, \quad \varepsilon_{\eta} = \omega \left(\eta + \frac{3}{2} \right) \quad (3.27)$$

where ω is the angular frequency, r is the radial distance from the position of equilibrium, and $\eta = 0, 1, \dots$ denotes the collective principal quantum number of the three oscillating modes. Furthermore, I let $w = \omega/2$ be defined as the curvature parameter of the potential. Hence, at curvature ω the $(\eta + 1)^{\text{th}}$ energy level is filled, where

$$\eta = \left\lfloor \frac{1}{2w} - \frac{3}{2} \right\rfloor. \quad (3.28)$$

Hence, systems with smaller w has a wider potential and contain more non-interacting particles. In figure 3.5 I show the slowly varying limit of an isotropic harmonic oscillator filled up to the 30th energy level close to the center, where the density is slowly varying. The results are analogous to what is found for the Bohr atom.

FIGURE 3.5: The ETF GE and the AG-GE of equation 3.25 for the case of an isotropic harmonic oscillator.

3.2.3 THE MATHIEU GAS

I now consider a infinite model system, the Mathieu gas (MG).³ Its potential is

$$v_{\text{KS}}^{\text{MG}}(\mathbf{r}) = \lambda [1 - \cos(pz)], \quad (3.29)$$

where λ is the amplitude and p is the wave vector assigned to the oscillation. In this case, I introduce the scaled parameters given by

$$\bar{\lambda} = \frac{\lambda}{\mu}, \quad \bar{p} = \frac{p}{2k_{\text{F}}}, \quad \text{and} \quad \bar{z} = k_{\text{F}}z, \quad (3.30)$$

where $k_{\text{F}} = \sqrt{2\mu}$ is the Fermi wave vector of the uniform electron gas, which in the semi-classical limit can be regarded as independent of position. Given the value of $\bar{\lambda}$ and $\bar{p}$, the chemical potential μ will be above or below the maximums of v_{KS} , i.e., the MG can be made to represent either a disturbed uniform electron gas or a system with an infinite number of classically forbidden regions.

The MG model system is a continuum of systems, parameterized in the two-dimensional parameter space, defined by the dimensionless quantities $\bar{\lambda}$ and $\bar{p}$. In the limit $\bar{\lambda} \rightarrow 0$, i.e., the limit of vanishing oscillation, the free electron gas is obtained. By contrast, when $\bar{\lambda} \rightarrow \infty$ the occupied energy levels in the z -direction become those of an harmonic oscillator. The parameter space of the

MG is shown in figure 3.6. Any sequence $\sqrt{2\bar{\lambda}\bar{p}^2} \rightarrow 0$ and $\bar{p} \rightarrow 0$ represents a possible limit of slowly varying density, of which there are infinitely many. A path such that $\bar{\lambda} > 1/2$ means the chemical potential stays below the maximum values of the potential, and the system will have infinitely many classically forbidden regions, repeated throughout the entire system. On the other hand, a path with $\bar{\lambda} < 1/2$ means the chemical potential never touches the maximum values of the potential, and the system approaches a slowly varying limit along

a path that resembles a perturbed uniform electron gas. The path which is given exactly by the intersecting value $\bar{\lambda} = 1/2$ represents systems where the chemical potential exactly tangents the maximum values of the potential. This path is indicated by a solid line of figure 3.6 and splits the entire parameter space into two distinct regions, the regime of harmonic oscillators (values of $\bar{\lambda} > 1/2$) and the free-electron regime (values of $\bar{\lambda} < 1/2$).

FIGURE 3.6: The parameter space of the MG. The shaded area corresponds to values of μ that belong to one of the possible energy bands of the energy spectrum in the z -dimension. The light area correspond to values of μ in the free-electron continuum in x and y between the bands. Yellow lines define the values of the chemical potential situated precisely on the band edges of the energy spectrum in z dimension. The straight lines in the parameter space indicate various paths along which the limit of slowly varying electron density is reached. The black line splits the parameter space into two distinct regions. The blue dashed line defines a path to the limit of slowly varying density for which $\bar{\lambda} < 1/2$, and the red dashed line defines a path to the limit of slowly varying density for which $\bar{\lambda} > 1/2$.

I will now discuss the performance of the ETF GE and the AG-GE in the slowly varying limit of the MG in three different cases. In the first case I let $\bar{\lambda} = 0.1$ and $\bar{p} = 0.02$. In this case, we are in the free-electron regime and thus the system does not have any classical turning points. The result is shown in figure 3.7. It becomes evident that the AG-GE does not describe the exact behavior of the MG in this case. Next we consider the case where $\bar{\lambda} = 0.5$ and $\bar{p} = 0.015$, for which the chemical potential μ precisely intersects the line between the harmonic oscillator regime and the free-electron regime respectively. Now, μ precisely tangents the maximum values of the cosine potential. The result is shown in figure 3.8.

FIGURE 3.7: The slowly varying limit of MG in the free-electron regime, where $\bar{\lambda} = 0.1$ and $\bar{p} = 0.02$.

In figure 3.7, i.e., the free-electron case, the ETF GE performs well. This is expected, since in this case, no classical turning points are present. On the other hand, while the AG-GE is similar in shape when compared to the exact refinement factor, there is an offset between the two curves. This behavior persists when classical turning points (and thus quantum oscillations) are introduced to the system. Indeed, from figures 3.8 and 3.9, it is seen that the amplitude of the exact oscillations are well captured by the AG-GE, however with an offset. This very same behavior is found in the Hermite gas model system, which is also studied in detail in Paper I.

FIGURE 3.8: The slowly varying limit of the MG in the free-electron regime, where $\bar{\lambda} = 0.5$, $\bar{p} = 0.015$.

FIGURE 3.9: The slowly varying limit of the MG in the harmonic oscillator regime, where $\bar{\lambda} = 0.9$, $\bar{p} = 0.02$.

3.2.4 CHARACTERIZATION OF MODEL SYSTEMS IN TERMS OF EIGENVALUES

While the results of Paper I suggest that there are clear cases for which the AG-GE of equation 3.25 is able to very accurately capture the exact quantum oscillations in the positive KED, it is not trivial to offer an a priori reason why this should be the case. In Paper I, however, some speculations are given.

One way to characterize a given model system is to examine the single-particle eigenvalues that correspond to the solutions of equation 3.2. It seems to be the case that when a system has, in addition to electronic edges, a fully isotropic set of eigenvalues, then the AG-GE is valid. This is the case for the AG, the large Bohr atom, and the three-dimensional isotropic harmonic oscillator, but not the case for the MG and the open harmonic oscillator (Hermite gas).

3.2.5 DEVELOPMENT OF A KINETIC ENERGY FUNCTIONAL

It is a primary goal in the further theoretical development of DFT to construct more accurate approximations to the xc energy $E_{xc}[n]$ and to the non-interacting kinetic energy $T_s[n]$. Essential to any semi-local approximation is the description of the slowly varying regime of an inhomogeneous system. The work I have presented here suggests that the slowly varying properties of the positive definite KED is influenced by the presence of electronic edges. This means that the GE of the ETF is insufficient, and needs to be adjusted to include quantum corrections of the same order in the reduced gradient s and Laplacian q .

The AG-GE of equation 3.25 includes the quantum oscillations that arise from a linear surface. Moreover, I have demonstrated that the AG-GE is surprisingly general and applies to systems that show a strong isotropy in the single-particle eigenvalue spectrum. However, it fails for systems without electronic edges, such as the MG model system. In addition, the mere presence of electronic edges seems not to be enough to the AG-GE to be valid, which again is the case for the MG, where there is a strong anisotropy in the eigenvalue spectrum. In this case, the anisotropy arises because of the fact that the eigenvalues in the z -direction are not densely spaced, as they are in the other two dimensions.

The question thus arises whether it is possible to construct a valid positive definite KED, which holds for all limits of slowly varying electron density. In other words, is it possible for such a semi-local approximation to differentiate between a situation where the ETF applies vs. when the AG-GE applies, based only on the semi-local information available in $n(\mathbf{r})$, s and q ?

One way to answer this question is to look at the path in a contour plot of s^2 and q as we take the limit of slowly varying density. In figure 3.10, I show the behavior of s^2 versus q for the Bohr atom model system. The slowly varying regime is reached in the limit $s^2, q \rightarrow 0$, and thus corresponds to the origin in the figure. As the Bohr atom becomes larger and larger with increasing quantum numbers N_{shell} , we see how loops start to form as we get closer and closer to the origin. This indicates that the path along which the slowly varying regime is reached may be highly non-trivial in its character, and may be hard trace out for more complex systems.

In order to investigate to what extent it is possible to obtain the information of the slowly varying properties of a system, I show in figure 3.11 the behavior

FIGURE 3.10: Contour plot of s^2 vs. q for all values of r for three Bohr atoms, filled up to the 1st, 2nd, and 10th shell. The number of loops increases with N_{shell} and the curve moves closer to the origin, which demonstrates how the limit of slowly varying density is reached in this system within an intermediate region of r values.

of s^2 versus q for both a jellium surface³⁵ and a particular choice of a MG. These systems are fundamentally different: while the slowly varying regime of the jellium surface does contain quantum oscillations, the MG does not. It becomes apparent that the jellium surface approaches the slowly varying limit in a much more complicated way, when compared to the MG. Moreover, the fact that the curves intersect with one another and the local values of the ETF GE and AG-GE are different suggests that the both expansions may not be the same in these points. The conclusion is thus that one single GE that is able to capture both types of slowly varying properties may not exist. This feature has to be dealt with in one way or another, when considering future semi-local approximations of the non-interacting KED.

Finally, while AG-KED captures the oscillations in the slowly varying limit of systems where electronic edges are present, it is less energetically favourable close to the surface region. By contrast, the ETF GE performs better when integrated over the entire system. This fact points in the general direction that a semi-local approximation of the positive definite KED which is aimed at capturing the properties of systems with electronic edges must somehow be modified so that it is able to both describe the oscillations far inside the system and the surface region. Such a construction could possibly result in more accurate total energies, as compared to a functional based on the ETF GE.

FIGURE 3.11: Contour plot s^2 versus q in the limit of slowly varying density for the case of a jellium surface ($r_s = 1$) as $z \rightarrow -\infty$ and a Mathieu gas ($\bar{\lambda} = 0.1$) as $\bar{p} \rightarrow 0$.

3.3 Parameterization of the ELF in the harmonic oscillator model system

The harmonic oscillator gas is a model system where the states in the z -direction are confined by the potential²⁰

$$v_{\text{KS}}(z) = \frac{\omega}{2} z^2, \quad (3.31)$$

where ω is the strength of the potential. It is convenient to introduce the scaled length-coordinate $\bar{z}$, so that

$$\bar{z} = \frac{z}{l}, \quad \text{where } l = \sqrt{\frac{1}{\omega}}. \quad (3.32)$$

The harmonic oscillator model is thus a two-parameter model, described in terms of the parameters $\alpha \in (0, \infty)$, and $\bar{z} \in (-\infty, \infty)$. The parameter α is defined via the relation²⁰

$$\mu = (\alpha + 1/2) \frac{1}{l^2}, \quad (3.33)$$

where μ is the chemical potential of equation 3.1, and which sets the degree of confinement (low α is highly confined) via $N = \lfloor \alpha \rfloor$, where N is the index of the highest level the states in the z -direction can occupy. The remainder $0 \leq \alpha - N < 1$ denotes the continuous bands in the xy -direction. Hence, for a given μ , a larger α corresponds to a wider parabola in the potential, which gives a less confined system. Within this context, I shall only be concerned with values of α that belong to the range $\alpha \in (0, 1]$. The density of the harmonic oscillator gas is given by

$$l^3 n(\bar{z}) = \begin{cases} \frac{\alpha}{\pi^{3/2}} e^{-\bar{z}^2}, & \alpha \leq 1 \\ \frac{(\alpha + 2(\alpha - 1)\bar{z}^2)}{\pi^{3/2}} e^{-\bar{z}^2}, & 1 \leq \alpha \leq 2. \end{cases} \quad (3.34)$$

Furthermore, the positive definite KED is

$$l^5 \tau(\bar{z}) = \frac{e^{-\bar{z}^2}}{2\pi^{3/2}} \alpha (\alpha + \bar{z}^2) \quad \alpha \leq 1, \quad (3.35)$$

and for $1 \leq 2$,

$$l^5 \tau(\bar{z}) = \frac{e^{-\bar{z}^2}}{2\pi^{3/2}} \left(\alpha^2 + 2(\alpha - 1) + (\alpha + 2(\alpha - 1)(\alpha - 3)) \bar{z}^2 + 2(\alpha - 1)\bar{z}^4 \right). \quad (3.36)$$

Now, D/D_h of equation 2.84 for $\alpha \in (0, 1]$ is given by

$$\frac{D}{D_h} = \frac{5}{3} \alpha \frac{1}{(3\sqrt{\pi} \alpha e^{-\bar{z}^2})^{2/3}}, \quad (3.37)$$

and for $\alpha \in [1, 2]$ the following expression holds:

$$\frac{D}{D_h} = \frac{5}{3} \frac{\left(\alpha(\alpha^2 + 2(\alpha - 1)) + 2\alpha(\alpha - 1)(2\alpha - 1)\bar{z}^2 + 4(\alpha - 1)^3 \bar{z}^4 \right)}{(\alpha + 2(\alpha - 1)\bar{z}^2)^2 (3\sqrt{\pi}(\alpha + 2(\alpha - 1)\bar{z}^2)e^{-\bar{z}^2})^{2/3}}. \quad (3.38)$$

FIGURE 3.12: Illustration of the electron localization function in the harmonic oscillator model system, for different values of α . When α exceeds the value 1, a new level gets occupied.

In Paper II, the harmonic oscillator model system is used to construct a density functional approximation of the ELF, in terms of the reduced density gradient s and Laplacian q of equation 2.69. The parameterization of ELF can thus be used in cases where the electron density is known, but not the kinetic energy density. In Paper II the resulting parameterization is evaluated on a number of systems.

PROPERTIES OF THE AK13 EXCHANGE FUNCTIONAL

Commonly used semilocal approximations to the xc energy in KS DFT generally work well for the prediction of properties related to the total energy of the system, such as equilibrium structures, phonon dispersions, or solid formation energies. While these approximations have been successfully applied across a wide range of systems, they typically fail to incorporate a number of important exchange features, such as the relative offset between the potentials in the molecular dissociation limit¹⁷, and the DD⁴⁹ of section 2.3. The DD has a special role in KS DFT — when added to the single-particle gap, the sum total equals the fundamental gap, which is one of the most important properties of any material.

The above mentioned exchange features can to various degrees be captured by methods that rely on higher-order theory.³⁴ Using these, however, invariably leads to an increase in computational effort. Thus, there is a significant incentive to incorporate the missing features into semi-local DFT methods.

A few steps in this direction has been taken with the introduction of a model potential of Becke and Johnson, as well as an exchange functional of Armiento and Kümmel functional. In this chapter, I will focus on the energetic properties of the AK13 functional. I will also demonstrate the capability of the BJ/AK13 type potentials to model the EXX DD, by only using the information contained in the KS eigenvalues.

4.1 The Becke-Johnson model potential

Becke and Johnson⁵ suggested an approximate exchange potential, that closely resembles that of the EXX within the OEP method, for radially symmetric atoms. They defined their potential as

$$v_{x\sigma}^{\text{BJ}}(\mathbf{r}) = v_{x\sigma}^{\text{Slater}}(\mathbf{r}) + v^c(\mathbf{r}), \quad (4.1)$$

where the first term given by

$$v_{x\sigma}^{\text{Slater}}(\mathbf{r}) = -\frac{1}{n_\sigma} \int \frac{|\sum_i m_{i\sigma} \phi_{i\sigma}^*(\mathbf{r}) \phi_{i\sigma}(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (4.2)$$

is the Slater exchange. Furthermore, $\phi_{i\sigma}$ is the i th KS orbital, with occupation number $m_{i\sigma}$, and σ refers to the spin. The second term of equation 4.1 is given by

$$v^c(\mathbf{r}) = C_{\Delta v} \sqrt{\frac{2\tau_\sigma(\mathbf{r})}{n_\sigma(\mathbf{r})}}, \quad \text{with } C_{\Delta v} = \frac{1}{\pi} \sqrt{\frac{5}{12}}. \quad (4.3)$$

Here, $\tau_\sigma(\mathbf{r})$ is the positive definite KED, discussed in section 2.6. A crucial feature of the Becke-Johnson (BJ) potential is its asymptotic properties: In the limit far outside a finite system, the potential takes on the positive, finite value²

$$\lim_{|\mathbf{r}| \rightarrow \infty} v_x^{\text{BJ}}(\mathbf{r}) = C \sqrt{-2\varepsilon_{I\sigma}}, \quad (4.4)$$

where $I = \text{HO}$ is the highest occupied KS eigenvalue. This feature is key, when it comes to e.g., mimicking the potential step structure of the EXX potential: it is seen from equation 4.4, that the asymptote of the potential changes discontinuously as a new KS orbital gets occupied. Since there is no effect on the electron density, upon a constant shift in the effective potential, BJ can effectively be shifted uniformly by the constant amount given by equation 4.4, to produce a DD, which is similar to that of EXX OEP, as was discussed in section 2.7.

4.2 The Armiento-Kümmel exchange functional

A serious drawback of the BJ model potential is that it is not the functional derivative of any total energy functional.²⁷ Armiento and Kümmel were able to derive an exchange functional with a well-defined functional derivative, which shares many of the attractive features of the BJ potential.¹ The starting point in their derivation is the asymptotic behavior of the density far outside a finite system:⁶⁶

$$n(\mathbf{r}) \rightarrow m_{\text{HO}} \mathcal{C}_N e^{-2\sqrt{-2\varepsilon_{\text{HO}}^0}|\mathbf{r}|}, \text{ as } |\mathbf{r}| \rightarrow \infty, \quad (4.5)$$

where $\mathcal{C}_N$ is a system-dependent constant of normalization, $\varepsilon_{\text{HO}}^0$ is the highest occupied KS eigenvalue that corresponds to a zero-aligned potential, and m_{HO} is the occupation number. Using equation 4.5, a refinement function can be derived.¹ It is given by

$$F_x^{\text{AK13}}(s) = 1 + B_1 s \ln(1 + s) + B_2 s \ln[1 + \ln(1 + s)], \quad (4.6)$$

where $B_2 = \mu_{\text{GE}} - B_1$, $B_1 = (3/5)\mu_{\text{GE}} + 8\pi/15$, and $\mu_{\text{GE}} = 10/81$ are non-empirical constants, derived from first principles arguments.¹ In the asymptotic limit, far outside the system, the Armiento-Kümmel functional (AK13) potential takes on the value which is given by

$$\lim_{|\mathbf{r}| \rightarrow \infty} v_x^{\text{AK13}}(\mathbf{r}) = \frac{A_x^2 Q_x^2}{2} \left(1 \pm \sqrt{1 - \frac{4\varepsilon_{\text{HO}}}{A_x^2 Q_x^2}} \right), \quad (4.7)$$

where

$$Q_x = \{\sqrt{2}/(3(3\pi^2)^{1/3})\}B_1, \quad (4.8)$$

and the plus sign generally applies. Equation 4.7 shows that since the asymptotic value of the potential depends on the highest occupied KS eigenvalue, it changes discontinuously as a new orbital gets occupied. The zero-aligned potential is given by

$$v_x^0[n](\mathbf{r}) = v_x^{\text{AK13}}[n](\mathbf{r}) + v_x^{\text{shift}}[n], \quad (4.9)$$

where $v_x^{\text{shift}}[n] \equiv -\lim_{|\mathbf{r}| \rightarrow \infty} v_x^{\text{AK13}}[n](\mathbf{r})$. From this, it follows that, as for usual potentials,⁴⁹

$$\lim_{|\mathbf{r}| \rightarrow \infty} v_x^0[n](\mathbf{r}) = 0. \quad (4.10)$$

A key difference between the AK13 functional and other semi-local approximations, such as PBE is its asymptotic behavior for large reduced density gradients s . While the refinement factor of PBE of section 2.6.3 remains bounded for large s , the AK13 diverges in the asymptotic limit which corresponds to finite step in its potential, far outside a finite system. This is illustrated in figure 4.1.

FIGURE 4.1: Exchange refinement factors of PBE and AK13. While the refinement factor of PBE remains finite for large reduced density gradients s , AK13 diverges asymptotically as $\sim s \ln s$. In this asymptotic limit, the AK13 potential takes on a finite value $\sim \sqrt{-\varepsilon_{\text{HO}}}$, where ε_{HO} is the highest occupied KS eigenvalue.

4.3 Dealing with inaccurate energetics

A serious drawback of the AK13 functional is that it gives inaccurate total energies and energetics. This heavily affects, e.g., the equilibrium properties of solids such lattice constants and bulk moduli, which of course is highly undesirable. In this section, I discuss a possible remedy to this issue. One may pose the question whether or not this deficiency can be said to originate from how the AK13 functional effectively changes the structure of the KS orbitals, compared to other common semi-local approximations. In order to determine this, one may apply some other xc-energy functional non-self-consistently on the obtained *self-consistent* AK13 density.

To set the stage for this discussion, we consider the general, non-unique partition of the xc energy

$$E_{xc}[n] = E_{xc}^{sl}[n] + E_{xc}^{nl}[n], \quad (4.11)$$

hence consisting of semi-local and non-local parts. Here, $E_{xc}^{nl}[n]$ is considered to be non-local in the sense that may depend on the density n throughout the entire system. Now, I let $E_x^{sl}[n] = E_x^{AK13}[n]$, and $E_c^{sl}[n]$ be some suitable semi-local correlation energy functional.

Now, given some external potential v , I denote by $n^{sl}[v]$ the density resulting from a self-consistent KS calculation of only the semi-local part of the problem. Doing this also gives me the reduced gradient $s^{sl}[v]$. The contribution from $E_{xc}^{nl}[n]$, such that the energetics of LDA is exactly reproduced is thus defined as

$$E_{xc}^{nl}[n, v] = -A_x \int n^{4/3}(\mathbf{r}) \left[B_1 s^{sl} \ln(1 + s^{sl}) + B_2 s^{sl} \ln(1 + \ln(1 + s^{sl})) \right] d^3r, \quad (4.12)$$

where A_x is given by equation 2.67, and $E_c^{sl}[n] = E_c^{LDA}[n]$. I emphasize that the dependence of v in equation 4.12 means the functional is not a valid KS functional.

4.3.1 SELF-CONSISTENT SCHEME FOR VOLUME RELAXATIONS OF SOLIDS

To illustrate how the computational scheme in equation 4.12 works for in practice for the energetics of solids, I now go on to consider structural relaxation of bulk Si. In figure 4.3 it is seen that a the AK13 functional combined with LDA correlation yields a much too large equilibrium lattice constant. However, when LDA is applied non-self-consistently on the converged AK13 density, the equilibrium lattice constant is substantially improved. As is seen from figure 4.2, the electron bands remain unaffected by this procedure defined by equation 4.12.

FIGURE 4.2: Band structure along high symmetry points of solid Si in the diamond structure, using self-consistent LDA (solid blue) and self-consistent AK13 (red dashed). The bands have been calculated at equilibrium lattice constants. In both cases, the exchange functionals have been paired with LDA correlation.

FIGURE 4.3: Equation of state of solid Si in the diamond structure. Using self-consistent AK13 (yellow dashed) results in a much too large equilibrium lattice constant. However, the application of LDA on the converged AK13 density (red dashed) results in a lattice constant much closer to that of self-consistent LDA (solid blue).

4.4 Modeling of the exact-exchange derivative discontinuity

As was discussed in section 4.1, the BJ potential shows a step structure in the potential because of its asymptotic dependence on the highest occupied KS eigenvalue. This kind of step structure is intimately related to the derivative discontinuity of the xc potential of section 2.3.1. It is interesting to ask to what extent the model of the discontinuous shift implicit in AK13 and BJ potentials is qualitatively correct or not. In terms of the zero-aligned eigenvalues $\varepsilon_{\text{HO}}^0$, the asymptotic value of the AK13 potential is

$$\lim_{|\mathbf{r}| \rightarrow \infty} v_{\text{x}}^{\text{AK13}}(\mathbf{r}) = A_{\text{x}} Q_{\text{x}} \sqrt{-\varepsilon_{\text{HO}}^0}, \quad (4.13)$$

where Q_{x} was defined in equation 4.8. The discontinuous shift $\Delta_{\text{x}}^{\text{AK13}}$ of the AK13 potential, as N passes through an integer M is given by the exchange part of equation 2.35, taken in the limit $|\mathbf{r}| \rightarrow \infty$,

$$\Delta_{\text{x}}^{\text{AK13}} = A_{\text{x}} Q_{\text{x}} \left(\sqrt{-\varepsilon_{\text{LU}}^0} - \sqrt{-\varepsilon_{\text{HO}}^0} \right), \quad (4.14)$$

where $\varepsilon_{\text{HO}}^0$ and $\varepsilon_{\text{LU}}^0$ is the highest occupied, and lowest unoccupied KS eigenvalue, respectively. In paper IV, I provide a detailed study of the capability of the shift given by equation 4.14 to model the EXX DD in the case of spherically symmetric ions.

MACHINE LEARNING

So far I have devoted the discussion to the formal treatment of many-electron systems in the context of DFT. In some situations, such as in a sophisticated stability analysis, it is required to compute the basic properties of a vast number of systems. This makes the use of first principles methods such as DFT restrictive, even at a semi-local level of theory. A very computationally effective way to deal with this issue is to resort to methods based on machine learning (ML). ML is a data-driven method, where predictions are based on induction from some prior existing set of data.

ML has been used increasingly in various contexts in theoretical materials science, only in the past few years.^{57,42,56,26} One of the reasons for this is its capability of making predictions at an extremely fast rate. While the use of first principles methods such as DFT requires the solution of self-consistent equations, such as the KS equations in order to obtain the ground state properties of a system, an ML algorithm relies only on matrix operations that typically only requires milliseconds to be executed. Recently, ML has been used to predict atomization energies of organic molecules.^{51,19} Thus far however, the work on extending ML methods for the treatment of periodic systems is still much in its cradle.

In this chapter, I introduce the main concepts of ML and define my notation. I discuss the concept of feature vector representations, and then go on to discuss the construction of ML models for the prediction of solid formation energies.

5.1 Kernel ridge regression

In this thesis, I will primarily focus on kernel ridge regression (KRR), which is a standard method used in ML. It is a nonlinear form of regression. The goal is to approximate some unknown function $y(x)$, where x is an input feature-vector, i.e., some representation of, e.g., a molecule, or crystal. Now, we let $\mathbf{y} = (y_1, y_2, \dots, y_n)$ denote n known data points, for a given set of n feature vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$. The sought approximation $f(x, \boldsymbol{\alpha})$ is then written in the form

$$f(x, \boldsymbol{\alpha}) = \sum_{i=1}^N \alpha_i k(x, x_i), \quad (5.1)$$

where k works as a function ansatz, and

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n), \quad (5.2)$$

is sought. The approximation is then found from the optimization problem

$$\min_{\boldsymbol{\alpha}} \left\{ \sum_{i=1}^n [y_i - f(x_i, \boldsymbol{\alpha})]^2 + \sum_{i,j=1}^n \alpha_i k(x_i, x_j) \alpha_j \right\}. \quad (5.3)$$

The solution is given by

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}, \quad (5.4)$$

where $\mathbf{I}$ is the identity matrix, and $\mathbf{K} = k(x_i, x_j)$ is the kernel matrix. The hyperparameter λ is a regularization parameter used to prevent over-fitting. I will restrict myself to the use of the Laplacian kernel function

$$k(x_i, x_j) = e^{-\frac{\|x_i - x_j\|_1}{\sigma}}, \quad (5.5)$$

where σ is a length scale hyperparameter, and $\|\cdot\|_1$ is the L^1 norm. While other types of kernels exist, the Laplacian kernel has previously successfully been used to compute molecular atomization energies.^{51,19}

5.1.1 GENERALIZATION ERROR

In order to evaluate the performance of an ML model, a convenient measure of the error in predicting new data outside of the set $\mathbf{x}$ is given by the so-called generalization error. Of all available data $\mathbf{x}$ and $\mathbf{y}$ from the solution of equation 5.4, a subset is excluded. Then the mean absolute error for the prediction of the excluded data is then evaluated. The subset of $\mathbf{x}$ used in equation 5.4 is called the training set $\mathbf{x}_{\text{train}}$, and the remaining set is called the test set $\mathbf{x}_{\text{test}}$. The generalization error is then given by

$$\text{MAE}_{\text{test}} = \frac{1}{m} \sum_{i \in \mathbf{x}_{\text{test}}} |y_i - f(x_i, \alpha)|. \quad (5.6)$$

Here, m is the size of the test set. In order for the generalization error be accurately estimated by equation 5.6, the set $\mathbf{x}_{\text{test}}$ needs to be of sufficient size. Naturally, the respective sizes of the sets $\mathbf{x}_{\text{test}}$ and $\mathbf{x}_{\text{train}}$ must be chosen with some care. However, it is possible to sidestep this issue by using k -fold cross-validation¹⁹, in combination with random sampling. This means that MAE_{test} is computed as the mean of the MAE_{test} of several independent ML runs. For each run, $\mathbf{x}_{\text{test}}$ is then constructed by selecting a specific number of entries randomly out of the full data set.

Similarly, $\text{MAE}_{\text{train}}$ is defined by substituting $\mathbf{x}_{\text{train}}$ for $\mathbf{x}_{\text{test}}$ in equation 5.6. This measure determines the precision with which the machine reproduces the data it has been trained on. If the value of $\text{MAE}_{\text{train}}$ too high, it is likely that the machine has too few degrees of freedom for the data to be accurately described. By contrast, a low value of $\text{MAE}_{\text{train}}$ suggests a high flexibility of the fitting function. In this case, there might be a potential risk of overfitting and that good transferability with a sufficiently large data set might be reached.

5.2 Feature vector representations

Of central importance to any ML scheme is the feature vector representation, also known as the descriptor. The descriptor works as a mapping from the atomistic description of the system into a form that is well-suited for statistical learning. Moreover, it should have some defining properties, in order for it to be rendered useful.

Besides being simple, i.e., generate as little computational effort as possible, any feature vector x should be

1. Complete and non-degenerate, i.e., it should include all features in the input associated with the underlying problem.
2. Compact and unique, i.e., there should be no two instances of different systems that result in the same x .
3. If two instances of input have similar y , then they should be represented by x in such a way that the distance $||x_1 - x_2||$ between them is sufficiently small.

In general, it is non-trivial to construct a useful descriptor. In addition, the above criterion introduce a certain arbitrariness to their construction: The precise nature of the physics one is interested in is what ultimately determines their usefulness in any given situation.

5.2.1 PRINCIPAL COMPONENT ANALYSIS

Once a feature vector representation has been constructed, it is often necessary to have at one's disposal some means of evaluating its performance. One such a mean is provided by principal component analysis (PCA). The simple idea behind PCA, is to reduce the dimensionality of the data set, while keeping intact as much of its information as possible. One starts with performing a singular value decomposition of $\mathbf{x}$, by writing

$$\mathbf{x} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^T. \quad (5.7)$$

The reduced PCA dataset $\mathbf{z}$ of rank k is subsequently formed from the k first rows of $\mathbf{U}$ and $\mathbf{\Sigma}$, so that

$$\mathbf{z} = \mathbf{U}_k\mathbf{\Sigma}_k. \quad (5.8)$$

The dataset $\mathbf{z}$ may often be used for visualisation purposes. Moreover, if the representation under consideration is believed to not be sufficiently compact, in the sense I have discussed above, one can try to replace $\mathbf{x}$ with its PCA counterpart, in order to extract a representative lower-dimensional part.

5.3 Coordinate-based representations of extended systems

In this section I discuss the main aspects of the feature vector representations presented in Paper V. I start with a brief review of the basic properties of the so-called Coulomb matrix representation, which has previously been used in the context of an accurate prediction of molecular atomization energies.

5.3.1 THE COULOMB MATRIX REPRESENTATION FOR MOLECULES

One of the main goals of this thesis has been to develop ML models for the accurate prediction of solid formation energies. This requires the construction of adequate vector representations for extended systems. As a first step to extend the framework to an atom by atom matrix-based representation of the crystal structure, I will consider a generalization of an ML method that has been applied to predict the atomization energies of molecules^{51,19,43}. These works considered a Coulomb matrix feature vector representation in the form

$$x_{ij} = \begin{cases} 0.5Z_i^{2.4} & i = j \\ Z_iZ_j\phi(\|\mathbf{r}_i - \mathbf{r}_j\|_2) & i \neq j, \end{cases} \quad (5.9)$$

where $\phi(r) = 1/r$ is the Coulomb potential, $\mathbf{r}_i$ and Z_i are the position and atomic number of the i th atom. Thus, the off-diagonal elements of this representation are the pair-wise Coulomb interactions between the atomic cores in the system. The diagonal elements are chosen in such a way that they result from an exponential fit to the potential energy of a free atom.

An unwanted feature of the Coulomb matrix vector representation is that a single molecule can give rise to several Coulomb matrices. This occurs because there are multiple ways to order the atoms. While there are some ways to deal with the issues of indexing the atoms, these solutions tend to introduce other issues, such as differentiability.^{44,52}

5.3.2 THE EWALD-SUM VECTOR REPRESENTATION

In Paper V, a number of extensions of the Coulomb matrix representation are presented and evaluated. In order to illustrate the general idea behind the construction of the vector representations in that paper, I shall here provide a brief summary of one of them, which is given by the Ewald-sum vector representation.

I start by considering an atom by atom matrix, like the one discussed in the previous section, and where each element now corresponds to each pair of atoms in a primitive unit cell, so that each element represents the full Coulomb interaction that results from all infinite repetitions of these two atoms in the lattice. Thus, each element can be viewed as the generalization of those of the Coulomb matrix of equation 5.9. In this case however, we need to take the infinite repetition of the entire lattice into account. This is accomplished by writing

$$x_{ij} = \frac{1}{N}Z_iZ_j \sum_{k \neq l} \varphi(\|\mathbf{r}_k - \mathbf{r}_l\|_2). \quad (5.10)$$

Here, summation runs over k for each atom i in the unit cell and its N closest equivalent atoms, and equivalently for l and j . Hence, taking the limit $N \rightarrow$

∞ represents the full electrostatic interaction between the infinitely repeated atoms, that are equivalent to atoms i and j in the primitive cell.

In order to deal with the unfavourable convergence which is associated with the sum that appears in equation 5.10, I proceed by invoking the Ewald summation technique.¹² Hence, let us write

$$x_{ij} = x_{ij}^{(r)} + x_{ij}^{(m)} + x_{ij}^0, \quad (5.11)$$

where the short range interaction $x_{ij}^{(r)}$, is given by

$$x_{ij}^{(r)} = Z_i Z_j \sum_{\mathbf{L}} \frac{\operatorname{erfc}(a \|\mathbf{r}_i - \mathbf{r}_j + \mathbf{L}\|_2)}{\|\mathbf{r}_i - \mathbf{r}_j + \mathbf{L}\|_2}, \quad i \neq j, \quad (5.12)$$

so that each $\mathbf{L}$ lie inside a sphere with some cut-off radius $L_{\max}$. The long range interaction $x_{ij}^{(m)}$ is given by the expression

$$x_{ij}^{(m)} = \frac{Z_i Z_j}{\pi V} \sum_{\mathbf{G}} \frac{e^{-\|\mathbf{G}\|_2^2 / (2a)^2}}{\|\mathbf{G}\|_2^2} \cos[\mathbf{G} \cdot (\mathbf{r}_i - \mathbf{r}_j)], \quad i \neq j, \quad (5.13)$$

where summation runs over all reciprocal lattice vectors $\mathbf{G}$ inside a sphere of a radius which is set by a cutoff $G_{\max}$, and V is the volume of the unit cell. The constant term x_{ij}^0 is written as

$$x_{ij}^0 = \begin{cases} -\left(Z_i^2 + Z_j^2\right) \frac{a}{\sqrt{\pi}} - \left(Z_i^2 + Z_j^2\right) \frac{\pi}{2Va^2} & i \neq j \\ -Z_i^2 \frac{a}{\sqrt{\pi}} - Z_i^2 \frac{\pi}{2Va^2} & i = j. \end{cases} \quad (5.14)$$

The number a works as a convergence parameter, which sets the overall rate of convergence of x_{ij} . While there have been some suggestions on how to set the value of a , I use in this work

$$a = \sqrt{\pi} \left(\frac{0.01M}{V} \right)^{1/6}, \quad (5.15)$$

where M is the number of atoms in the unit cell.

5.3.3 GENERAL RESULTS ON DATASETS

The data set that was used to evaluate the feature vector representations of Paper V was extracted from the Materials Project database.⁴⁶ It consists of roughly 4k different solids, that were sampled without obvious order. All of the representations discussed in Paper V give similar results: In all cases, and with a training set size of 3k, the generalization error MAE_{test} is of the order of about 0.5eV/atom. This is to be compared with much smaller generalization error that was obtained for the molecular atomization energies⁵¹, which is of about 0.02 eV/atom.

The apparent discrepancy between the molecular atomization energies and the results of Paper V can to some degree be explained by the fact that the training set that was considered was too limited. More importantly though, is the fact that the composition and diversity between the two data sets is starkly different.

5.4 *Symmetry-restricted vector representations*

The results of Paper V suggests that the sparseness in the chemical space might affect the performance of the training of the machine. One reason for this sparseness, is that the training set itself that was used in Paper V is too limited. While it is possible to further improve on the generalization error by considering a larger training set size, this improvement alone is not sufficient.

One way of accounting for the sparseness of the chemical space is to restrict the training to one single crystal structure, pertaining to a certain symmetry. This is the main idea behind the results of Paper VI. Paper VI is specifically concerned with the Elpasolite structure, which is the most abundant crystal that appears in the ICSD database. In that paper, it is demonstrated that the generalization error can be made arbitrarily small, given a sufficiently large training set. In particular, for a training set size of 10k, we reach a generalization error of 0.1eV, for the prediction of about 2M Elpasolites.

CONCLUSIONS

In this thesis, I have presented a number of results within two main areas of research: density functional theory and machine learning. The results of Paper I provide further insight into properties of systems with slowly varying electron densities in vicinity of electronic edges. Furthermore, the gradient approximation derived in Paper I can be a basis for, e.g., a functional for the non-interacting kinetic energy density, possibly in combination with results of Paper II. Paper II also suggests that it may be possible to model the electron localization function in terms of semi-local density-based quantities. While this result may be of importance in and of itself, the electron localization function could also be used as an important tool in the ever continuing search for more accurate semi-local functionals. Several works use or have suggested to use ELF or ELF-like indices within the construction of approximate exchange-correlation functionals.

Paper III and Paper IV deal with the general problem of incorporating the exact exchange properties in semi-local density-functional theory. This is no doubt a relevant endeavour, albeit highly non-trivial. Further work will be necessary to fully reconcile an improved orbital description, with sufficiently accurate energetics, on a semi-local level of theory.

The machine learning models developed in Paper V and VI depict a promising path towards an accurate and swift determination of material properties, entirely from a data-driven approach. Still, there are many challenges to address. While this ML model delivers impressive accuracy in its predictions of substitutions into a single crystal structure, this is often not the most central question to address in an applied problem, e.g., to make predictions for all crystal structures within a given phase diagram for a compound. Future work is needed towards more generally applicable ML models.

APPENDIX

7.1 Positive definite KED of closed, radially symmetric systems

Consider N non-interacting fermions in the spherically symmetric potential $v_s(r)$. The separable solutions to $\hat{H}_s \phi_\nu = \epsilon_\nu \phi_\nu$ are

$$\phi_\nu = \phi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi), \quad (7.1)$$

where each R_{nl} is a radial distribution function and Y_{lm} are spherical harmonics. We wish to calculate the positive kinetic energy density

$$\tau(r, \theta, \phi) = \sum_{nlm} |\nabla \phi_{nlm}|^2. \quad (7.2)$$

To this end, we calculate the gradient $\nabla \phi_{nlm}$ using the product rule

$$\nabla \phi_{nlm} = \nabla R_{nl} Y_{lm} + R_{nl} r \nabla Y_{lm} \quad (7.3)$$

where $\Psi_{lm} = r \nabla Y_{lm}$ is the vector spherical harmonics along the ϕ -direction. The Ψ_{lm} obey Unsöld's theorem⁶⁴, i.e.,

$$\sum_{m=-l}^l |\Psi_{lm}|^2 = \frac{1}{4\pi} (2l+1)(l+1)l, \quad (7.4)$$

which mirrors the fact that spatial densities must be radially symmetric. In the same way, the Y_{lm} 's obey the well known addition theorem i.e.,

$$\sum_{m=-l}^l |Y_{lm}|^2 = \frac{1}{4\pi} (2l+1) \quad (7.5)$$

Hence, we are left with the expression for the radial KED:

$$\begin{aligned} \tau^{\text{Radial}}(r) &= \frac{1}{4\pi} \sum_{n=1}^N \sum_{l=0}^{n-1} (2l+1) \\ &\times \left(\left[\frac{\partial}{\partial r} R_{nl}(r) \right]^2 + l(l+1) \left[\frac{R_{nl}(r)}{r} \right]^2 \right). \end{aligned} \quad (7.6)$$

7.1.1 THE BOHR ATOM

Consider non-interacting electrons that are bound by the hydrogen-like potential

$$v_s^{\text{Bohr}}(\mathbf{r}) = -\frac{Z}{|\mathbf{r}|}, \quad (7.7)$$

where Z is the atomic number. The system consists of a finite number of N electrons and the KS orbitals are the familiar functions

$$\phi_{\eta lm}(r, \theta, \phi) = R_{\eta l}(r)Y_{lm}(\theta, \phi) \quad (7.8)$$

where the $R_{\eta l}$'s are proportional to Laguerre polynomials and the Y_{lm} 's are the normalized spherical harmonics. The eigenvalues are

$$\epsilon_{\eta} = -\frac{Z^2}{2} \frac{1}{\eta^2}. \quad (7.9)$$

We take the system to be filled with particles up to principal quantum number $\eta = N_{\text{shell}}$, which means that it contains $N = 2 \sum_{\eta=1}^{N_{\text{shell}}} \eta^2 = (1/3)N_{\text{shell}}(N_{\text{shell}} + 1)(2N_{\text{shell}} + 1)$ particles (including the spin degree of freedom). Furthermore, as explained in the main text, the atomic number Z scales the system, so we set $Z = N_{\text{shell}}^2$ to simplify the problem. Using Eq. (7.6) the positive KED τ becomes

$$\begin{aligned} \tau^{\text{Bohr}}(r) &= \frac{1}{4\pi} \sum_{\eta=1}^{N_{\text{shell}}} \sum_{l=0}^{\eta-1} (2l+1) \left(\frac{2Z}{\eta} \right)^3 \frac{(\eta-l-1)!}{2\eta[(\eta+l)!]} \\ &\times \left(\left[\frac{\partial}{\partial r} \left(e^{-\frac{Zr}{\eta}} \left(\frac{2Zr}{\eta} \right)^l L_{\eta-l-1}^{2l+1} \left(\frac{2Zr}{\eta} \right) \right) \right]^2 \right. \\ &\left. + l(l+1) \left[\frac{1}{r} e^{-\frac{Zr}{\eta}} \left(\frac{2Zr}{\eta} \right)^l L_{\eta-l-1}^{2l+1} \left(\frac{2Zr}{\eta} \right) \right]^2 \right). \end{aligned} \quad (7.10)$$

7.2 Properties of the Mathieu gas

Given a potential which is constant in two spatial directions and varies along the third, z say, according to

$$v_s^{\text{MG}}(z) = \lambda(1 - \cos(pz)), \quad (7.11)$$

where λ is the amplitude and p is the wave vector of the oscillation, the solutions to the corresponding eigenvalue problem are of the form

$$\begin{aligned} \varphi_\eta^{\text{MG}}(z) &= \frac{1}{L_3} \left[ce_\eta \left(z, -\frac{1}{2} \frac{\bar{\lambda}}{\bar{p}^2} \right) + ise_\eta \left(z, -\frac{1}{2} \frac{\bar{\lambda}}{\bar{p}^2} \right) \right] \\ &= \frac{1}{L_3} e^{i\eta\bar{p}\bar{z}} \sum_{k \in \mathbb{Z}} c_{2k}^\eta e^{i2k\bar{p}\bar{z}}, \end{aligned} \quad (7.12)$$

where $\eta\bar{p}k_F L_3 = 2\pi n_3$ ($n_3 \in \mathbb{Z}$), L_3 is the size of the system, measured in units of z . Here, the functions ce_η and se_η are the real even and odd Mathieu functions respectively. We have introduced the scaled parameters

$$\bar{\lambda} = \frac{\lambda}{\mu}, \quad \bar{p} = \frac{p}{2k_F}, \quad \bar{z} = k_F z, \quad (7.13)$$

where $k_F = \sqrt{2\mu}$ is the Fermi wave vector of the uniform electron gas. The c_{2k}^η are found from the relation

$$(2k + \eta)^2 c_{2k}^\eta - \frac{\bar{\lambda}}{2\bar{p}} (c_{2k-2}^\eta + c_{2k+2}^\eta) = a \left(\eta, \frac{\bar{\lambda}}{2\bar{p}^2} \right) c_{2k}^\eta, \quad (7.14)$$

and they are normalized according to

$$\sum_{k \in \mathbb{Z}} |c_{2k}^\eta|^2 = 1. \quad (7.15)$$

The eigenvalue associated with the φ_η are

$$\frac{\epsilon_\eta}{\mu} = \bar{\lambda} + \bar{p}^2 a \left(\eta, \frac{\bar{\lambda}}{2\bar{p}^2} \right). \quad (7.16)$$

The positive kinetic energy density becomes

$$\begin{aligned} \frac{\tau^{\text{MG}}(z)}{\tau_u} &= \frac{5}{2} \bar{p} \int_0^{\eta_m} d\eta \left\{ \frac{1}{2} \left(1 - \bar{p}^2 a_\eta - \bar{\lambda} \right)^2 \right. \\ &\quad \times \left[ce_\eta^2(\bar{p}\bar{z}, \bar{q}) + se_\eta^2(\bar{p}\bar{z}, \bar{q}) \right] + \bar{p}^2 (1 - \bar{p}^2 a_\eta - \bar{\lambda}) \\ &\quad \times \left. [ce_\eta'^2(\bar{p}\bar{z}, \bar{q}) + se_\eta'^2(\bar{p}\bar{z}, \bar{q})] \right\}, \end{aligned} \quad (7.17)$$

where $\tau_u = k_F^5 / (10\pi^2)$ and a_η are the eigenvalues of Eq. (7.16), η_m is the energy of the highest occupied state, and $\bar{q} = -(1/2)\bar{\lambda}/\bar{p}^2$.

7.3 Positive definite KED of the Hermite gas

For the HG, the potential varies along the z -axis as an HO, i.e.,

$$v_s^{\text{HG}}(z) = \frac{\omega^2}{2} z^2. \quad (7.18)$$

The eigenfunctions are the familiar

$$\varphi_\eta^{\text{HG}}(z) = \left(\sqrt{\frac{\omega}{\pi}} \frac{1}{2^\eta \eta!} \right)^{1/2} H_\eta(\sqrt{\omega} z) e^{-\frac{\omega z^2}{2}}, \quad (7.19)$$

with corresponding eigenvalues

$$\epsilon_\eta = \omega \left(\eta + \frac{1}{2} \right), \quad (7.20)$$

for $\eta = 0, 1, \dots$. Introducing the scaled parameters

$$w = \frac{\omega}{\mu}, \quad N(\mu) = \left\lfloor \frac{1}{w} - \frac{1}{2} \right\rfloor, \quad \bar{z} = k_F z, \quad (7.21)$$

where $k_F = \sqrt{2\mu}$, and $N(\mu)$ is the number of occupied z -orbitals respectively. Division with the KED of the free electron gas yields the dimensionless quantity

$$\begin{aligned} \frac{\tau^{\text{HG}}(\bar{z})}{\tau_u} &= \frac{1}{4\pi} \sqrt{\frac{w}{2\pi}} \sum_{\eta=0}^{N(\mu)} \frac{1}{2^\eta \eta!} \left[1 - w \left(\eta + \frac{1}{2} \right) \right] \\ &\times \left[\frac{1}{2} H_\eta^2 \left(\sqrt{\frac{w}{2}} \bar{z} \right) e^{-\frac{1}{2} w \bar{z}^2} \left(1 - w \left(\eta + \frac{1}{2} \right) \right) \right. \\ &\left. + \left[\frac{d}{d\bar{z}} \left(H_\eta \left(\sqrt{\frac{w}{2}} \bar{z} \right) e^{-\frac{\omega \bar{z}^2}{4}} \right) \right]^2 \right]. \end{aligned} \quad (7.22)$$

The curvature parameter w that describes the wideness of the potential parabola determines the number of occupied orbitals in the z -direction.

BIBLIOGRAPHY

- [1] R. Armiento and S. Kümmel. Orbital Localization, Charge Transfer, and Band Gaps in Semilocal Density-Functional Theory. *Physical Review Letters*, 111(3):036402, 2013.
- [2] R. Armiento, S. Kümmel, and T. Körzdörfer. Electrical response of molecular chains in density functional theory: Ultranonlocal response from a semilocal functional. *Physical Review B*, 77(16):165106, 2008.
- [3] R. Armiento and A. E. Mattsson. Subsystem functionals in density-functional theory: Investigating the exchange energy per particle. *Physical Review B*, 66(16):165117, 2002.
- [4] A. D. Becke and K. E. Edgecombe. A simple measure of electron localization in atomic and molecular systems. *The Journal of Chemical Physics*, 92:5397, 1990.
- [5] A. D. Becke and E. R. Johnson. A simple effective potential for exchange. *The Journal of Chemical Physics*, 124:221101, 2006.
- [6] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, 1927.
- [7] D. M. Ceperley and B. J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Physical Review Letters*, 45(7):566–569, 1980.
- [8] M. K. Y. Chan and G. Ceder. Efficient band gap prediction for solids. *Physical Review Letters*, 105:196403, 2010.
- [9] P. A. M. Dirac. Note on Exchange Phenomena in the Thomas Atom. *Mathematical Proceedings of the Cambridge Philosophical Society*, 26(03):376–385, 1930.
- [10] R. M. Dreizler and E. K. U. Gross. *Density Functional Theory : An Approach to the Quantum Many Body Problem*. Springer-Verlag, Berlin, 1990.
- [11] H. Eschrig. *The fundamentals of density functional theory*. Teubner-Texte zur Physik, Teubner, 1996.
- [12] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.
- [13] E. Fermi. Un Metodo Statistico per la Determinazione di alcune Prioprietà dell’Atomo. 6:602–607, 1927.
- [14] C. Fiolhais, F. Nogueira, and M. Marques. *A Primer in Density Functional Theory*. Springer-Verlag, NY, 2003.
- [15] P. Fulde. *Electron Correlations in Molecules and Solids*. Springer-Verlag, Berlin, 1991.
- [16] C. Gatti. Chemical bonding in crystals: new directions. *Zeitschrift für Kristallographie-Crystalline Materials*, 220:399–457, 2005.

- [17] O. V. Gritsenko and E. J. Baerends. Effect of molecular dissociation on the exchange-correlation Kohn-Sham potential. *Physical Review A*, 54(3):1957–1972, 1996.
- [18] E. K. U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory*. Hilger, Bristol, 1991.
- [19] K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko, and K-R. Müller. Assessment and validation of machine learning methods for predicting molecular atomization energies. *Journal of Chemical Theory and Computation*, 9(8):3404–3419, 2013.
- [20] F. Hao, R. Armiento, and A. E. Mattsson. Subsystem functionals and the missing ingredient of confinement physics in density functionals. *Physical Review B*, 82(11):115103, 2010.
- [21] J. Harris. Adiabatic-connection approach to Kohn-Sham theory. *Physical Review A*, 29(4):1648–1659, 1984.
- [22] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York, 2009.
- [23] O. J. Heilmann and E. H. Lieb. Electron density near the nucleus of a large atom. *Physical Review A*, 52(5):3628–3643, 1995.
- [24] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Physical Review*, 136(3B):B864–B871, 1964.
- [25] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Reviews of Modern Physics*, 87:897, 2015.
- [26] Vu K., J. C. Snyder, L. Li, M. Rupp, B. F. Chen, T. Khelif, K-R. Müller, and K. Burke. Understanding kernel ridge regression: Common behaviors from simple functions to density functionals. *International Journal of Quantum Chemistry*, 115:1115–1128, 2015.
- [27] A. Karolewski, R. Armiento, and S. Kümmel. Electronic excitations and the becke-johnson potential: The need for and the problem of transforming model potentials to functional derivatives. *Physical Review A*, 88:052519, 2013.
- [28] W. Kohn. Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms. *Physical Review Letters*, 76(17):3168–3171, 1996.
- [29] W. Kohn. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Reviews of Modern Physics*, 71(5):1253–1266, 1999.
- [30] W. Kohn and L. J. Sham. Quantum Density Oscillations in an Inhomogeneous Electron Gas. *Physical Review*, 137(6A):A1697–A1705, 1965.

- [31] W. Kohn and L. J. Sham. Quantum Density Oscillations in an Inhomogeneous Electron Gas. *Physical Review*, 137(6A):A1697–A1705, 1965.
- [32] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133–A1138, 1965.
- [33] Walter Kohn and Ann E. Mattsson. Edge Electron Gas. *Physical Review Letters*, 81(16):3487–3490, 1998.
- [34] S. Kümmel and L. Kronik. Orbital-dependent density functionals: Theory and applications. *Reviews of Modern Physics*, 80(1):3–60, 2008.
- [35] N. D. Lang and W. Kohn. Theory of Metal Surfaces: Charge Density and Surface Energy. *Physical Review B*, 1(12):4555–4568, 1970.
- [36] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 76(12):6062–6065, 1979.
- [37] M. Levy and J. P. Perdew. Hellmann-Feynman, virial, and scaling requisites for the exact universal density functionals. Shape of the correlation potential and diamagnetic susceptibility for atoms. *Physical Review A*, 32(4):2010–2021, 1985.
- [38] E. H. Lieb. Thomas-fermi and related theories of atoms and molecules. *Reviews of Modern Physics*, 53(4):603–641, 1981.
- [39] E. H. Lieb. Density functionals for coulomb systems. *International Journal of Quantum Chemistry*, 24(3):243–277, 1983.
- [40] E. H. Lieb and B. Simon. Thomas-Fermi Theory Revisited. *Physical Review Letters*, 31(11):681–683, 1973.
- [41] E. H. Lieb and B. Simon. The Thomas-Fermi theory of atoms, molecules and solids. *Advances in Mathematics*, 23(1):22–116, 1977.
- [42] A. Lopez-Bezanilla and O. A. von Lilienfeld. Modeling electronic quantum transport with machine learning. *Physical Review B*, 89:235411, 2014.
- [43] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K-R. Müller, and O. A. von Lilienfeld. Machine learning of molecular electronic properties in chemical compound space. *New Journal of Physics*, 15(9):095003, 2013.
- [44] J. E. Moussa. Comment on “fast and accurate modeling of molecular atomization energies with machine learning”. *Physical Review Letters*, 109:059801, 2012.
- [45] G. L. Oliver and J. P. Perdew. Spin-density gradient expansion for the kinetic energy. *Physical Review A*, 20:397–403, 1979.

- [46] S. P. Ong, W.D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *Computational Materials Science*, 68:314–319, 2013.
- [47] R. G. Parr and W. Yang. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, 1989.
- [48] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(3865), 1996.
- [49] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz. Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy. *Physical Review Letters*, 49(23):1691–1694, 1982.
- [50] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244–13249, 1992.
- [51] M. Rupp, A. Tkatchenko, K-R. Müller, and O. A. von Lilienfeld. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical Review Letters*, 108:058301, 2012.
- [52] M. Rupp, A. Tkatchenko, K-R. Müller, and O. A. von Lilienfeld. Rupp et al. reply:. *Physical Review Letters*, 109:059802, 2012.
- [53] A. Savin, O. Jepsen, J. Flad, O. K. Andersen, H. Preuss, and H. G. and v. Schnering. Electron localization in solid-state structures of the elements: the diamond structure. *Angewandte Chemie International Edition in English*, 31(87):187, 1992.
- [54] A. Savin, R. Nesper, S. Wengert, and T. F. Fässler. Elf: The electron localization function. *Angewandte Chemie International Edition in English*, 36:1808–1832, 2003.
- [55] L. Schimka, J. Harl, and G. Kresse. Improved hybrid functional for solids: the hsesol functional. *The Journal of Chemical Physics*, 134(2):024116, 2011.
- [56] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross. How to represent crystal structures for machine learning: Towards fast prediction of electronic properties. *Physical Review B*, 89:205118, 2014.
- [57] J. C. Snyder, Rupp M., Hansen K., K-R. Müller, and Burke K. Finding density functionals with machine learning. *Physical Review Letters*, 108:253002, 2012.
- [58] L. Spruch. Pedagogic notes on Thomas-Fermi theory (and on some improvements): atoms, stars, and the stability of bulk matter. *Reviews of Modern Physics*, 63(1):151–209, 1991.
- [59] E. Teller. On the Stability of Molecules in the Thomas-Fermi Theory. *Reviews of Modern Physics*, 34(4):627–631, 1962.

- [60] L. H. Thomas. The Calculation of Atomic Fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23(05):542–548, 1927.
- [61] C. Toher, Filippetti A., S. Sanvito, and Burke K. Self-interaction errors in density-functional calculations of electronic transport. *Physical Review Letters*, 95:146402, 2005.
- [62] F. Tran and P. Blaha. Accurate Band Gaps of Semiconductors and Insulators with a Semilocal Exchange-Correlation Potential. *Physical Review Letters*, 102(22):226401, 2009.
- [63] F. Tran, P. Blaha, and Schwarz K. Band gap calculations with becke–johnson exchange potential. *Journal of Physics: Condensed Matter*, 19(19):196208, 2007.
- [64] A. Unsöld. Beiträge zur quantenmechanik der atome. *Annalen der Physik*, 82, 1927.
- [65] R. van Leeuwen. Density functional approach to the many-body problem: Key concepts and exact functionals. *Advances in Quantum Chemistry*, 43:25–94, 2003.
- [66] R. van Leeuwen and E. J. Baerends. Exchange-correlation potential with correct asymptotic behavior. *Physical Review A*, 49(4):2421–2431, 1994.
- [67] C. F. v. Weizsäcker. Zur Theorie der Kernmassen. *Zeitschrift für Physik*, 96(7-8):431–458, July 1935.

LIST OF FIGURES

2.1	Total energy as a function of the time-averaged particle number	12
3.1	Example of model potentials in one dimension	24
3.2	Slowly varying limit of the Airy gas	28
3.3	Surface region of the Airy gas	28
3.4	Slowly varying limit of the Bohr atom	29
3.5	Slowly varying limit of the isotropic harmonic oscillator	30
3.6	Parameter space of the Mathieu gas model systems	31
3.7	Slowly varying limit of the MG in the free-electron regime . . .	32
3.8	The MG where the harmonic oscillator regime intersects with the free-electron gas regime	32
3.9	A MG in the harmonic oscillator regime	32
3.10	Scaled gradient s^2 versus Laplacian q for the Bohr atom	34
3.11	s^2 vs. q in the limit of slowly varying density for a jellium surface system and the MG	35
3.12	Illustration of the electron localization function in the harmonic oscillator model system	37
4.1	Exchange refinement factors of PBE and AK13	42
4.2	Electron bands of Si for LDA and AK13	44
4.3	Volume relaxation of Si in the diamond structure	44

Papers

LIST OF PUBLICATIONS AND MY CONTRIBUTION

- Paper I I derived the expressions for almost all of the model systems that we studied. I did all the numerical calculations, and took part in the analysis and discussion of the results. I wrote the first draft, and the text was finalized jointly.
- Paper II I took part in the planning of the project, the derivation of key results, and produced some of the data. I took part in the interpretation of the results, and is writing the paper.
- Paper III I performed all the electronic structure calculations, and took the major role in their analysis and discussion. I did most of the writing.
- Paper IV I performed all the electronic structure calculations, and took the major role in their analysis and discussion. I wrote the paper.
- Paper V I took part in the original planning of the project, the analysis and interpretation of the results. The paper was written jointly.
- Paper VI I took part in the planning of the project. I helped with setting up the electronic structure calculations. I participated in the analysis, and interpretation of the results. The paper was written jointly.

Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

[http://urn.kb.se/resolve? urn:nbn:se:liu:diva-139767](http://urn.kb.se/resolve?urn:nbn:se:liu:diva-139767)