HARVARD UNIVERSITY
Graduate School of Arts and Sciences

THESIS ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Engineering and Applied Sciences

have examined a thesis entitled

"Efficient Multiscale Simulation of Simple Metallic Systems"

presented by Nicholas Isaac Choly

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature .....
Typed name: Professor E. Kaxiras

Signature .....
Typed name: Professor M. Aziz

Signature .....
Typed name: Professor H. Ehrenreich

Signature .....
Typed name: Professor Z. Suo

Date: June 9, 2004

Efficient Multiscale Simulation of Simple Metallic Systems

A thesis presented

by

Nicholas Isaac Choly

to

The Division of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Applied Physics

Harvard University

Cambridge, Massachusetts

2004

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Thesis advisor

Efthimios Kaxiras

Author

Nicholas Isaac Choly

Efficient Multiscale Simulation of Simple Metallic Systems

Abstract

The steady increase in computational resources and numerical sophistication has brought about a new approach in physical simulation. The methods that comprise this approach are known as multiscale methods, and have the defining characteristic of combining several simulation methods together, rendering tractable physical problems that no single simulation method can resolve. We have developed an approach for coupling quantum-mechanical and classical methods for the efficient simulation of multiscale problems in simple metals.

The present multiscale method employs orbital-free density functional theory, in which fictitious orbitals are never introduced. We review the theory, and describe the state-of-the-art functionals asociated with it. We have developed an efficient simulation code for performing orbital-free density functional theory calculations, and we describe the methods developed to treat the functional minimization problem.

One of the biggest barriers hindering the widespread use of orbital-free methods is that only local pseudopotentials can be used, and hence the powerful machinery of norm-conserving pseudopotentials is inapplicable. We develop a similar machinery for local pseudopotentials, and we report on the application of these methods.

We solve several problems associated with the efficient use of orbital-free density functional methods. Certain orbital-free methods are formulated in reciprocal space

and are applicable to periodic systems. Incorporation of these methods in a multiscale setting requires that the effects of periodicity be absent. A direct translation of the methods to real space is extremely inefficient. Motivated by these considerations, we have developed an efficient method for applying orbital-free methods to non-periodic systems.

We also overcome an algorithmic problem with the calculation of ionic forces in grid-based electronic structure methods in general. We develop and test an efficient method for computing ionic forces that scales quasi-linearly ($O(N \log N)$) with the system size.

We develop and examine a multiscale method for the simulation of simple metallic systems. Part of the system is treated quantum-mechanically, and part classically, with consideration made to the coupling of the two. We apply this method to several physical systems.

Contents

List of Figures

List of Tables

Citations to Previously Published Work

Large portions of Chapters 4, 5, and 6 have appeared in the following papers:

N. Choly and E. Kaxiras, "Kinetic energy density functionals for non-periodic systems," Sol. Sta. Comm. **121**, 281 (2002), `cond-mat/0112040`.

N. Choly and E. Kaxiras, "Fast method for force computations in electronic structure calculations," Phys. Rev. B **67** 155101 (2003), `cond-mat/0210389`.

N. Choly, G. Lu, W. E, and E. Kaxiras, "Multiscale simulations in simple metals: a density-functional based methodology," submitted to Phys. Rev. B, `cond-mat/0404414`.

Acknowledgments

When I think about my time in graduate school, I realize that I'm lucky to have participated in the experience. I learned much had a lot of fun, and this is due in large part to the kindness and wisdom of those around me.

I have to thank Umesh Waghmare, who patiently showed me the ropes of density functional theory. Paul Maragakis, you showed me how useful computer skills can be, and you were a most reliable source for interesting conversation. Gang Lu, thank you for being a great friend and collaborator.

I am grateful for the financial support I received from the NSF, and later from Lawrence Livermore Labs. I am indebted to Riad Manaa for bringing the latter source of funding to my attention, and for helping me with all things Livermore.

Susan Wieczorek has been a tremendous ally on the administrative side of things. I think that without Susan, the Division of Engineering and Applied Sciences would long ago have exploded in a shower of improperly done or undone paperwork.

I am thankful for all of the great scientists who have lent their ear and given me invaluable feedback, including Weinan E, Emily Carter, Bill Curtin, Henry Ehrenreich, and Mike Aziz. To Prof. Kaxiras I owe a lot of thanks, for letting me be a part of his research group, for showing me how to stay focused on the important things, and for just being an all-around excellent person.

My parents are largely to thank for any success that I have achieved. I attribute my love of learning to them, as they have always had interesting things to teach me. Like any worthy challenge, this PhD has been stressful at times, and my family has always been there to make it easier. Knowing that I have another home, even if it is far away, has always been a comfort that I can't imagine being without.

I am very fortunate that I have had the companionship of Jeannie Park for almost all of my time at Harvard. I don't like to think about where I would be without her constant emotional support, love, and energy. It is much more pleasant to think about where we've come together.

Dedicated to my family.

Chapter 1

Introduction and background

1.1 A look toward the future of electronic structure calculation

As of the writing of this thesis, the behavior and properties of a great many materials and molecules are known. It stands as a testament to the success of theoretical physics that a great many of these properties can be found by purely theoretical methods. One of the more important players in the arena of quantitative theoretical methods is the density functional theory (DFT) of (primarily) Hohenberg, Kohn, and Sham[5, 6], which has been successful for treating many condensed matter systems.

Since the formal inception of DFT in the 1960's, it has evolved into a general-purpose tool that many researchers can easily use to study many condensed matter or molecular systems. Its historical growth is closely tied with that of the computer, and the success of DFT is due in no small part to the many numerical techniques and im-

plementational details that have been developed along the way. These methods have evolved somewhat of a "standard model" of DFT. In the condensed matter-materials science world, the standard model of DFT involves iterative minimization of the Kohn-Sham energy in a plane-wave basis with ions represented by norm-conserving pseudopotentials. In the chemistry community, this framework is modified slightly, and there the basis set of choice consists of Gaussian orbitals. The degree of maturity of these methods is exemplified by the widespread use of a number of robust and powerful software packages for DFT calculations, including VASP[7], CASTEP[8], and GAUSSIAN[9].

However as good as established DFT methods are at reproducing results from the past, they will probably not be up to the demands of future (and many current) simulation needs. The reason for this is tied up in the computational cost of the Kohn-Sham method. The number of floating-point operations required to compute the energy of a physical system scales with the cube of the system size, assuming that the system is of roughly constant density. And although the speed of computers has been growing faster than any power of time (i.e. exponentially), this will not continue forever, and the situation is still such that presently "standard model" DFT calculations cannot exceed around one thousand atoms.

Many important physical problems that would benefit from a density functional treatment, however, exceed this size. For instance the behavior of important biological molecules, the microscopic machines underlying all life, is expected to be a consequence of the quantum mechanical laws governing their constituent electrons and nuclei, but many of these molecules consist of many thousands of atoms. In the

realm of materials science, the elements of dislocation theory have long been known, and it is well established that dislocations and their interactions play a primary role in determining the toughness of materials. But in order to theoretically determine these interactions, a quantum mechanical treatment of the systems is necessary; and unfortunately the smallest system of atoms in which some of these interactions can be represented also consist of many thousands of atoms. And obviously there are many other instances where the exploration of some phenomenon requires the examination of a system of large size, or at the least, examination of a large-sized system is the only apparent route of elucidating the phenomenon. Thus a means of extending DFT simulations to larger systems would be extremely useful, and the study of methods that achieve this has become a field unto itself.

In the past decade or so there has been much activity in this field. One might distinguish several broad categories of these approaches (see Fig. 1.1). One line of attack involves making clever approximations to the standard Kohn-Sham method that allows for a more efficient solution algorithm. Another category of approaches consists of the "multiscale" methods, applicable to certain systems in which the detailed physics occurs only in a small subset of the entire system. Ideas from both of these categories are involved in the present work, and in fact end up intertwining in an interesting manner.

1.1.1 Approximations resulting in efficient algorithms

The methods belonging to the first of these two categories themselves fall into several sub-categories. The different sub-categories are distinguished by main ideas be-

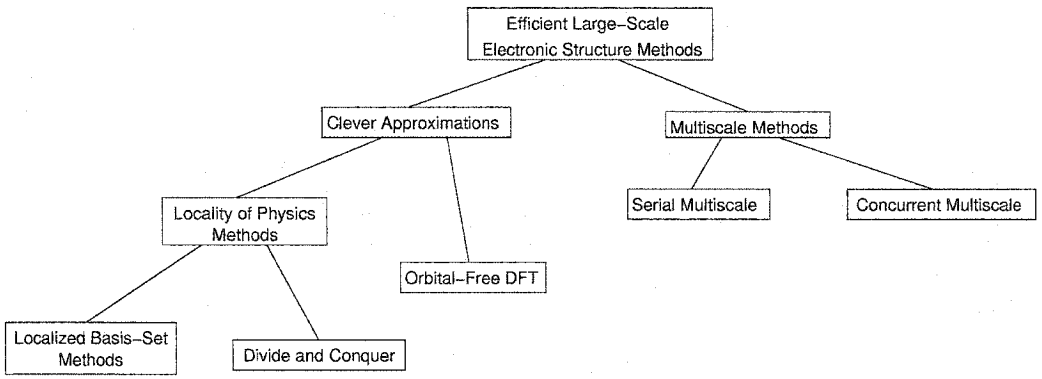

Figure 1.1: A hierarchical view of efficient methods for large-scale electronic structure calculations.

hind them. One idea that has resulted in the development of many similar methods[10, 11, 12, 13, 14, 15] in different guises is that of electronic "locality of physics"[14]- the property possessed by some systems in which the properties of electrons at one point in space are uncorrelated with properties of electrons at another point when the two points are sufficiently far away. Unlike the phenomenon of Anderson localization, the present form of localization occurs even in perfectly-ordered systems. Efficient methods have been developed to take advantage of this localization property. They work typically by dividing the system into overlapping regions of a size comparable to that of the localization length, and solving the electronic problem in each region somewhat independently, i.e. a "divide and conquer" approach. These approaches, while they are much more efficient than the Kohn-Sham method in the limit of very large system size, have a "crossover" size below which they are no faster than the Kohn-Sham method, and often this crossover size itself is at the limit of the current computational capability.

Another approach, one that has received less attention, is quite different. While the Kohn-Sham method results from mapping the interacting electron problem onto

a density problem, and then further onto a non-interacting electron problem, another approach is to consider the intermediate density problem itself. While Kohn himself developed the Kohn-Sham method to overcome the inadequacies of the methods available at the time for treating the pure density problem, there has been significant progress since then in pure density functional methods. And many of these methods are more efficient than the Kohn-Sham method; in fact they are potentially much more efficient, and without a crossover size. It is this class of methods, the orbital-free density functional methods, that is studied presently.

1.1.2 Multiscale methods

The second general category, that of the multiscale approaches, attacks the problem of large systems from a different philosophical standpoint. In contrast to the types of methods discussed in the previous section, the multiscale approaches do not necessarily involve any new algorithms *per se*, but seek to apply existing methods in a less wasteful manner.

Two distinct types of schemes are currently being termed "multiscale". The first type, which will be referred to here as *serial multiscale* methods, are methods that derive properties of a system from an accurate, small-scale calculation, and pass these properties to a larger-scale scale theory for which these parameters are givens. For instance, one might compute the elastic moduli of a material from DFT calculation of a single crystal unit cell, and then use these moduli in a continuum simulation of a macroscopic piece of that material. The problem with the serial multiscale category is that if taken too literally, it encompasses almost all of physics; almost every theory

takes as given some result from a lower-level theory.

A distinct category of techniques are the *parallel* or *concurrent* multiscale methods. In a concurrent multiscale method, two (or more) different types of simulation are combined to treat a single system. The concurrent multiscale approach is applicable to systems in which the detailed physics is confined to a small region of the whole system, but nonetheless depends on the less-detailed physics outside of the small region. The approach then is to apply an accurate (but undoubtedly computationally expensive) method to the detailed region of the system, and treat the rest of the system with a computationally cheaper (but undoubtedly less accurate) method. An example system (which has become a classic of the field) where a concurrent multiscale approach makes sense is a material containing a crack. While the crack propagation itself is a detailed atomistic phenomenon, the stimulus of the propagation can be macroscopic. Applied strains on the macroscopic scale are carried down to the tip of the crack, where the stress becomes concentrated, causing the bonds to rupture and the crack to travel.

The development of a new concurrent multiscale simulation need not involve the development of a new type of simulation. The novelty that is needed is a means of seamlessly coupling different simulation types. This is a challenging problem generally, made even more challenging by the fact that the problem is somewhat nebulous- there is no consensus as to what constitutes a "seamless" coupling.

1.2 An overview of the present work

The developments of the present thesis are in both the area of approximate DFT and multiscale simulation; in fact, the two are brought together in the multiscale method presently constructed. A theme that pervades all topics presented here is orbital-free density-functional theory (OF-DFT). It is useful both as an efficient density-functional method, and, it will be shown, as a means of coupling quantum mechanics to classical mechanics. The theory of OF-DFT and a description of a method for performing simulations with OF-DFT is presented in chapter 2.

A feature of OF-DFT methods is that, because there are no orbitals involved, non-local pseudopotentials do not apply. A study of local pseudopotentials was undertaken, and is presented in chapter 3. A set of methods were developed for the development of local pseudopotentials that attempt to capture the essential notions that have made modern norm-conserving pseudopotentials so useful. These methods were used to investigate the realm of local pseudopotentials, and interesting conclusions are drawn regarding the common notions behind norm-conserving pseudopotentials.

With an eye towards multiscale application, a method was developed for performing OF-DFT calculations fully in *real space*. This is rather non-trivial, as OF-DFT methods are typically derived in reciprocal space and the direct translations of them into real space can be very inefficient. Our method allows for the simulation of clusters and other finite systems with a wide variety of orbital-free functionals. This method is described in chapter 4.

Efficient electronic structure methods often focus on the calculation of the total energy of the system, and sometimes the calculation of the forces within these

formalisms is not efficient. We have developed a method explicitly for the efficient computation of forces. This method is applicable to any electronic structure method that utilizes a uniform grid in space. The method is presented in chapter 5.

We have developed a class of multiscale methods for coupling quantum and classical simulation for simple metallic systems. The coupling has a density-functional foundation. It inherently requires a mapping from a charge density to an energy, which the Kohn-Sham method cannot provide, but which the orbital-free DFT methods can. Thus OF-DFT is a necessary ingredient in the present multiscale method. The formalism and analysis of the coupling methods are presented in chapter 6. The coupling methods are applied to several physical systems in chapter 7. Finally in chapter 8 we summarize, speculate, and conclude.

Chapter 2

Orbital-free density functional theory

A recurring theme of this thesis is orbital-free density functional theory (OF-DFT). It is our aim to give a general description of OF-DFT in this chapter. But before we do this, we first describe the many-body quantum problem we are ultimately trying to solve. We then briefly describe the density functional approach to solving it. This allows us to present the orbital-free density functional theory, a particular manifestation of DFT which, in contrast to more traditional DFT methods, does not involve any fictitious single-particle wavefunctions. We emphasize the difference between OF-DFT and the standard Kohn-Sham approach. Finally, we provide some details of the actual implementation of OF-DFT for practical calculations of condensed matter systems.

2.1 The Schrödinger equation

For much of chemistry and condensed matter physics, a complete description of a system is provided by its energy and electronic wavefunctions in its ground and excited states. Very often knowledge of the ground state suffices. The ground-state energy of a system of electrons and nuclei is found by solving the Schrödinger equation, written here in variational form:

$$\begin{aligned} E &= \min_{\Psi} \langle \Psi | \hat{H} | \Psi \rangle \\ \hat{H} &\equiv -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} - \frac{1}{2m_p}\sum_i \nabla^2_{\mathbf{R}_i} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i<j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} - \sum_{i,j} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \end{aligned} \tag{2.1}$$

where the minimization is over all normalized $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\})$ that are antisymmetric in the variables $\{\mathbf{r}_i\}$ describing electron coordinates. The $\{\mathbf{R}_j\}$ describe the nuclear coordinates, and m_p is the proton mass. Also note that in Eq. (2.1) and throughout, *atomic units*, in which $\hbar = e = m_e = 1$, are employed.

Eq. (2.1) is one of the most important equations in all of science, as its solutions describe all of chemistry, as famously noted by Dirac. Indeed, almost all characteristics of matter at "reasonable" temperatures and pressures, including all of the exotic low-temperature properties such as superconductivity, the quantum Hall effects, etc., can be shown to be consequences of Eq. (2.1). Just as the few rules of chess result in a game of great complexity and depth, the simplicity of Eq. (2.1) belies the incredibly diverse world of its solutions.

Given its relevance, Eq. (2.1) has been the subject of constant study since it was written down. Many methods have been developed for its approximate solution,

and many of these are relevant only to some particular physical situation. Progress has typically been made by making an approximation to Eq. (2.1) or its solutions that enable some sort of analytical attack. We now describe one such important approximation, the Born-Oppenheimer approximation, which, in a sense, cuts the problem in half.

2.1.1 The Born-Oppenheimer approximation

A key observation regarding Eq. (2.1) is that $m_p >> 1$. It can be shown[16] that this allows for a very accurate solution with a separable wavefunction:

$$\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_j\}) \rightarrow \Psi_{\mathrm{e}}(\{\mathbf{r}_i\})\,\Psi_N(\{\mathbf{R}_j\}) \tag{2.2}$$

The ground-state electronic wavefunction Ψ_{e} can be shown to minimize, to a very good approximation:

$$\begin{aligned} E_{\mathrm{e}}[\mathbf{R}_j] &= \min_{\Psi_{\mathrm{e}}} \left\langle \Psi_{\mathrm{e}} \left| -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + V_{\mathrm{e-i}}(\mathbf{r}_i) + V_{\mathrm{e-e}}(\mathbf{r}_i) \right| \Psi_{\mathrm{e}} \right\rangle, \\ V_{\mathrm{e-i}}(\mathbf{r}_i) &\equiv \sum_{i,j} \frac{-Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}, \\ V_{\mathrm{e-e}}(\mathbf{r}_i) &\equiv \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \end{aligned} \tag{2.3}$$

where now the minimization is over normalized electronic wavefunctions $\Psi_{\mathrm{e}}(\{\mathbf{r}_i\})$ that are antisymmetric in the $\{\mathbf{r}_i\}$. This is the essence of the Born-Oppenheimer approximation. We may treat the nuclear coordinates simply as parameters of the potential, and the electrons, to a high degree of accuracy, live in the ground state of the potential $V_{\mathrm{e-i}}$ of the nuclei at given positions $\{\mathbf{R}_j\}$.

The Born-Oppenheimer approximation retains the important physics of most con-

densed matter systems. Although the Born-Oppenheimer approximation greatly simplifies the problem, Eq. (2.3) is still incredibly complicated.

2.1.2 The Hartree energy

The complexity of the solutions of Eq. (2.3) are due in large part to the electron-electron interaction $V_{\mathrm{e-e}}$. The electron-electron energy of a many-body electron wavefunction Ψ can be approximately calculated by evaluating the *Hartree* energy E_{H}:

$$\begin{aligned} \langle\Psi|V_{\mathrm{e-e}}|\Psi\rangle &\simeq E_{\mathrm{H}}[\rho_{\Psi}(\mathbf{r})], \\ \rho_{\Psi}(\mathbf{r}) &\equiv \int |\Psi(\mathbf{r},\mathbf{r}_2,\ldots,\mathbf{r}_n)|^2 \mathrm{d}\mathbf{r}_2\cdots\mathrm{d}\mathbf{r}_n, \\ E_{\mathrm{H}}[\rho(\mathbf{r})] &\equiv \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \end{aligned} \tag{2.4}$$

Note that the Hartree energy is a density functional, mapping a charge density to a real number. It is simply the classical electrostatic energy of the charge density. The Hartree energy is often represented in terms of the Hartree potential $V_{\mathrm{H}}(\mathbf{r})$:

$$E_{\mathrm{H}}[\rho(\mathbf{r})] = \frac{1}{2}\int \rho(\mathbf{r})V_{\mathrm{H}}(\mathbf{r})d\mathbf{r} \tag{2.5}$$

where

$$V_{\mathrm{H}}(\mathbf{r}) \equiv \frac{\delta E_{\mathrm{H}}}{\delta\rho(\mathbf{r})} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' \tag{2.6}$$

The Hartree energy functional can lead to a method in and of itself, the Hartree theory, but we will not discuss it here. The Hartree energy serves as a good example of a density functional, and it plays an important role in both the Kohn-Sham method and orbital-free DFT methods.

2.2 Density functional theory

Density functional theory (DFT) is a family of methods for approximately solving Eq. (2.3) that have been very successful in yielding, to a good approximation, the chemical and physical properties of a great many materials and molecules. These methods ultimately resulted in the awarding of the 1998 Nobel Prize in chemistry to Walter Kohn, one of the founding fathers of DFT.

2.2.1 Thomas-Fermi theory

The first glimmers of density functional theory came as early as 1927, merely one year after the Schrödinger equation was written down. Thomas[17] and Fermi[18] developed a statistical theory of atoms based on the approximation that the electrons distribute themselves uniformly in phase space. From the modern density functional theory point of view, the Thomas-Fermi theory is based on using a local density functional to treat the kinetic energy of a system. The Thomas-Fermi functional can be derived as follows. Assume that N non-interacting Fermions are placed in a cubic box of edge length L. If periodic boundary conditions are imposed, the electrons assume plane-wave states, with wave-vectors determined by the boundary conditions:

$$\begin{aligned} \psi_k(\mathbf{r}) &= L^{-3/2} e^{i\mathbf{k}\cdot\mathbf{r}}, \\ \mathbf{k} &= \frac{2\pi}{L}(n_1, n_2, n_3), \text{ where } n_i \in \mathbb{Z} \end{aligned} \tag{2.7}$$

The kinetic energy of $\psi_k(\mathbf{r})$ is given by $k^2/2$. Thus the kinetic energy of the N lowest-energy particles (keeping in mind spin degeneracy) is thus:

$$T_{\mathrm{TF}} = \int_{|\mathbf{k}|<k_F} 2\frac{k^2}{2}\left(\frac{L}{2\pi}\right)^3 d\mathbf{k} = \frac{L^3 k_F^5}{10\pi^2} \tag{2.8}$$

where k_F, the Fermi wave-vector, is determined by requiring that the number of particle states of smaller wave-vector (keeping in mind spin degeneracy) is given by N:

$$N = \int_{|\mathbf{k}|<k_F} 2\left(\frac{L}{2\pi}\right)^3 d\mathbf{k} = \frac{L^3 k_F^3}{3\pi^2} \tag{2.9}$$

We can thus express the kinetic energy in terms of the density:

$$\begin{aligned} T_{\mathrm{TF}} &= V C_{\mathrm{TF}} \rho^{5/3}, \\ C_{\mathrm{TF}} &\equiv \frac{3}{10}(3\pi^2)^{2/3}, \\ \rho &\equiv N/L^3 \end{aligned} \tag{2.10}$$

The Thomas-Fermi functional is an approximation for the kinetic energy of an arbitrary, spatially inhomogeneous charge density $\rho(\mathbf{r})$, in which the kinetic energy of every infinitesimal volume of space is assumed to be that of a uniform charge density:

$$T_{\mathrm{TF}}[\rho(\mathbf{r})] \equiv C_{\mathrm{TF}} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} \tag{2.11}$$

The Thomas-Fermi functional obviously gives exactly the kinetic energy of a uniform charge density $\rho(\mathbf{r}) = \rho_0$. However, without the formalism of density functional theory, it is not even clear what the Thomas-Fermi functional yields for an arbitrary $\rho(\mathbf{r})$, as there are many different sets of orthonormal wavefunctions that yield a density $\rho(\mathbf{r})$, yet have different kinetic energies. This issue is completely cleared up by the theory of Hohenberg and Kohn.

2.2.2 The Hohenberg-Kohn theorem

The quantum-mechanical properties of a system can be examined by using basic degrees of freedom other than the wavefunction. Theories have been built around

Green functions and reduced density matrices. Hohenberg and Kohn[5] made rigorous the use of the density as the basic variable. Since their original work, the arguments have been simplified by Levy[19] and Lieb[20] via the constrained-search formalism, which we reproduce here.

First we note that we can re-express the ground-state electronic energy, Eq. (2.3), as:

$$\begin{aligned} E_{\mathrm{e}}[\mathbf{R}_j] &= \min_{\Psi_{\mathrm{e}}} \left[\left\langle \Psi_{\mathrm{e}} \left| \hat{T} + V_{\mathrm{e-e}} \right| \Psi_{\mathrm{e}} \right\rangle + \int \rho_{\Psi}(\mathbf{r}) V_{\mathrm{e-i}}(\mathbf{r}) \mathrm{d}\mathbf{r} \right], \\ \hat{T} &\equiv -\frac{1}{2} \sum_i \nabla^2_{\mathbf{r}_i} \end{aligned} \tag{2.12}$$

We then define a density functional $F[\rho]$:

$$F[\rho] \equiv \min_{\Psi} \left\langle \Psi \left| \hat{T} + V_{\mathrm{e-e}} \right| \Psi \right\rangle \text{ s.t. } \rho_{\Psi} = \rho \tag{2.13}$$

i.e. we perform a *constrained search* over all normalized, antisymmetric wavefunctions that yield a density ρ, and pick the one with the smallest combined kinetic and electron-electron energy. We also note that the definition of $F[\rho]$ establishes a mapping from any density to a particular wavefunction:

$$\rho \rightarrow \Psi_{\rho} \tag{2.14}$$

the density of which is ρ, and which minimizes the kinetic plus electron-electron energy. Thus we can express $F[\rho]$ as:

$$F[\rho] = \langle \Psi_{\rho} | \hat{T} | \Psi_{\rho} \rangle + \langle \Psi_{\rho} | V_{\mathrm{e-e}} | \Psi_{\rho} \rangle \tag{2.15}$$

With this functional in hand, the ground-state electronic energy of the system is:

$$E_{\mathrm{e}}[\mathbf{R}_j] = \min_{\rho} \left[F[\rho] + \int \rho(\mathbf{r}) V_{\mathrm{e-i}}(\mathbf{r}) \mathrm{d}\mathbf{r} \right] \text{ s.t. } \int \rho(\mathbf{r}) \mathrm{d}\mathbf{r} = N \tag{2.16}$$

Thus the density ρ can serve as the basic variable of quantum mechanics. Unfortunately, the full complexity of the many-body problem is wrapped up in $F[\rho]$, and without further information about this functional, we are no closer to being able to calculate properties of quantum-mechanical systems.

2.2.3 The Kohn-Sham method

We can gain insight into the nature of $F[\rho]$ in several ways. First, as noted previously, the Hartree energy $E_{\mathrm{H}}[\rho]$ provides an approximation to $\langle\Psi_\rho|V_{\mathrm{e-e}}|\Psi_\rho\rangle$. Furthermore, we note that while the particular details of $F[\rho]$ depend on the form of the electron-electron interaction, our *ability* to define $F[\rho]$ did not. Thus for any form of particle-particle interaction we can construct a corresponding density functional $F[\rho]$. In particular, if there is zero particle-particle interaction, we can define a functional known as the non-interacting kinetic energy functional:

$$T_{\mathrm{s}}[\rho] \equiv \min_{\Psi}\langle\Psi|\hat{T}|\Psi\rangle \text{ s.t. } \rho_\Psi = \rho \tag{2.17}$$

We digress a moment to reveal some properties of T_{s}. Because of the lack of any particle-particle interaction, Eq. (2.17) is in general minimized with a single Slater determinant $\Phi_\rho = \mathrm{Sl}[\{\phi_i\}]$ of orthonormal single-particle orbitals $\{\phi_i\}$. Thus we can re-express T_{s} in terms of single-particle orbitals:

$$\begin{aligned} T_{\mathrm{s}}[\rho] &= \min_{\phi_i}\sum_i\langle\phi_i| - \tfrac{1}{2}\nabla^2|\phi_i\rangle \\ \text{s.t. } \langle\phi_i|\phi_j\rangle &= \delta_{ij} \text{ and } \sum_i|\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \end{aligned} \tag{2.18}$$

The non-interacting kinetic energy is useful because the *interacting* kinetic energy $\langle\Psi_\rho|\hat{T}|\Psi_\rho\rangle$ should be reasonably approximated by $T_{\mathrm{s}}[\rho]$. This observation led Kohn

and Sham to the following expression for the electronic energy of a system:

$$E_{\rm e} = \min_{\rho} \left[T_{\rm s}[\rho] + E_{\rm H}[\rho] + E_{\rm xc}[\rho] + \int \rho(\mathbf{r}) V_{\rm e-i}(\mathbf{r}) \mathrm{d}\mathbf{r} \right] \quad (2.19)$$

$$\text{s.t.} \int \rho(\mathbf{r}) \mathrm{d}\mathbf{r} = N,$$

$$E_{\rm xc} \equiv F[\rho] - T_{\rm s}[\rho] - E_{\rm H}[\rho] \quad (2.20)$$

The *exchange-correlation* functional $E_{\rm xc}[\rho]$ is thus defined as the small part of $F[\rho]$ not accounted for by the non-interacting kinetic energy and the Hartree energy. Now we use our single-particle orbital description of $T_{\rm s}$, Eq. (2.18), within Eq. (2.19). We also express $\rho(\mathbf{r})$ in terms of single-particle orbitals in the last term of Eq. (2.19); the result is:

$$E_{\rm e} = \min_{\phi_i} \left[\sum_i \langle \phi_i | -\tfrac{1}{2}\nabla^2 | \phi_i \rangle + E_{\rm H}[\rho] + E_{\rm xc}[\rho] + \sum_i \int |\phi_i(\mathbf{r})|^2 V_{\rm e-i}(\mathbf{r}) \mathrm{d}\mathbf{r} \right] \quad (2.21)$$

$$\text{such that } \langle \phi_i | \phi_j \rangle = \delta_{ij} \quad (2.22)$$

$$\text{and } \rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2 \quad (2.23)$$

If we impose the orthonormality constraints, Eq. (2.22), with a set of Lagrange multipliers ϵ_{ij}, and we take the variation of Eq. (2.21) with respect to the ϕ_i, the result is:

$$\left[-\tfrac{1}{2}\nabla^2 + V_{\rm H}[\rho;\mathbf{r}] + V_{\rm xc}[\rho;\mathbf{r}] + V_{\rm e-i}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \sum_j \epsilon_{ij} \phi_j(\mathbf{r})$$

$$\text{s.t. } \rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2, \quad (2.24)$$

$$\text{where } V_{\rm H}[\rho;\mathbf{r}] \equiv \frac{\delta E_{\rm H}}{\delta \rho(\mathbf{r})} = \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}' \quad (2.25)$$

$$\text{and } V_{\rm xc}[\rho;\mathbf{r}] \equiv \frac{\delta E_{\rm xc}}{\delta \rho(\mathbf{r})} \quad (2.26)$$

We can perform a unitary transformation of the ϕ_i in order to diagonalize ϵ_{ij}:

$$\begin{aligned} \phi_i &= \sum_j U_{ij}\tilde{\phi}_j, \\ \sum_{k,l} U^*_{ki}\epsilon_{kl}U_{lj} &= \delta_{ij}\tilde{\epsilon}_i \end{aligned} \tag{2.27}$$

In terms of the $\tilde{\phi}_i$, Eqs. (2.24) reduce to:

$$\left[-\tfrac{1}{2}\nabla^2 + V_{\rm H}[\rho;\mathbf{r}] + V_{\rm xc}[\rho;\mathbf{r}] + V_{\rm e-i}(\mathbf{r}) - \tilde{\epsilon}_i\right]\tilde{\phi}_i(\mathbf{r}) = 0 \tag{2.28}$$

$$\text{s.t. } \rho(\mathbf{r}) = \sum_i |\tilde{\phi}_i(\mathbf{r})|^2 \tag{2.29}$$

Eqs. (2.28) are known as the Kohn-Sham equations, and the $\tilde{\phi}_i$ as the Kohn-Sham orbitals. The Kohn-Sham equations provide a formally exact method for treating the interacting electron problem. Again, the complexities of many-body phenomena are now thrown onto the exchange-correlation functional, $E_{\rm xc}[\rho]$. But definite progress has been made, because the uncertain part of the problem, $E_{\rm xc}$, is a small quantity.

The Kohn-Sham method proceeds in practice as follows. An approximate form is chosen for $E_{\rm xc}[\rho]$. An initial guess is made for the charge density, $\rho^{(0)}$. Then the potentials $V_{\rm H}[\rho^{(0)}]$ and $V_{\rm xc}[\rho^{(0)}]$ are constructed. Then Eqs. (2.28) are solved for the ϕ_i; these are simply a set of single-particle Schödinger equations for which there are many robust numerical solution methods. From these ϕ_i, a new density $\rho^{(1)}$ is constructed. Then the process repeats itself until self-consistency is achieved, i.e. until $\rho^{(i+1)}$ and $\rho^{(i)}$ are not much different.

The Kohn-Sham method can be seen as a mapping from an interacting-electron problem to a non-interacting electron problem with the same charge density. The kinetic energy of the many-body wavefunction Ψ_ρ is similar to that of the non-

interacting wavefunction Φ_ρ, the electron-electron energy of Ψ_ρ is close to $E_{\rm H}[\rho]$, and the errors are formally corrected by $E_{\rm xc}$.

2.2.4 The exchange-correlation functional and the LDA

The simplest and most frequently used approximate form for the exchange-correlation functional is the *local density approximation*, or LDA. The essence of the LDA is that the contribution that $\rho(\mathbf{r})$ within any infinitesimal volume $\mathrm{d}\mathbf{r}$ makes to $E_{\rm xc}[\rho(\mathbf{r})]$ is equal to the exchange-correlation energy of a uniform electron gas of density $\rho_0 = \rho(\mathbf{r})$ and volume $\mathrm{d}\mathbf{r}$:

$$E_{\rm xc}^{\rm LDA}[\rho(\mathbf{r})] = \int f_{\rm xc}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \tag{2.30}$$

where $f_{\rm xc}(\rho)$ is the exchange-correlation energy density of a uniform electron gas of density ρ. Eq. (2.30) is typically expressed as:

$$E_{\rm xc}^{\rm LDA}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})\epsilon_{\rm xc}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \tag{2.31}$$

Where $\epsilon(\rho)$ is the exchange-correlation energy density per particle. We have already encountered an LDA approximation to a functional: the Thomas-Fermi functional, Eq. (2.11), which is the LDA approximation of the non-interacting kinetic energy functional $T_{\rm s}$. The energy of a uniform interacting electron gas (i.e. $F[\rho_0]$) has been calculated to a fair degree of accuracy with quantum Monte Carlo techniques in an important calculation by Ceperley and Alder[21]. By subtracting off the non-interacting-kinetic and Hartree energies, the result, via Eq. (2.20), is the exchange-correlation energy.

The popularity of density functional theory is due largely to the surprising success of the LDA for the exchange-correlation energy. This success is somewhat understood via the adiabatic connection formalism of DFT [22, 23, 24]: the LDA exchange-correlation functional is found to satisfy a particular sum rule that the true functional satisfies. The smallness of the exchange-correlation energy also helps reduce the impact of any errors that an approximate XC functional may have.

2.2.5 Pseudopotentials

An integral part of the "standard model" of electronic structure calculations is the pseudopotential, which immensely reduces the amount of computation required to simulate many systems. The pseudopotential concept stems from the notion of core and valence electrons. When a single-particle method such as the Kohn-Sham method is used to calculate the electronic structure of an atom, one typically finds that some of the single-particle states are characterized by highly negative energies and strong localization near the nucleus, while the remaining states extend further from the nucleus and have energies much closer to zero. These are respectively the *core* and *valence* states. The valence states are delocalized from the nucleus because they must be orthogonal to the core states, and they do so by occupying a region of space other than the core. This results in their energy being much higher.

When atoms are brought together to form molecules or solids, the valence electrons overlap and mix to form extended states. But typically the core states of each atom are affected only very slightly. The properties of the bonded system are almost entirely determined by the valence states, and the core states serve merely to form an

"orthogonality hole" around the nucleus that the valence electrons must avoid, and to partially shield the (Coulombic) electrostatic potential of the nucleus.

This leads to the notion of the *pseudopotential*, a potential that is used to represent the combination of the nucleus and the core electrons. The tail form of the pseudopotential should behave like the screened nuclear Coulombic tail, and thus should have the form:

$$V_{\mathrm{psp}}(r) \sim -\frac{N_{\mathrm{n}} - N_{\mathrm{c}}}{r} \text{ as } r \to \infty \tag{2.32}$$

where N_{n} is the nuclear charge, and N_{c} is the number of core electrons. The core of the pseudopotential should have the right amount of repulsion to mimic the orthogonality hole of the core electrons. Beyond that, the exact form of the pseudopotential does not matter that much. The bonding properties of a system are not much affected by the exact shape of the valence states within the core region. This freedom in the core region is exploited to construct pseudopotentials that result in pseudo-valence states that match the true valence states outside the core, but are optimally smooth inside the core. This allows for an easier numerical representation of the pseudo-valence orbitals; a basis-function expansion of the resulting pseudo-orbitals requires fewer functions, and spatial grid resresentations can get away with coarser grids. Thus pseudopotentials result in a large reduction in computational effort: not only are the core electrons removed from the problem entirely, but the number of degrees of freedom needed to represent the remaining valence states can be made optimally small.

2.2.6 Kinetic energy functionals and orbital-free DFT

The Kohn-Sham method, with all of its strengths, still leaves room for improvement in terms of numerical efficiency, even when combined with the pseudopotential concept. The problem lies with the solution of the Kohn-Sham equations, Eqs. (2.28). They can be solved by exact diagonalization of the Hamiltonian matrix; the Hamiltonian matrix is an $N \times N$ matrix, where N is proportional to the number of valence electrons in the system (and hence the number of atoms in the system). Exact diagonalization requires a computational effort proportional to N^3. A better method is to iteratively solve Eqs. (2.28), which takes advantage of the sparseness of the Hamiltonian. But orthogonality of the Kohn-Sham orbitals must be maintained, which requires computation of the overlap matrix $S_{ij} \equiv \langle \phi_i | \phi_j \rangle$. There are N^2 such matrix elements, and the calculation of each one requires a computation proportional to N. Thus for iterative solution methods the overall computation scales as N^3 as well.

A computational effort that scales as N^3 mounts quickly with growing system size. The best algorithmic scaling one could possibly have is a linear dependence on the system size, as the amount of memory required to store a description of the system scales as N, and the time to access all that memory scales as N.

One route to achieving a linear-scaling algorithm is through wavefunction localization. If the wavefunctions are each confined to a fixed volume, the overlap matrix becomes sparse, and the N^3 orthogonalization bottleneck can be overcome.

Another route to achieving algorithmic efficiency comes from approaching the density functional problem from a different angle. The Kohn-Sham method results from mapping the many-body problem, Eq. (2.3), to a density problem, Eq. (2.19),

and further mapping this to a non-interacting Fermion problem, Eq. (2.28). But one may choose to attack the density problem itself.

The idea of orbital-free density functional theory (OF-DFT) is to calculate the energy by performing the minimization of Eq. (2.19) directly in terms of the density. This requires one to have an explicit functional form for $T_s[\rho]$ and $E_{xc}[\rho]$. We can again use the LDA XC functional, but this leaves the problem of approximating T_s.

The issue of finding a good approximate T_s is one of the central problems of OF-DFT. It is at once more and less difficult than the exchange-correlation functional problem. On one hand, the kinetic energy is a much larger contribution to the total energy than the exchange-correlation energy. On the other hand, T_s is inherently much simpler to analyze than E_{xc}, since it does not involve electron-electron interaction and all of the complexities associated with it. As such, a fair amount has been learned about the structure and limiting forms of the exact functional T_s.

A general strategy that has been employed for the construction of many approximate (non-interacting) kinetic energy functionals is to construct an explicit density functional that exhibits the correct behavior in one or more of the known limits of the exact functional. The primary functionals used in the present thesis are those of Wang et al.[25, 26], which exhibit properties of the exact functional that are relevant to simple metallic systems. They are modern functionals (as of this thesis) that represent a good balance between quality and computational efficiency.

Perhaps the most important limit that any approximate kinetic energy functional must exhibit is the LDA limit of T_s, namely T_{TF}. That is, for uniform charge densities $\rho(\mathbf{r}) = \rho_0$, the approximate functional must equal $T_{TF}[\rho_0]$.

The functionals employed presently take several other limits of the exact functional into account. Another limit that is easily accounted for is the case when $\rho(\mathbf{r})$ consists of a single orbital, i.e. one or two electrons. In this case we have $\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$, and the kinetic energy is given by the *von Weizsäcker* functional[27]:

$$T_{\mathrm{vW}}[\rho] = -\frac{1}{2}\int \rho^{1/2}(\mathbf{r})\nabla^2\rho^{1/2}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{2.33}$$

This functional also captures the kinetic energy properly in several other situations. When the density has an oscillation of a wave-vector $k > 2k_{\mathrm{F}}$, then this can only result from an orbital of wave-vector $k/2$ being present; T_{vW} properly captures its kinetic energy. Also at surfaces, vacancies or anywhere there the density has an expontential tail, the orbital with highest energy (and hence the slowest expontential decay) dominates the density and the kinetic energy, and again the energy is properly treated by T_{vW}.

Another known limit, the linear-response regime, is important, especially for metallic systems. The linear response of T_{s} is its second functional derivative with respect to the density, evaluated at a uniform charge density ρ_0:

$$\left.\frac{\delta^2 T_{\mathrm{s}}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})}\right|_{\rho_0} \tag{2.34}$$

We can derive the exact linear response as follows. First we note that the density-functional form of the non-interacting Fermion problem is given by:

$$\min_{\rho}\left[T_{\mathrm{s}}[\rho] + \int \rho(\mathbf{r})V(\mathbf{r})\mathrm{d}\mathbf{r}\right] \tag{2.35}$$

or equivalently:

$$\frac{\delta T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r})} = -V(\mathbf{r}) \tag{2.36}$$

thus we have:

$$\frac{\delta^2 T_s}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})} = -\frac{\delta V(\mathbf{r})}{\delta\rho(\mathbf{r}')} \tag{2.37}$$

Thus if we need to relate the change in the potential at a point $\mathbf{r}$ when the density is changed at point $\mathbf{r}'$. We can find this relation via perturbation theory. If we perturb our box of plane waves used in deriving T_{TF} (see Eqs. (2.7)-(2.10)) with a potential of the form:

$$V(\mathbf{r}) = \sum_{\mathbf{q}} \delta V_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \tag{2.38}$$

perturbation theory yields the following perturbed wavefunctions:

$$\psi_{\mathbf{k}} = L^{-3/2} e^{i\mathbf{k}\cdot\mathbf{r}} + 2L^{-3/2} \sum_{\mathbf{k}'\neq\mathbf{k}} \frac{\delta V_{\mathbf{k}'-\mathbf{k}}}{k^2 - k'^2} e^{i\mathbf{k}'\cdot\mathbf{r}} \tag{2.39}$$

By summing up the densities of each of the occupied perturbed wavefunctions, we find:

$$\begin{aligned} \rho(\mathbf{r}) &= \rho_0 + \delta\rho(\mathbf{r}), \\ \delta\rho(\mathbf{r}) &= \frac{8}{L^3} \sum_{|\mathbf{k}|<k_{\mathrm{F}}} \sum_{\mathbf{k}'\neq\mathbf{k}} \frac{\delta V_{\mathbf{k}'-\mathbf{k}}}{k^2 - k'^2} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} \\ &= \sum_{\mathbf{k}} \chi_{\mathrm{Lind}}(k)\, \delta V_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, \end{aligned} \tag{2.40}$$

where $\chi_{\mathrm{Lind}}(x)$ is the Lindhard response function[28], given by:

$$\begin{aligned} \chi_{\mathrm{Lind}}(k) &= -\frac{1}{\pi^3} \int_{|\mathbf{k}|<k_{\mathrm{F}}} \frac{\mathrm{d}\mathbf{k}}{q^2 + 2\mathbf{k}\cdot\mathbf{q}} \\ &= -\frac{k_{\mathrm{F}}}{\pi^2} F_{\mathrm{Lind}}\left[\frac{k}{2k_{\mathrm{F}}}\right], \\ F_{\mathrm{Lind}}(x) &\equiv \frac{1}{2} + \frac{1-x^2}{4x} \ln\left|\frac{1+x}{1-x}\right| \end{aligned} \tag{2.41}$$

Thus we can relate the change in potential to the change in density:

$$\delta V_{\mathbf{k}} = \frac{1}{\chi_{Lind}(k)} \delta\rho_{\mathbf{k}} \tag{2.42}$$

Then from Eq. (2.37) we have:

$$\hat{\mathcal{F}}\left[\frac{\delta^2 T_{\mathrm{s}}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})}\right] = -\frac{1}{\chi_{Lind}(k)} \tag{2.43}$$

where $\hat{\mathcal{F}}[f(\mathbf{r},\mathbf{r}')] \equiv \int f(\mathbf{r},\mathbf{r}')e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\mathrm{d}(\mathbf{r}-\mathbf{r}')$ is the Fourier transform. Eq. (2.43) is the exact linear response of the homogeneous, non-interacting Fermi gas about any uniform density. The value of the uniform density determines k_{F}.

2.2.7 Forces within OF-DFT

The Hellmann-Feynman theorem is used in quantum mechanics to find the atomic forces within the Born-Oppenheimer approximation. However, the derivation of the Hellmann-Feynman theorem involves wavefunctions, and it is worthwhile to demonstrate how forces can be calculated within the purely density-based framework of DFT.

The total energy of a combined electronic-ionic system (within the Born-Oppenheimer approximation) is given by the sum of the electronic energy and the ion-ion or Ewald energy:

$$\begin{aligned} E_{\mathrm{tot}}[\{\mathbf{R}_j\}] &= \min_{\rho} E'_{\mathrm{elec}}[\rho;\{\mathbf{R}_j\}] + E_{\mathrm{i-i}}[\{\mathbf{R}_j\}], \\ E'_{\mathrm{elec}} &\equiv \min_{\rho}\left[F[\rho] + \int \rho(\mathbf{r})V_{\mathrm{e-i}}(\mathbf{r};\{\mathbf{R}_j\})\mathrm{d}\mathbf{r}\right], \\ E_{\mathrm{i-i}}[\{\mathbf{R}_j\}] &\equiv \sum_{i<j}\frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \end{aligned} \tag{2.44}$$

where the electronic energy E' depends on the ionic coordinates because $V_{\mathrm{e-i}}(\mathbf{r})$ is a superposition of potentials centered at the nuclear coordinates, and the Z_i are the nuclear charges (or valence charges, in the case of a pseudopotential calculation). The force on the ith atom is given by:

$$\begin{aligned}
\mathbf{F}_i &\equiv -\frac{\partial E}{\partial \mathbf{R}_i} && (2.45)\\
&= -\frac{\partial E'}{\partial \mathbf{R}_i} - \int \frac{\delta E'}{\delta \rho(\mathbf{r})} \frac{\partial \rho_{\min}(\mathbf{r})}{\partial \mathbf{R}_i} \mathrm{d}\mathbf{R} - \frac{\partial E_{\mathrm{i-i}}}{\partial \mathbf{R}_i} && (2.46)\\
&= -\frac{\partial E'}{\partial \mathbf{R}_i} - \frac{\partial E_{\mathrm{i-i}}}{\partial \mathbf{R}_i} && (2.47)\\
&= -\int \rho_{\min}(\mathbf{r}) \frac{\partial V_{\mathrm{e-i}}(\mathbf{r})}{\partial \mathbf{R}_i} \mathrm{d}\mathbf{r} - \frac{\partial E_{\mathrm{i-i}}}{\partial \mathbf{R}_i} && (2.48)
\end{aligned}$$

where the second term in Eq. (2.46) vanishes because E' is to be minimized with respect to ρ, and hence $\delta E'/\partial \rho(\mathbf{r}) = 0$.

This issue of how to compute the atomic forces *efficiently* remains, and is the subject of chapter 5.

2.3 The implementation of OF-DFT

Given approximate functionals for T_{s} and E_{xc}, an OF-DFT calculation proceeds by carrying out the minimization of Eq. (2.19). Many details lie between the conceptual level of Eq. (2.19) and an actual calculation.

2.3.1 Representation of the density

A major issue is the representation of the basic variable ρ. Practitioners of the Kohn-Sham method have used a large variety of representations of their basic variable,

the Kohn-Sham orbitals, including plane waves, Gaussian basis sets, and real-space grids. We have developed a number of different codes based on different representations of the density, but in our most efficient code a dual real-reciprocal representation is used to simulate periodic systems. The density is represented on a uniform grid within a periodic parallelepiped unit cell with basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$:

$$\rho(\mathbf{r}) \rightarrow \rho_{l_1 l_2 l_3} = \rho\left(\frac{l_1}{N_1}\mathbf{a}_1 + \frac{l_2}{N_2}\mathbf{a}_2 + \frac{l_3}{N_3}\mathbf{a}_3\right) \tag{2.49}$$

Local functionals such as the exchange-correlation energy within the LDA (Eq. 2.31) are computed with a real space sum:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})] \rightarrow \frac{\Omega}{N}\sum_{l_1 l_2 l_3} \rho_{l_1 l_2 l_3}\epsilon_{\mathrm{xc}}\left(\rho_{l_1 l_2 l_3}\right) \tag{2.50}$$

where $\Omega = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$ is the unit cell volume, and $N = N_1 N_2 N_3$.

But some of the needed quantities are more efficiently calculated in reciprocal space. Because the grid is uniform, the reciprocal space representation of any spatially-dependent quantity can be obtained efficiently by using the fast Fourier transform (FFT). If we define the reciprocal lattice vectors $\mathbf{b}_i$ such that $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$, then the Fourier transform of quantities defined on the unit cell is approximated by:

$$\begin{aligned}
\tilde{f}_{m_1 m_2 m_3} &\equiv \tilde{f}(m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3) && (2.51)\\
&\equiv \frac{1}{\Omega}\int_{\mathrm{cell}} f(\mathbf{r}) e^{i(m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3)\cdot\mathbf{r}} d\mathbf{r} && (2.52)\\
&\simeq \mathcal{F}[f_{l_1 l_2 l_3}] && (2.53)
\end{aligned}$$

where $\mathcal{F}$ is the discrete Fourier transform:

$$\begin{aligned}
\mathcal{F}[f_{l_1 l_2 l_3}] &= \frac{1}{N}\sum_{l_1=0}^{N_1-1}\sum_{l_2=0}^{N_2-1}\sum_{l_3=0}^{N_3-1} f_{l_1 l_2 l_3}\\
&\times\ e^{2\pi i\left(\frac{l_1 m_1}{N_1} + \frac{l_2 m_2}{N_2} + \frac{l_3 m_3}{N_3}\right)} && (2.54)
\end{aligned}$$

Its inverse, $\mathcal{F}^{-1}$, is given by:

$$\begin{aligned}\mathcal{F}^{-1}[\tilde{f}_{m_1m_2m_3}] &= \sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1}\sum_{m_3=0}^{N_3-1}\tilde{f}_{m_1m_2m_3} \\ &\times\ e^{-2\pi i(\frac{l_1m_1}{N_1}+\frac{l_2m_2}{N_2}+\frac{l_3m_3}{N_3})}\end{aligned} \tag{2.55}$$

$\mathcal{F}$ and $\mathcal{F}^{-1}$ can be computed in $O(N\log N)$ time with the fast Fourier transform. Thus convolutions, which become products in reciprocal space, can be efficiently computed:

$$(f\star g)(\mathbf{r}) \equiv \int f(\mathbf{r}')g(\mathbf{r}-\mathbf{r}')d\mathbf{r}' \tag{2.56}$$

$$= \mathcal{F}^{-1}[\tilde{f}\tilde{g}] = \mathcal{F}^{-1}\left[\mathcal{F}[f]\mathcal{F}[g]\right] \tag{2.57}$$

By this means the Hartree potential is efficiently computed by using the FFT:

$$(V_{\mathrm{H}})_{l_1l_2l_3} = \mathcal{F}^{-1}\left[\frac{\tilde{\rho}_{m_1m_2m_3}}{4\pi|m_1\mathbf{b}_1+m_2\mathbf{b}_2+m_3\mathbf{b}_3|^2}\right] \tag{2.58}$$

The FFT is also used to efficiently evaluate convolution terms in the approximate kinetic energy functionals.

The true non-interacting kinetic energy functional $T_{\mathrm{s}}[\rho]$ is not defined for negative ρ. However, most approximate forms of T_{s} can be analytically continued to negative ρ, and thus when carrying out the minimization of Eq. (2.19) over the domain of all real functions $\rho(\mathbf{r})$ using such an approximate functional, the possibility exists that the minimizing $\rho(\mathbf{r})$ could be negative in some regions of space. To eliminate this non-physical possibility, we instead work with the real quantity $\phi(\mathbf{r})$, and define:

$$\rho(\mathbf{r}) \equiv \phi^2(\mathbf{r}). \tag{2.59}$$

By minimizing with respect to real ϕ, the minimization domain is restricted to only positive densities.

2.3.2 Minimization of the energy

We minimize $E[\rho(\phi(\mathbf{r}))]$ with respect to $\phi(\mathbf{r})$ using a specialized form of the conjugate gradients algorithm. At each relaxation step, the gradient vector is calculated:

$$G(\mathbf{r}) = \frac{\delta E}{\delta \phi(\mathbf{r})} \tag{2.60}$$

then G is projected onto the surface of constant particle number:

$$G'(\mathbf{r}) = G(\mathbf{r}) - \frac{\langle \phi | G \rangle}{\langle \phi | \phi \rangle} \phi(\mathbf{r}) \tag{2.61}$$

In this manner, the number of particles in $\phi + \epsilon G'$ is the same as the number of particles in ϕ for small ϵ. If gradient preconditioning is to be performed, the following reciprocal-space preconditioner is used (this preconditioner is based on removing the high-frequency components of the gradient that come from the von Weizsäcker energy):

$$G''(\mathbf{q}) = \frac{1}{q^2 + 1} G'(\mathbf{q}) \tag{2.62}$$

Then we orthogonalize again:

$$G^{(3)}(\mathbf{r}) = G''(\mathbf{r}) - \frac{\langle \phi | G'' \rangle}{\langle \phi | \phi \rangle} \phi(\mathbf{r}) \tag{2.63}$$

Then we compute the conjugate gradient direction:

$$G^{(4)}(\mathbf{r}) = G^{(3)}(\mathbf{r}) - \frac{\langle G^{(3)} | G^{(3)} \rangle}{\langle G^{(3)}_{old} | G^{(3)}_{old} \rangle} G^{(4)}_{old}(\mathbf{r}) \tag{2.64}$$

Eq. (2.64) is the essence of the conjugate gradient method. The simpler and less-effective steepest descent method would result if we set $G^{(4)}(\mathbf{r}) = G^{(3)}(\mathbf{r})$. Then we orthogonalize one more time:

$$G^{(5)}(\mathbf{r}) = G^{(4)}(\mathbf{r}) - \frac{\langle \phi | G^{(4)} \rangle}{\langle \phi | \phi \rangle} \phi(\mathbf{r}) \tag{2.65}$$

Then we do line minimizations along the $G^{(5)}$ direction. We parameterize the new ϕ:

$$\phi^{new}(\mathbf{r}) = \cos(\theta)\phi(\mathbf{r}) + \sin(\theta)G^{(5)}(\mathbf{r}) \tag{2.66}$$

which by construction maintains the correct particle number. Then we minimize the energy with respect to θ, and the procedure repeats until sufficient minimization is achieved.

We can perform molecular dynamics with the Verlet method:

$$\mathbf{R}(t+\Delta t) = 2\mathbf{R}(t) - \mathbf{R}(t-\Delta t) + \frac{\mathbf{F}(t)}{m}(\Delta t)^2 \tag{2.67}$$

and we have found the limited-memory BFGS algorithm[29] to be efficient and useful for minimizing the energy with respect to the nuclear coordinates.

Chapter 3

Local pseudopotentials

The use of pseudopotentials, as described in 2.2.5, make electronic structure calculations much more efficient and allow much larger systems to be studied. In this chapter we review the theory behind the construction of pseudopotentials and the reasons behind their success. We then discuss why the pseudopotentials that result from most construction methods are unsuitable for use within OF-DFT. Then we present our method for generating pseudopotentials that are compatible with OF-DFT.

3.1 Norm-conserving pseudopotentials

The purpose of the pseudopotential is to incorporate the effects of the core electrons with an effective atomic potential, entirely eliminating the need to include the core wavefunctions in a calculation. Thus central to the theory of pseudopotentials is the study of the effects of the core.

As mentioned before, for most atoms there is a clear distinction between core and

valence electrons. The core electrons are localized near the nucleus and, being deep in the nucleus' Coulomb potential, are very tightly bound. When additional electrons are added to the atom, they can choose to also live deep in the Coulombic hole and maintain orthogonality with the core states by having either radial or angular oscillations (i.e. having a different radial or angular quantum number) at the expense of kinetic energy, or they can choose to live further away from the nucleus, avoiding the need to have high-energy wiggles but at the expense of not enjoying the deep Coulombic well; the electrons of course do whatever results in the lowest energy. The valence states are those electrons that live away from the nucleus.

When two atoms are brought together, the nuclear separation at which the energy is lowest is several Ångstroms. At this separation, and all larger separations, the core states are typically unaltered and do not overlap. This implies that the binding energy curve between the atoms is determined by the form of the valence wavefunctions in the interstitial regions outside of the core. The situation is similar when many atoms are brought together to form molecules or solids.

The combination of the core electrons and the nuclear potential can be replaced by a potential, or more generally a Hermitian operator, that "acts like the core," in that electrons scatter off the potential in the same way that they would scatter off the core.

An important result in the theory of wave scattering is that the nature of a spherical scatterer is contained in the scattering phase shifts between an incoming spherical wave and the resulting outgoing spherical wave. These phase shifts are found by determining the right linear combination of spherical waves outside the

scatterer that continuously and smoothly join to a solution within the scatterer; the phase shifts are thus determined by the logarithmic derivatives of the solutions at the boundary of the scatterer. There is a phase shift $\delta_l(E)$ associated with each angular momentum l value, and for the continuous spectrum of incoming energies E.

This line of thought suggests that in order to behave like the core, a pseudopotential should mimic its scattering phase shifts, or, equivalently, the logarithmic derivatives of its solutions to the radial Schrödinger equation. There exists an important identity relating these logarithmic derivatives at a radius R to the charge within R:

$$\begin{aligned} \frac{\mathrm{d}x[E]}{\mathrm{d}E} &= -\frac{2}{(r\phi_E)^2}\int_0^R (r\phi_E)^2 \mathrm{d}r, \\ \text{where } x[E] &\equiv \frac{\mathrm{d}}{\mathrm{d}r}\ln\phi_E(r)\Big|_R \end{aligned} \tag{3.1}$$

where $\phi_E(r)$ is a solution of the radial Schrödinger equation at an energy E (not necessarily an eigenvalue). Eq. (3.1) indicates a direct relationship between the logarithmic derivatives at R and the amount of charge contained within R. Eq. (3.1) provides an indirect means of controlling the variation of the logarithmic derivative at R with respect to the energy: instead one may vary the amount of charge within a sphere of radius R. The identity of Eq. (3.1) has been extended to include higher derivatives of $\ln\phi_E(R)$ with respect to E[30].

Another fact important to the theory of pseudopotentials is that a pseudopotential need not be a local potential, but in general may be any kind of operator. The form taken by most pseudopotentials used today is:

$$\langle \mathbf{r}|V^{\mathrm{psp}}|\mathbf{r}'\rangle = \sum_{lm} Y_{lm}^*(\Omega)V_l^{\mathrm{psp}}(r)\frac{\delta(r-r')}{r^2}Y_{lm}(\Omega') \tag{3.2}$$

This form of pseudopotential acts separately on the components of a wavefunction

with different total orbital angular momentum. This endows the pseudopotential with more degrees of freedom, allowing it so satisfy more desired properties.

Most pseudopotentials employed in electronic structure calculations are of the basic style of the norm-conserving pseudopotentials (NCPPs) developed by Hamann, Schlüter, and Chiang. The NCPP construction scheme proceeds as follows:

- An *all-electron* (AE) Kohn-Sham calculation is performed on a single atom, resulting in radial wavefunctions $u(r) = r\phi(r)$ that satisfy the radial version of the Kohn-Sham equations:

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V^{\mathrm{eff}}(r) + \frac{l(l+1)}{2r^2} - \epsilon_{nl}\right] u_{nl}^{\mathrm{AE}}(r) = 0,$$
$$V^{\mathrm{eff}}(r) \equiv V_{\mathrm{H}}[\rho^{\mathrm{AE}}; r] + V_{\mathrm{xc}}[\rho^{\mathrm{AE}}; r] - \frac{Z}{r} \tag{3.3}$$

  where V_{H} and V_{xc} are the Hartree and exchange-correlation potentials defined in Eqs. (2.25) and (2.26) respectively, and Z is the charge of the nucleus.

- A *pseudo-wavefunction* $u_l^{\mathrm{psp}}(r)$ is constructed for each different l quantum number represented among the valence states. The pseudo-wavefunction is constructed such that it is equal to the corresponding valence all-electron function $u_{nl}^{\mathrm{AE}}(r)$ for $r > R_c$, where R_c is the *core radius*, a radius chosen that defines the extent of the core, typically around one or two Ångstroms. Within R_c the pseudo-wavefunctions are chosen to have an arbitrary smooth form, but with a total charge within R_c equal to that of $u_{nl}^{\mathrm{AE}}(r)$:

$$\int_0^{R_c} |u_{nl}^{\mathrm{psp}}(r)|^2 \mathrm{d}r = \int_0^{R_c} |u_{nl}^{\mathrm{AE}}(r)|^2 \mathrm{d}r \tag{3.4}$$

  This is the norm-conserving condition that gives these pseudopotentials their name.

- An effective pseudopotential $V_l^{\mathrm{psp,eff}}$ is constructed by inverting the Schrödinger equation for each pseudo-wavefunction:

$$V_l^{\mathrm{psp,eff}}(r) \equiv \frac{\left[\frac{1}{2}(\mathrm{d}^2/\mathrm{d}r^2) + \epsilon_{nl}\right] u_{nl}^{\mathrm{AE}}(r)}{u_{nl}^{\mathrm{AE}}(r)} \tag{3.5}$$

- From each $V_l^{\mathrm{psp,eff}}$, we find the actual pseudopotential components V_l^{psp} that, when used in a Kohn-Sham calculation of the atom, yield the pseudo-wavefunctions $u_l^{\mathrm{psp}}(r)$ at the proper energies ϵ_{nl}:

$$V_l^{\mathrm{psp}}(r) = V_l^{\mathrm{psp,eff}} - V_{\mathrm{H}}[\rho^{\mathrm{psp}}; r] - V_{\mathrm{xc}}[\rho^{\mathrm{psp}}; r] - \frac{l(l+1)}{2r^2} \tag{3.6}$$

  These pseudopotentials, placed in Expr. (3.2), define the NCPP operator.

We note some properties of the NCPPs. By construction, they reproduce the valence eigenvalues of the all-electron atom. The wavefunctions have the same shape as the all-electron wavefunctions outside the core. This implies that the logarithmic derivatives at the eigen-energies are correct:

$$x^{\mathrm{psp}}[\epsilon_{nl}] = x^{\mathrm{AE}}[\epsilon_{nl}] \tag{3.7}$$

But because of the norm-conserving condition, Eq. (3.4), the logarithmic derivatives of the solutions of the pseudopotential are correct at energies *near* the ϵ_{nl} as well, via the identity (3.1):

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} x^{\mathrm{psp}}[\epsilon]\Big|_{\epsilon=\epsilon_{nl}} = \frac{\mathrm{d}}{\mathrm{d}\epsilon} x^{\mathrm{AE}}[\epsilon_{nl}]\Big|_{\epsilon=\epsilon_{nl}} \tag{3.8}$$

Thus a NCPP will scatter electrons in a way very similar to the combination of the nucleus and the core.

We note that a large degree of flexibility exists still in the construction of the pseudopotential. The pseudo-wavefunctions can have any convenient form within R_c as long as they are norm-conserving. This flexibility has usually been exploited to make calculations more efficient: by choosing the pseudo-wavefunctions to be as smooth a possible, one can usually then get away with using as few basis functions as possible in a calculation that uses these pseudopotentials.

3.2 Local pseudopotentials

Unfortunately the operator (3.2) cannot be used within an OF-DFT calculation. It only makes sense within a framework that uses orbitals, as it involves orbital projectors, and the whole framework of NCPPs is grounded in the theory of wave scattering. In fact, because the NCPPs are in general operators that are non-local in real space, i.e. $\langle \mathbf{r}|V^{\mathrm{psp}}|\mathbf{r}'\rangle \neq 0$ for $\mathbf{r} \neq \mathbf{r}'$, they are not rigorously applicable within the Hohenberg-Kohn density functional theory, as Eq. (2.16) would not hold if $V_{\mathrm{e-i}}$ were a non-local potential; one has to reformulate the theory with the basic variable being the single-particle density matrix $\gamma(\mathbf{r}, \mathbf{r}')$. Most practitioners of DFT nonetheless use NCPPs within the Hohenberg-Kohn-Sham framework, as one could take the point of view that a Kohn-Sham calculation using non-local pseudopotentials and only valence electrons is a "numerical trick" to approximating the important aspects of a (formally problem-free) Kohn-Sham calculation using all of the electrons and local Coulombic potentials to represent the nuclei.

If it happened that all of the $V_l^{\mathrm{psp}}(r)$ were identical, then the pseudopotential

would be a local operator:

$$\langle \mathbf{r}|V^{\mathrm{psp}}|\mathbf{r}'\rangle = V^{\mathrm{loc}}(r)\sum_{lm} Y^*_{lm}(\Omega)\frac{\delta(r-r')}{r^2}Y_{lm}(\Omega') = V^{\mathrm{loc}}(r)\delta(\mathbf{r}-\mathbf{r}'), \tag{3.9}$$

and the electron-ion energy would be expressible purely in terms of the density and not the Kohn-Sham orbitals:

$$E_{\mathrm{e-i}} = \int \rho(\mathbf{r})V^{\mathrm{psp}}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.10}$$

The goal of the work of this chapter is the construction of pseudopotentials that maintain the important properties satisfied by NCPPs, but that exploit the remaining freedom for pseudopotential construction in order to bring the different l-components into coincidence; i.e. the construction of local pseudopotentials. First we discuss some of the limitations that locality entails, and what has been done previously in the realm of local pseudopotentials.

3.2.1 Fundamental limitations of local pseudopotentials

The eigenvalues of all radial local potentials satisfy the following:

$$l > l' \Rightarrow \epsilon^{\mathrm{loc}}_{nl} > \epsilon^{\mathrm{loc}}_{nl'} \tag{3.11}$$

This result is easily demonstrated by noting that the radial Schrödinger equation contains the orbital angular momentum term:

$$\frac{l(l+1)}{2r^2} \tag{3.12}$$

which increases with l, and thus going from l' to $l > l'$ is a perturbation process that increases the energy. Non-local pseudopotentials, on the other hand, act separately on each l component and can be designed to have any spectrum whatsoever.

Eq. (3.12) bears unfortunate relevance to the possibilities of local pseudopotentials. If we consider a typical transition metal like Cobalt, which has an electronic structure of $[\mathrm{Ar}]3d^7 4s^2$, we see that its outermost electrons are the $3d$ and $4s$ electrons; both sets of electrons are important in determining the properties of Co, and hence must both be included in any electronic structure calculation involving Co, whether it uses pseudopotentials or not. But the $3d$ electrons have a higher l-number but a lower energy than the $4s$ electrons. Clearly a local pseudopotential cannot recreate this situation, unless, as proposed by Starkloff and Joannopolous[31], the entire $l = 3$ shell is included in the calculation as well. When this sort of thing is done, however, the computational savings from using pseudopotentials begins to slip away.

3.2.2 Previous work

The majority of pseudopotentials used in electronic structure calculations are of the basic norm-conserving form (although the "ultrasoft" pseudopotentials of Vanderbilt, and the closely-related PAW method are becoming mainstream). This is mainly due to their strong performance. Consequently the body of work dealing with local pseudopotentials is rather limited.

Prior to the development of NCPPs, local pseudopotentials were more common. A family of local pseudopotentials that are Coulombic for $r > R_c$ but equal to a constant for $r < R_c$ were developed. These potentials incorporate the idea that the valence electrons are repelled from the core by a particular amount, but do not invoke the full power of scattering theory. The Starkloff-Joannopoulos local pseudopotentials were designed to allow for the treatment of some transition metals, but have computational

disadvantages in that in reciprocal-space they decay slowly and hence require a large basis set or a fine grid.

Since the development of NCPPs, Perdew, Fiolhais, and co-workers[32, 33, 1, 34] have developed a form of local pseudopotential that they have dubbed the "evanescent core" pseudopotential (ECP), which is intended for the treatment of simple metals. They have the following mathematical form:

$$V^{\mathrm{ECP}}(r) = -\frac{Z}{R_c}\left\{\frac{1}{x}\left[1-(1+\beta x)e^{-\alpha x}-Ae^{-x}\right]\right\}, \tag{3.13}$$

$$x \equiv r/R_c \tag{3.14}$$

R_c is a core radius, and Z is the valence of the element in question. The potential is required to have vanishing first and third derivatives at $r = 0$, which determines the parameters A and β in terms of α. This results in a rapid decay in reciprocal space, $\tilde{V}^{\mathrm{ECP}}(q) \sim q^{-8}$ as $q \to \infty$. The remaining parameters, α and R_c, are set by requiring that (1) the pseudopotential yield the correct lattice constant for the $T = 0$ phase of the element, and (2) the pseudopotential should yield the correct "valence interstitial electron number" N_{int}, which is defined as the number of valence electrons in the interstitial region between the polyhedral surface of the Wigner-Seitz cell of a monatomic crystal and the surface of the largest inscribed sphere.

Given that the ECPs are constructed to reproduce certain bulk properties, it is not surprising that their general performance with bulk properties is good, including bulk properties that are not guaranteed simply by construction. The ECPs constructed for simple metals exhibit bulk moduli (and higher derivatives of the energy with respect to crystal volume) with an accuracy very similar to that of NCPPs. Most pseudopotentials are like ECPs in that they are constructed to behave well in one

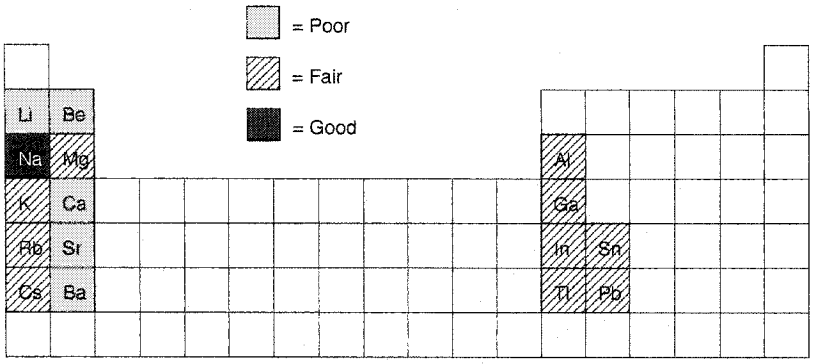

Figure 3.1: Qualitative summary of the transferability tests of the evanescent core pseudopotentials made by Nogueira et al.[1].

type of atomic environment; NCPPs by construction reproduce free atom valence properties well. The ability of a pseudopotential to properly mimic the true atomic core outside of this environment is the pseudopotential's *transferability*. Nogueira et al. tested the transferability of the ECPs by performing cluster and excited-state calculations for a number of simple metals and comparing to all-electron calculations. They find a range of transferability from poor to good. Their findings are qualitatively summarized in Fig. 3.1 The performance of the ECPs is encouraging in that they perform well enough to be useful, which is surprising in light of their simplicity. This gives hope that further improvement might be found with a more sophisticated pseudopotential.

We also mention the work of Watson et al.[35], who developed local pseudopotentials specifically for orbital-free DFT methods. With their method, a local pseudopotential is derived for use in an OF-DFT calculation with one particular kinetic energy functional, and in order to use a different functional, the pseudopotential must be re-derived. This is not a serious drawback; similarly, NCPPs are derived with respect

to a particular approximate exchange-correlation functional. In their approach, the first step is to perform a Kohn-Sham DFT calculation of an elemental crystal, using a reliable NCPP. This yields a charge density $\rho_{\mathrm{KS}}^{\mathrm{cryst}}(\mathbf{r})$ that has the translational symmetry of the crystal lattice. Then the unique local crystal potential, $V_{\mathrm{OF}}^{\mathrm{cryst}}(\mathbf{r})$ that yields $\rho_{\mathrm{KS}}^{\mathrm{cryst}}(\mathbf{r})$ in an OF-DFT calculation is found:

$$\frac{\delta T_{\mathrm{s}}^{\mathrm{approx}}}{\delta\rho}[\rho_{\mathrm{KS}}^{\mathrm{cryst}}(\mathbf{r})] + \frac{\delta E_{\mathrm{H}}}{\delta\rho}[\rho_{\mathrm{KS}}^{\mathrm{cryst}}(\mathbf{r})] + \frac{\delta E_{\mathrm{xc}}}{\delta\rho}[\rho_{\mathrm{KS}}^{\mathrm{cryst}}(\mathbf{r})] + V_{\mathrm{OF}}^{\mathrm{cryst}} = 0 \tag{3.15}$$

$V_{\mathrm{OF}}^{\mathrm{cryst}}(\mathbf{r})$ also has the symmetry of the crystal lattice, and can be seen as the convolution of an atomic potential with the structure factor S:

$$V_{\mathrm{OF}}^{\mathrm{cryst}}(\mathbf{r}) = (V_{\mathrm{OF}}^{\mathrm{atom}} * S)(\mathbf{r}) \tag{3.16}$$

and thus the atomic potential can be obtained by deconvolving in reciprocal space:

$$V_{\mathrm{OF}}^{\mathrm{atom}}(\mathbf{q}) = V_{\mathrm{OF}}^{\mathrm{cryst}}(\mathbf{q})/S(\mathbf{q}) \tag{3.17}$$

but in this manner $V_{\mathrm{OF}}^{\mathrm{atom}}$ is only defined at $\mathbf{q}$ vectors where the structure factor is non-zero, i.e. at the reciprocal lattice vectors. At these points, $V_{\mathrm{OF}}^{\mathrm{atom}}$ is invariant under the operations of the point group of the crystal lattice. Watson et al. obtain a local pseudopotential $V_{\mathrm{OF}}^{\mathrm{psp}}$ from $V_{\mathrm{OF}}^{\mathrm{atom}}$ by averaging over all reciprocal lattice vectors of equal length. This still defines $V_{\mathrm{OF}}^{\mathrm{psp}}(|\mathbf{q}|)$ at a discrete set of $|\mathbf{q}|$, and Watson et al. essentially interpolate to define $V_{\mathrm{OF}}^{\mathrm{psp}}(|\mathbf{q}|)$ for all $|\mathbf{q}|$. Later, in chapter 6 we use the same spherical symmetrization procedure to obtain an atomic charge density from a crystal charge density, and we show that this procedure is optimal in a particular sense.

Watson et al. show how the above method results in local pseudopotentials that show decent performance for bulk properties like bulk moduli and phonon dispersion

curves. However they do not report on non-bulk properties, leaving the degree of transferability of their potentials an open question. Their work is interesting in that they generate pseudopotentials in a unique way from the bulk environment, and that they use the density as the basic variable, which they connect to the potential via density functional theory.

3.3 Present work

Although not nearly as general or as transferable as NCPPs, the evanescent core potentials perform surprisingly well given their simplicity. We explore the possibilities of local pseudopotentials by utilizing a more sophisticated form that has much more flexibility and spans much more of "pseudopotential space." The pseudopotentials of Watson et al. introduce the use of the charge density as the basic variable determining the pseudopotential; we utilize density-based potentials as a starting point in our search for better local pseudopotentials.

We have developed a method for generating local pseudopotentials that proceeds roughly as follows: an initial-guess local pseudopotential is found by performing a charge density $\rightarrow$ potential inversion procedure similar to Watson et al., but with important differences. Then this initial pseudopotential is slowly perturbed in order to optimize its scattering properties. The present pseudopotentials are thus derived with reference to the free atom, and they take full advantage of the freedom available for the form of the pseudo-wavefunctions within the core.

More precisely, our method for generating pseudopotentials can be described in four steps:

1. A reference atomic or ionic configuration is chosen, and an all-electron calculation is performed for this configuration, resulting in a self-consistent AE potential $V^{\mathrm{AE}}(r)$ and an AE *valence* charge density $\rho^{\mathrm{AE}}(r)$. The charge density is "pseudized" to a smooth, nodeless form, with the same normalization as $\rho^{\mathrm{AE}}(r)$, inside some core radius r_c. This pseudo-density will be denoted by $\rho^{\mathrm{psp}}(r)$.

2. The pseudo-density $\rho^{\mathrm{psp}}(r)$ is inverted to yield the *unique* local potential $V^{\mathrm{I}}(r)$ for which $\rho^{\mathrm{psp}}(r)$ is the ground state density. Then the tail of $V^{\mathrm{I}}(r)$ is altered to yield a potential $V^{(0)}(r)$ that serves as an initial guess for the next stage.

3. A parameterized local potential $V(\alpha_i; r)$ is considered, where the parameters α_i determine the shape of the potential in the core region. For the values $\alpha_i = 0$ the parameterized potential corresponds to the initial guess potential, $V(\alpha_i = 0; r) = V^{(0)}(r)$. Starting with an initial guess of $\alpha_i = 0$, the α_i are varied to match the scattering properties of $V(\alpha_i; r)$ to those of $V^{\mathrm{AE}}(r)$ at several selected energies. The result of this optimization will be denoted by $V^{(1)}(r) \equiv V(\alpha_i^{\mathrm{opt}}; r)$.

4. Steps 1-3 are carried out for a number of reference electronic configurations $i = 1, ..., N$ of the atom, resulting in a set of optimized potentials $V_i^{(1)}(r)$. A linear combination of these potentials that produces the best bulk properties is taken, resulting in the final local pseudopotential,

$$V(r) = \sum_{i=1}^{N} \beta_i V_i^{(1)}(r). \tag{3.18}$$

Next we describe and justify these steps in detail.

3.3.1 AE calculation and pseudization

The details of AE calculations are well established and are available in the literature[36]. The pseudized charge density $\rho^{\text{psp}}(r)$ is constructed by pseudizing the AE wavefunctions in the manner described by Bachelet et al. in [4], and summing the charge densities from each pseudo-wavefunction multiplied by their occupation.

3.3.2 Inversion of the pseudodensity

A potential must be found that yields $\rho^{\text{psp}}(r)$ as the ground state charge density. The Hohenberg-Kohn theorem[5] guarantees that should such a (local) potential exist, it is unique, up to an additive constant. The process of finding this unique potential is non-trivial, and a number of methods have been developed to carry out this inversion [37, 38, 39, 40, 41, 42, 43]. We have found these methods difficult to converge numerically, and we have developed our own method that behaves in a satisfactory manner.

We find the potential that generates $\rho^{\text{psp}}(r)$ by exploiting a theorem due to Levy and Perdew[44] that equates this potential with a Lagrange multiplier which constrains the density. As described in chapter 2, in the constrained-search formalism of DFT the non-interacting kinetic energy functional is defined as:

$$T_s\left[\rho(\mathbf{r})\right] = \min\left[\sum_i \langle\phi_i| - \tfrac{1}{2}\nabla^2|\phi_i\rangle\right] \tag{3.19}$$

$$\text{such that } S_{ij} \equiv \langle\phi_i|\phi_j\rangle = \delta_{ij} \tag{3.20}$$

$$\text{and } \rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2 \tag{3.21}$$

As with the derivation of the Kohn-Sham method, we impose the orthogonality con-

straints with Lagrange multipliers ϵ_{ij}. However, now we additionally impose the density constraint, Eq. (3.21) with a continuous Lagrange multiplier $\lambda_\rho(r)$, and we find that the ϕ_i's minimize the following functional:

$$T_\lambda\left[\rho(\mathbf{r})\right] \equiv \sum_i \langle\phi_i| - \tfrac{1}{2}\nabla^2|\phi_i\rangle + \int d\mathbf{r}\lambda_\rho(\mathbf{r})\left[\sum_i |\phi_i(\mathbf{r})|^2 - \rho(\mathbf{r})\right] + \sum_{ij}\epsilon_{ij}\left[\langle\phi_i|\phi_j\rangle - \delta_{ij}\right] \tag{3.22}$$

Taking the variation with respect to ϕ_i^* gives:

$$\left[-\tfrac{1}{2}\nabla^2 + \lambda_\rho(\mathbf{r})\right]\phi_i(\mathbf{r}) = -\sum_j \epsilon_{ij}\phi_j(\mathbf{r}) \tag{3.23}$$

With a unitary transformation the matrix ϵ_{ij} can be made diagonal with eigenvalues $-\tilde{\epsilon}_i$ and the corresponding transformed eigenfunctions ϕ_i's satisfy:

$$\left[-\tfrac{1}{2}\nabla^2 + \lambda_\rho(\mathbf{r})\right]\tilde{\phi}_i(\mathbf{r}) = \tilde{\epsilon}_i\tilde{\phi}_i(\mathbf{r}) \tag{3.24}$$

This is a Schrödinger equation, with $\lambda_\rho(\mathbf{r})$ playing the role of the potential. The Hohenberg-Kohn theorem ensures that if such a Lagrange multiplier can be found, it is the unique local potential that generates $\rho(\mathbf{r})$. [1]

To find the correct form of the Lagrange multipliers, and hence find the effective potential that generates a particular charge density, we employ the augmented Lagrangian method of constrained minimization [45]. This is an algorithm for constrained optimization that is designed to enforce constraints with Lagrange multipliers, and simultaneously find the correct values that the multipliers should have. To the functional T_λ we add two terms:

$$T'_\lambda\left[\rho(\mathbf{r})\right] \equiv T_\lambda\left[\rho(\mathbf{r})\right] + \tfrac{1}{2}c_\lambda\int d\mathbf{r}\left[\sum_i |\phi_i(\mathbf{r})|^2 - \rho(\mathbf{r})\right]^2 + \tfrac{1}{2}c_\epsilon\sum_{ij}\left[S_{ij} - \delta_{ij}\right]^2 \tag{3.25}$$

[1] It is possible, in the case of non-V-representable $\rho(\mathbf{r})$, to arrive at a $\lambda_\rho(\mathbf{r})$ for which eqns. (3.24) are satisfied, but the resulting ϕ_i's will not be the *ground state* eigenfunctions of the system. This is easily checked, however, and was not found to occur.

These additional terms are penalty functions. Their effect is to drive the minimum of T'_λ closer to the desired constraints. The larger the values of c_λ and c_ϵ are, the closer the minimizing ϕ_i's will satisfy the density and orthonormality constraints. The minimization proceeds as follows: an initial guess is made for the Lagrange multipliers, $\lambda_\rho^{(0)}(\mathbf{r})$ and $\epsilon_{ij}^{(0)}$. with this choice of Lagrange multipliers $T'_\lambda[\rho(\mathbf{r})]$ is minimized with respect to the ϕ_i, resulting in a set of orbitals $\phi_i^{(0)}$ with overlaps $S_{ij}^{(0)}$. Then the Lagrange multipliers are updated as follows:

$$\lambda_\rho^{(n+1)}(\mathbf{r}) = \lambda_\rho^{(n)}(\mathbf{r}) + c_\lambda \left[\sum_i \left| \phi_i^{(n)}(\mathbf{r}) \right|^2 - \rho(\mathbf{r}) \right]$$
$$\epsilon_{ij}^{(n+1)} = \epsilon_{ij}^{(n)} + c_\epsilon \left[S_{ij}^{(n)} - \delta_{ij} \right] \tag{3.26}$$

where the superscripts are an iteration index. The Lagrange multipliers are updated by an amount proportional to the degree to which the minimum avoids the constraint manifold.

Following the update of the Lagrange multipliers, $T'_\lambda[\rho(\mathbf{r})]$ must be minimized again, and the process repeats itself. From the form of Lagrange multiplier update formula, Eq. (3.26), it is clear that the Lagrange multipliers continue to be updated until the minimizing ϕ_i's satisfy both the density and the orthonormality constraint. Using larger values for c_λ and c_ϵ will result in fewer iterations of this loop, but each minimization of $T'_\lambda[\rho(\mathbf{r})]$ will be more ill-conditioned. In practice, we have found that c_λ and c_ϵ can be chosen such that each minimization can be carried out in a reasonable number of steps, and the update loop need only be carried out between five and ten times to achieve satisfactory convergence of the Lagrange multipliers.

A method of determining the potential from the density due to Zhao, Morrison, and Parr[42] also employs a type of penalty functional, but in this method the result

of increasing the strength of the penalty function to infinity must be extrapolated (in the present method, this would correspond to increasing c_λ, $c_\epsilon \rightarrow \infty$). This results in an ill-conditioned minimization; the present method bypasses this difficulty by including Lagrange multiplier terms explicitly that aid in the constraint process.

The method was implemented for the special case of spherically symmetric charge densities and potentials, and utilized to find the effective potential $V^{\mathrm{I}}(r)$ that generates $\rho^{\mathrm{psp}}(r)$. It is natural to ask why $V^{\mathrm{I}}(r)$ is not a suitable pseudopotential itself. For all densities we have inverted, we found that the resulting potential does not have a correctly-shaped tail. When this potential is descreened,

$$V^{ID}(r) \equiv V^{\mathrm{I}}(r) - V_{\mathrm{H}}\left[\rho^{\mathrm{psp}}(r)\right] - V_{\mathrm{xc}}\left[\rho^{\mathrm{psp}}(r)\right] \tag{3.27}$$

the resulting potential $V^{\mathrm{ID}}(r)$, if it is to be a realistic pseudopotential, should rapidly approach $-Z/r$ outside of the core of the atom, where Z is the valence of the element under consideration. However, all inverted, descreened potentials $V^{\mathrm{ID}}(r)$ that we have generated deviate appreciably from $-Z/r$ at values of r significantly outside the core of the atom (up to $\sim$ 10 Bohr, see Fig. (3.2)). Examining the situation closer, one can see that there is no fundamental reason why $V^{\mathrm{ID}}(r)$, generated as described above, would have the proper Coulombic tail. The exponentially decaying tail of $\rho^{\mathrm{psp}}(r)$ should come from a potential that goes to zero as $r \rightarrow \infty$, but not necessarily as $-Z/r$. When such a potential is used in a bulk calculation, the deviation from the proper Coulombic form, although seemingly slight, provides enough repulsion to prevent the solid from forming.

For each $V^{\mathrm{ID}}(r)$ generated, a "fixed" potential $V^{(0)}(r)$ was constructed by match-

ing $V^{\mathrm{ID}}(r)$ to $-Z/r$ with a smooth step function:

$$V^{(0)}(r) = f(r/r_f)V^{\mathrm{ID}}(r) + [1 - f(r/r_f)]\frac{-Z}{r} \tag{3.28}$$

we have used an exponential form for $f(x)$:

$$f(x) = \exp\left(-x^{6.5}\right) \tag{3.29}$$

The "fixing" radius r_f is chosen to be the radius at which $\left|V^{\mathrm{ID}}(r) + Z/r\right|$ is minimized; such radii are typically $r_f \sim 2$ Bohr. Thus the fixed potentials $V^{(0)}(r)$ are explicitly constructed to have the correct Coulombic tail, while retaining the form of the inverted potential in the core.

This inversion procedure can be viewed as analogous to the method of Watson et al., except that it is with respect to the free atom, and that it is an inversion using the *exact* non-interacting kinetic energy functional.

3.3.3 Scattering optimization

In this step, a parameterized form for the potential is developed, and the parameters are varied to optimize its scattering properties. After completing this work, it was discovered that a similar optimization approach has been used before, but in the context of non-local pseudopotentials[46, 47]. The following parameterized form has been chosen:

$$V(\alpha_i; r) = V^{(0)}(r) + \sum_{i=1}^{N_B} \alpha_i B_i(r) \tag{3.30}$$

where the $B_i(r)$ are basis functions that quickly approach zero for $r > r_c$, so that the form of the potential *within* r_c may be varied by varying the parameters α_i, but the tail of $V(\alpha_i; r)$ will always be equal the the tail of $V^{(0)}(r)$.

Various forms for the $B_i(r)$ were tried, including Gaussians and polynomials; the results were not very sensitive to the choice of basis functions. Our final choice was damped cosines:

$$B_j(r) = f(r/r_c)\cos\left(\frac{\pi(j-1)r}{r_c}\right), j = 1, ..., N_B \tag{3.31}$$

with the same $f(x)$ from Eq. (3.29). A set of energies E_j at which scattering properties should be matched to AE scattering properties is chosen. Note that the scattering logarithmic derivatives $x_l[V(\alpha_i; r); E_j]$ at the chosen energies E_j are determined by the potential $V(\alpha_i; r)$, and thus are functions of the parameters α_i. Thus we can denote the scattering logarithmic derivatives of $V(\alpha_i, r)$ by $x_l^{\text{param}}[\alpha_i; E_j]$. We define the penalty function:

$$P_1(\alpha_i) = \sum_j \sum_{l=0}^{l_{max}} \left(x_l^{\text{param}}[\alpha_i; E_j] - x_l\left[V^{\text{AE}}; E_j\right]\right)^2 \tag{3.32}$$

which measures deviation from the desired scattering properties. However, unrestricted minimization of P_1 results in unphysical potentials with wild variations. This can be remedied by requiring the potential to be smooth, which yields a more physical functional form, and additionally allows the potential to be represented with a smaller number of plane waves. The smoothness requirement is implemented with a second penalty function, involving the potential:

$$P_2(\alpha_i) = \int_0^\infty dr \left(\frac{d^2}{dr^2} V(\alpha_i; r)\right)^2 \tag{3.33}$$

The two penalty functions are combined into an overall penalty function:

$$P(\alpha_i) = P_1(\alpha_i) + \gamma P_2(\alpha_i) \tag{3.34}$$

and γ is chosen to balance the strength of the two penalty functions.

The optimized pseudopotentials $V(\alpha_i; r)$ were found by minimizing $P(\alpha_i)$ starting from an initial guess of $\alpha_i = 0$. Generally, $P(\alpha_i)$ is a very complicated function of its arguments so gradient-based minimization techniques like conjugate gradients algorithms are probably not effective here. Instead, a simple scheme of following the downhill gradient was employed to minimize $P(\alpha_i)$. At each step in the optimization of the α_i, each parameter was varied slightly, $\alpha_i \rightarrow \alpha_i + \Delta\alpha_i$, the scattering properties $x_l[..., \alpha_i + \Delta\alpha_i, ...; E_j]$ were recalculated, and the change in the penalty function $(\Delta P)_i$ was obtained. The gradient of P was approximated as $(\nabla_\alpha I_{tot})_i \simeq \frac{(\Delta P)_i}{\Delta\alpha_i}$. Then a step was taken in α_i-space, $\vec{\alpha} \rightarrow \vec{\alpha} - \kappa\vec{\nabla}_\alpha I_{tot}$ for some small κ. The minimization was pursued until the scattering penalty, $P_1(\alpha_i)$ did not decrease significantly. A useful optimization schedule was found: one optimizes P with one basis function, $N_B = 1$, and one uses the resulting potential as an initial guess potential and one optimizes P with $N_B = 2$, and so on, until the improvement of P is slight, or until the resulting optimized potentials start to contain small-amplitude, high-frequency oscillations that P_2 does not suppress.

3.3.4 Linear combination of optimized potentials

In general, although the potentials $V(\alpha_i; r)$ emerging from the previous step were found to match the scattering properties of the corresponding AE potentials quite well, they did not accurately reproduce properties of the bulk phase. However, it was found that when optimized potentials were generated for several electronic and ionic configurations of the element under consideration, the lattice constants and bulk moduli obtained from them were consistently "harder" (smaller lattice constant and

higher bulk modulus) or "softer" than the experimental values. Furthermore, it was found that by taking a linear combination of optimized potentials, as in Eq. (3.18), a potential with accurate bulk properties was obtained. It was this linear combination that was taken to be the final local pseudopotential.

Although the scattering logarithmic derivatives are not linearly related to the potential, and thus taking a linear combination of potentials will not result in a linear combination of logarithmic derivatives, they will be roughly interpolated by the linear combination. Thus, by taking a linear combination of potentials optimized to different reference states, one can expect that the resulting potential will match the scattering properties of all of the reference states, albeit to a lesser degree.

It has been argued that the the reproduction of the AE chemical hardness is an important pseudopotential transferability criterion[46, 48]. The elements of the chemical hardness matrix are derivatives of the eigenvalues of a system with respect to the orbital occupation numbers. Matching the eigenvalues of a particular configuration ensures, via the arguments of section 3.1, good scattering properties of this configuration. By matching the derivatives of the eigenvalues with respect to occupation numbers ensures that the scattering properties remain well-matched as the atomic configuration changes. The linear combination of potentials derived from different configurations in the present method is, in effect, another way to achieve good scattering properties simultaneously for different atomic configurations with a single potential.

3.4 Application to Al

The above method was first be applied to the case of Al. AE calculations were carried out with the program described in [36]. The pseudo-density was constructed from the pseudo-wavefunctions, and inverted to yield a local potential. This potential, inverted from a ground state Al:$3s^2 3p^1$ pseudodensity, deviates from the correct $-Z/r$ behavior, which is shown in Fig. 3.2. After fixing its tail as described above, this potential was used as an initial guess for the next stage of the procedure, in which the potential's scattering properties are matched to those of the Al:$3s^2 3p^1$ AE potential. The optimized potential is plotted in Fig. 3.4, and Fig. 3.3 shows a plot of its scattering logarithmic derivatives as a function of energy; the results from the pseudopotential compare well with the scattering logarithmic derivatives of the AE potential.

An optimized potential was similarly determined for the excited ionic configuration Al:$3s^1 3p^1$, which is also shown in Fig. 3.4. Despite the success in optimizing the scattering properties of both potentials, neither potential simulates bulk fcc Al very well. Bulk calculations were done using the exchange-correlation functional of Perdew and Zunger[49, 21] and Monkhorst-Pack k-point sampling[50]. The potential optimized to ground state AE scattering predicts a bulk phase for Al with a lattice constant 9% smaller than experiment, and a bulk modulus nearly a factor of 2 larger than experiment. On the other hand, the potential optimized to the excited ionic AE scattering predicts a bulk phase with a lattice constant 9% larger than the experimental value, and a bulk modulus nearly a factor of 1/2 smaller than experiment.

By taking the average of these two potentials ($\beta_1 = \beta_2 = 0.5$ in Eq. (3.18)), we

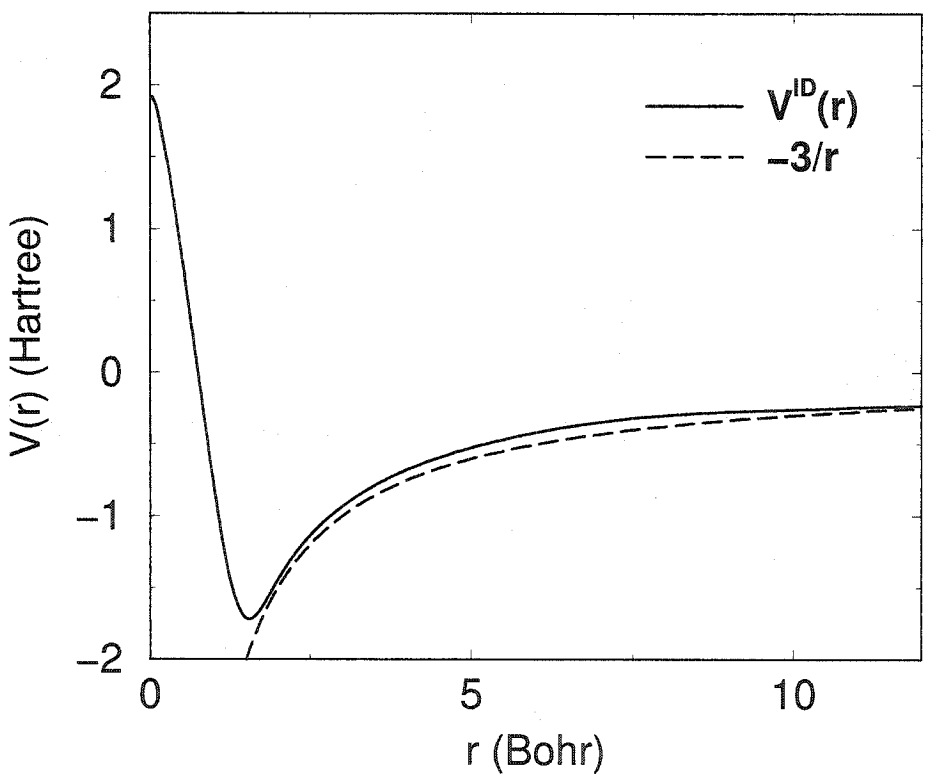

Figure 3.2: Potential resulting from inverting the Al:$3s^2 3p^1$ (ground state) pseudo-density. Note the deviation from the correct $-3/r$ Coulombic form in the tail to large values of r.

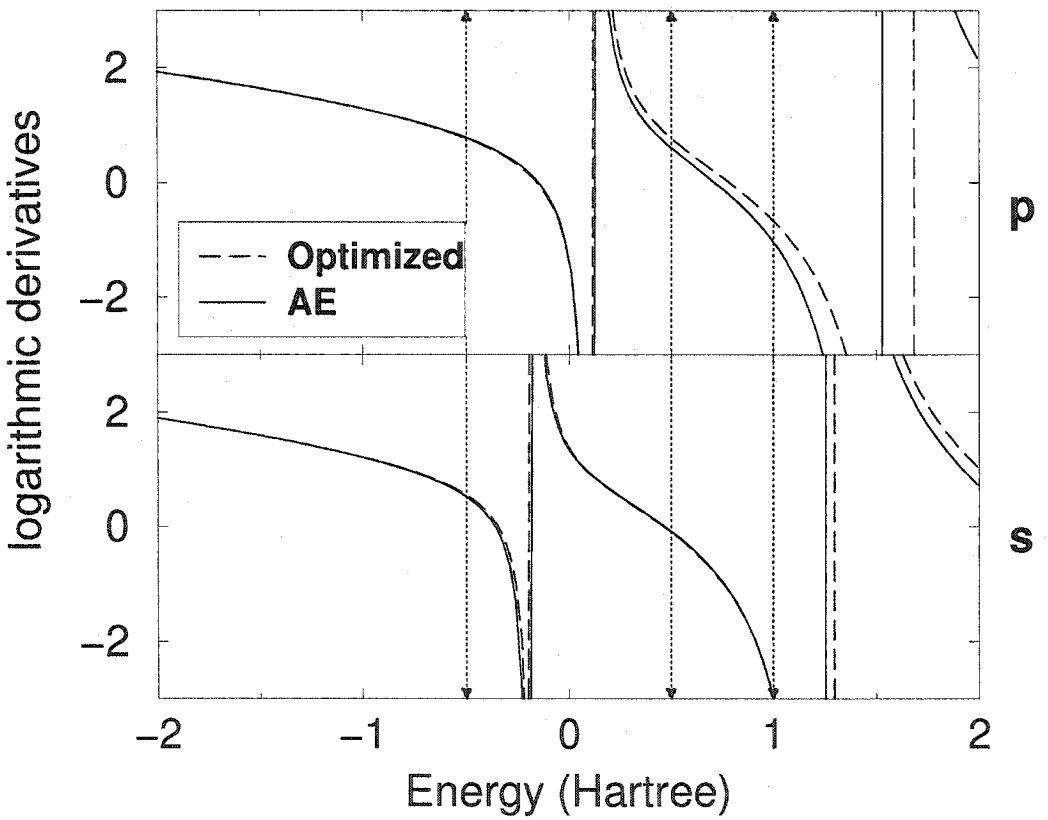

Figure 3.3: Scattering logarithmic derivatives of the optimized potential compared to that of the AE potential of Al:$3s^2 3p^1$. The logarithmic derivatives were sampled at a radius of $r_s = 4.0$ Bohr. The vertical arrows indicate the energies E_j at which the logarithmic derivatives were evaluated.

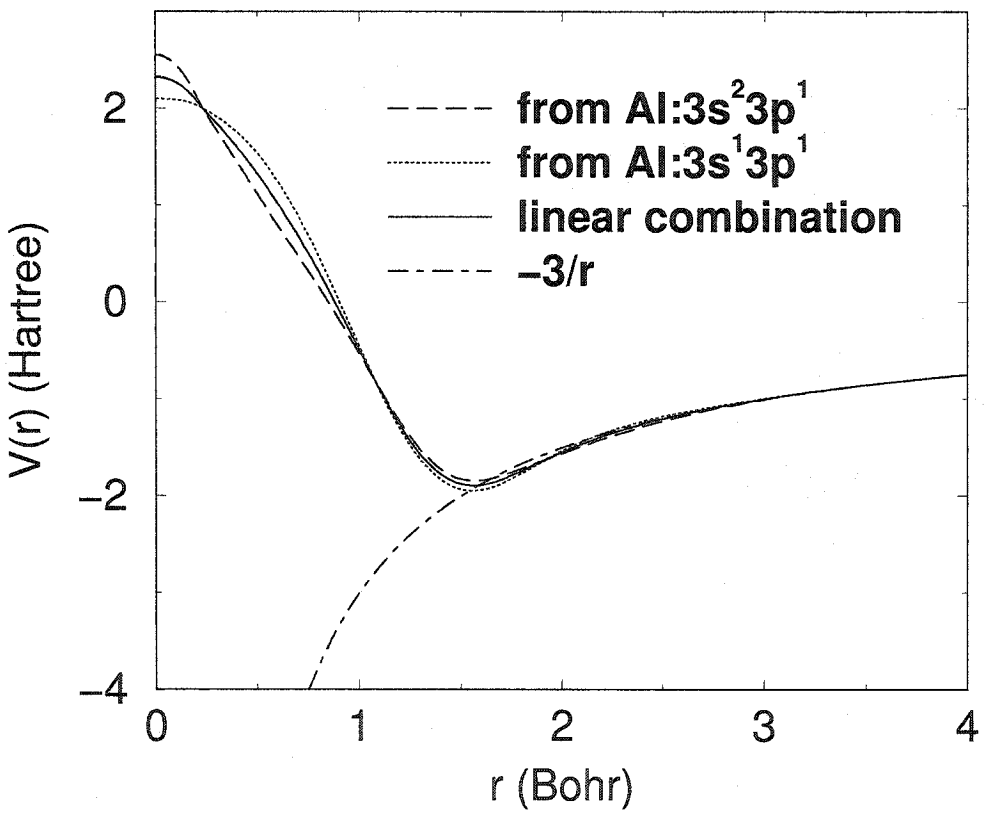

Figure 3.4: Optimized potentials for aluminum with scattering properties matched to those of the Al:$3s^2 3p^1$ and Al:$3s^1 3p^1$ AE potentials; also shown is the linear combination of these two optimized potentials.

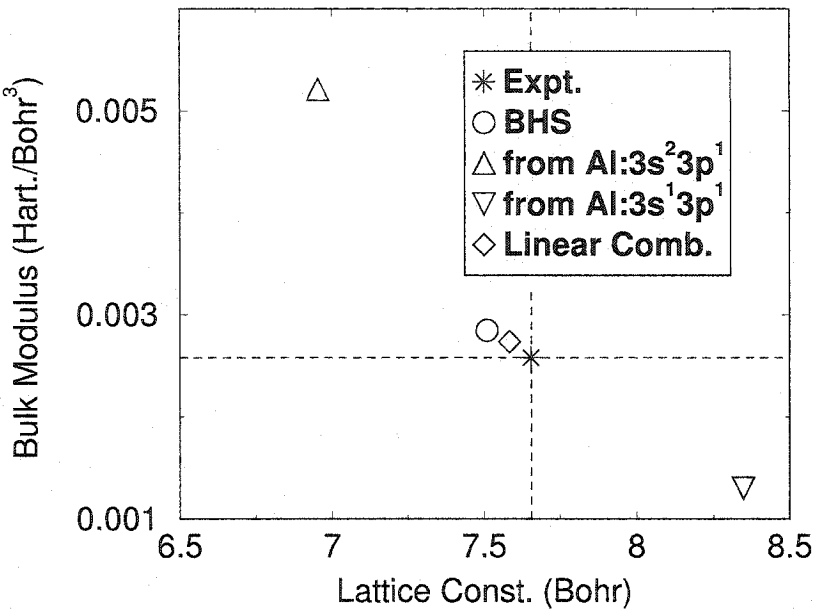

Figure 3.5: Physical properties of the optimized potentials and the linear combination potential compared with the experimental values. Also shown for comparison are the physical properties resulting from a non-local norm-conserving pseudopotential of the BHS type.

obtained a potential with very satisfactory bulk properties; the lattice constant is smaller than experiment by merely 1%, and the bulk modulus larger than experiment by 6%. These results are actually better than what is obtained from a non-local Bachelet-Hamann-Schlüter (BHS) pseudopotential, as shown in Fig. 3.5. Fig. 3.6 displays the energy per atom (relative to the energy of the isolated atom) as a function of the volume per atom, for the fcc, bcc, hcp, and sc structures, for the linear combination pseudopotential, a non-local, norm-conserving BHS pseudopotential, and the Goodwin-Needs-Heine (GNH) local pseudopotential [2]. Physical properties calculated with the present pseudopotential match the BHS potential's properties more closely than does the GNH pseudopotential, as shown in table 3.1.

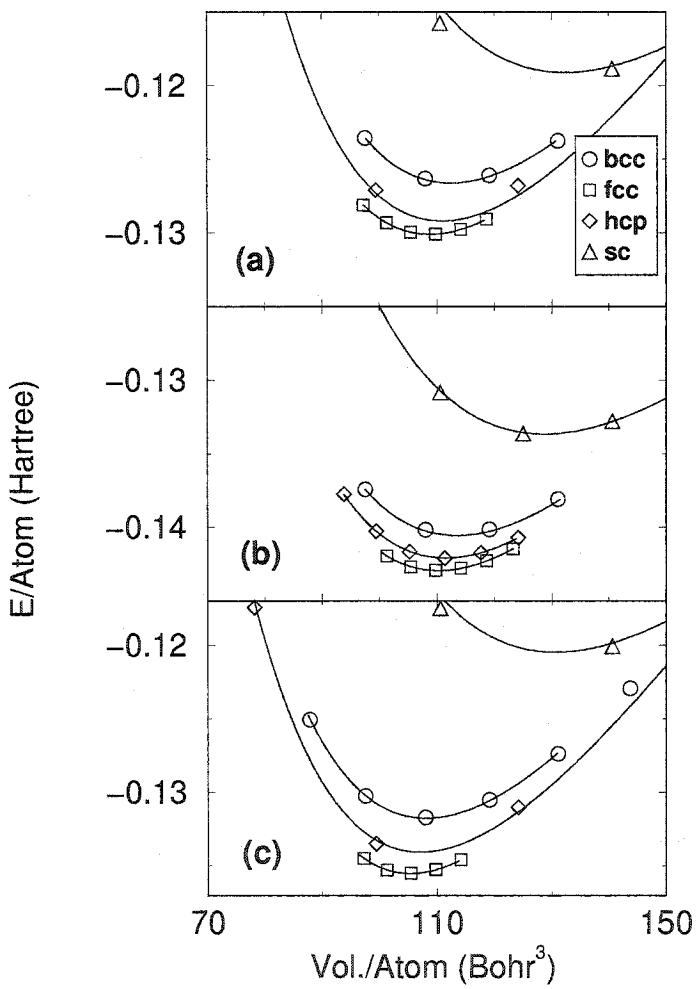

Figure 3.6: Energy/atom plotted versus volume/atom for aluminum in four different crystal structures, for (a) the present potential, (b) the Goodwin-Needs-Heine[2] local pseudopotential, and (c) a non-local BHS pseudopotential.

Table 3.1: Calculated physical properties for bulk aluminum for the Bachelet-Hamann-Schlüter[4] (BHS) non-local pseudopotential, the Goodwin-Needs-Heine[2] (GNH) local pseudopotential, and the present local pseudopotential. Atomic units are used. The ideal $c/a = \sqrt{8/3}$ was maintained in the hcp calculations.

crystal struct.	latt. const. (Bohr)			compressibility (Bohr^3/Hartree)			cohesive energy (Hartree)		
	BHS	GNH	Present	BHS	GNH	Present	BHS	GNH	Present
fcc	7.50	7.61	7.58	347	426	371	-0.135	-0.143	-0.130
hcp	4.75	4.81	4.80	386	435	376	-0.134	-0.142	-0.129
bcc	6.00	6.10	6.08	398	461	405	-0.131	-0.140	-0.127
sc	5.07	5.05	5.10	538	556	538	-0.120	-0.134	-0.119

In conclusion, we have developed a method for the generation of local pseudopotentials for electronic structure calculations. Although there are many possible choices to make during the course of generating a pseudopotential (for instance the scattering optimization energies E_j, the number of optimization parameters, and the reference states chosen), we have shown that with judicious parameters and a well-motivated optimization procedure, these pseudopotentials perform well. We demonstrated that for the test case of Al, the method is very effective in producing a useful local pseudopotential. Typical norm-conserving pseudopotentials are inherently non-local; we believe that the present method provides a general scheme for producing local pseudopotentials that optimally reproduce characteristics of the corresponding AE potential, while maintaining the spirit of *ab initio* calculations to the extent possible. In particular, we note that no experimental input is needed to construct the optimized potentials $V(\alpha_i; r)$, while the linear combination of such potentials can be chosen in the manner that suits best specific applications, providing additional flexibility in the definition of the final potential.

Chapter 4

Orbital-free methods in real space

Electronic structure calculations based on density functional theory[5, 6] (DFT) and employing an approximate kinetic energy functional[51, 52, 53, 54, 55, 56, 25, 26, 3, 57, 58] have been shown to yield accurate energetics for a variety of physical systems, with a considerably smaller expenditure of computer effort than traditional schemes. A drawback of many existing kinetic energy functionals is the necessity to evaluate them in part in reciprocal space. The benefits of doing electronic structure calculations exclusively in real space are several. Foremost among them is their ability to efficiently simulate isolated systems, whereas reciprocal space methods would require a large supercell with a significant portion of the volume devoted to uninteresting vacuum. A direct extension of this feature is the possibility of using arbitrary boundary conditions rather than the strict periodic boundary conditions underlying reciprocal space approaches; this should be of paramount importance in dealing with complex structures which cannot be accommodated by simple periodic boundary conditions, such as dislocations, cracks, etc. Finally, real space methods can be readily

parallelized for efficient computations on parallel computer architectures.

In this chapter we examine the reasons for the evaluation in reciprocal space of many kinetic energy functionals and propose a new functional form that can reproduce the energetics of those functionals but does not require any reciprocal space evaluations. The performance of the new functionals is evaluated and compared to the existing reciprocal space functionals as well as to the traditional approaches based on the calculation of electronic states.

DFT in its usual guise consists of solving the Kohn-Sham (KS) equations[6]. The fictitious KS wavefunctions allow the exact evaluation of $T_s[\rho]$, the kinetic energy of non-interacting fermions with a density $\rho(\mathbf{r})$. Recent calculations have shown that for certain systems the solution of the KS equations, usually the most computationally demanding part of electronic structure calculations, can be bypassed using an approximate form for $T_s[\rho]$[58]. A fair number of approximate non-interacting kinetic energy functionals have been proposed for use within such orbital-free methods. A subset of them are similar in form[55, 51, 54, 52, 56, 25, 26, 3, 57, 58], sharing the following common traits:

1. A major ingredient of $T_s[\rho]$ is the Thomas-Fermi (TF) energy[17, 18]:

$$T_{TF}[\rho] = C_{TF} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} \tag{4.1}$$

where $C_{TF} = \frac{3}{10}(3\pi^2)^{2/3}$. Here and throughout this article atomic units ($\hbar = m_e = e = 1$) are employed. The TF energy approximates the kinetic energy in an element of space $d\mathbf{r}$ with that of a homogeneous non-interacting electron gas with a density $\rho(\mathbf{r})$. Hence for homogeneous densities $\rho(\mathbf{r}) = \rho_0$, $T_{TF}[\rho(\mathbf{r})]$ is exact.

2. Another important contribution is the von Weizsäcker energy[27], given by:

$$T_{vW}[\rho] = \frac{1}{2}\int \sqrt{\rho(\mathbf{r})}\,\nabla^2\sqrt{\rho(\mathbf{r})}\,d\mathbf{r} \tag{4.2}$$

It can be readily shown that the von Weizsäcker energy yields the correct kinetic energy for a non-interacting fermion density that consists of a single orbital, i.e. a one- or two-electron density. Also, for any density, the von Weizsäcker energy yields the energy that a system of non-interacting bosons of density $\rho(\mathbf{r})$ would have. This term is a lower bound of $T_s[\rho]$.

3. The response of a homogeneous non-interacting Fermi gas to a small perturbation is known exactly[28]. This response can be related to the second functional derivative of $T_s[\rho]$ evaluated at uniform density $\rho(\mathbf{r}) = \rho_0$. The Fourier transform of this functional derivative is given by:

$$\hat{\mathcal{F}}\left[\left.\frac{\delta^2 T_s}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0}\right] = -\frac{1}{\chi_{Lind}(q)} \tag{4.3}$$

$$\chi_{Lind}(q) = -\frac{k_F}{\pi^2}\left[\frac{1}{2} + \frac{1-q^2}{4q}\ln\left|\frac{1+q}{1-q}\right|\right] \tag{4.4}$$

where $\chi_{Lind}(q)$ is the Lindhard response function, $\hat{\mathcal{F}}[f(\mathbf{r})] = \int f(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r}$ denotes the action of the Fourier transform on the function $f(\mathbf{r})$, $\mathbf{q} = \mathbf{k}/2k_F$, and $q = |\mathbf{q}|$. In order to satisfy Eq.(4.3) in addition to items 1 and 2 above, the kinetic energy functionals under consideration here include a term $T_K[\rho]$, which will be hereafter referred to as the *kernel energy*:

$$T_K[\rho] = \int f\left(\rho(\mathbf{r})\right) K(|\mathbf{r}-\mathbf{r}'|)g(\rho(\mathbf{r}'))d\mathbf{r}\,d\mathbf{r}' \tag{4.5}$$

where for the moment $f(\rho)$, $g(\rho)$ are arbitrary functions that can be chosen to satisfy known limits of the exact $T_s[\rho]$.

The total kinetic energy functional is taken to be the sum of these terms:

$$T_s[\rho] \simeq T_{TF}[\rho] + T_{vW}[\rho] + T_K[\rho] \tag{4.6}$$

By plugging Eq.(4.6) into Eq.(4.3), it is seen that $T_s[\rho]$ exhibits the correct linear response, provided:

$$\begin{aligned}\hat{\mathcal{F}}\left[\left.\frac{\delta^2 T_K}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0}\right] &= -\frac{1}{\chi_{Lind}(q)} \\ -\hat{\mathcal{F}}\left[\left.\frac{\delta^2 T_{TF}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0}\right] &- \hat{\mathcal{F}}\left[\left.\frac{\delta^2 T_{vW}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0}\right]\end{aligned} \tag{4.7}$$

The functional derivatives of T_{TF} and T_{vW} are well known, and the second functional derivative of T_K can, by design, be easily evaluated, so that Eq. (4.7) takes the form:

$$\frac{2k_F}{\pi^2} f'(\rho_0) g'(\rho_0) K(q) = \hat{K}(q) \tag{4.8}$$

$$\hat{K}(q) \equiv -\frac{k_F}{\pi^2 \chi_{Lind}(q)} - 1 - 3q^2 \tag{4.9}$$

For any choice of the functions f and g, $K(q)$ can be readily chosen such that the total kinetic energy functional exhibits the correct linear response. The different kinetic energy functionals considered in this paper differ mostly in the choice of the functions f and g.

In order to use kinetic energy functionals in actual electronic structure calculations, an algorithm must be developed for their action on discrete representations of the charge density. It is clear that $T_{TF}[\rho(\mathbf{r})]$ and $T_{vW}[\rho(\mathbf{r})]$ can be computed easily and efficiently on a grid in real space:

$$T_{TF}[\rho_i] = \Omega C_{TF} \sum_i \rho_i^{5/3} \tag{4.10}$$

$$T_{vW}[\rho_i] = \frac{\Omega}{2} \sum_{ij} \sqrt{\rho_i}\, \Delta_{ij} \sqrt{\rho_j} \tag{4.11}$$

where Ω is the volume per grid point, and Δ_{ij} is a discrete representation of the Laplacian operator. In principle, we could also compute the kernel energy $T_K[\rho(\mathbf{r})]$ in real space as follows:

$$T_K[\rho_i] = \sum_{ij} f\left(\rho_i\right) K(|\mathbf{r}_j - \mathbf{r}_j|) g\left(\rho_j\right) \tag{4.12}$$

However, this is impractical because of the specific form of $K(r)$. $K(r)$ is the Fourier transform of $\hat{K}(q)$ (see Eq.(4.9)), which is a non-trivial Fourier transform. Herring[59] has shown how it may be evaluated numerically. The quantity $K(r)$ is shown in Fig. 4.1, indicating that $K(r)$ does not decay rapidly. In order to evaluate

$$\int K(|\mathbf{r} - \mathbf{r}'|) h(\mathbf{r}') d\mathbf{r}' \tag{4.13}$$

accurately, the integral needs to be evaluated over a sphere centered at $\mathbf{r}$ of a radius equal to λ, with λ a large value. Such a convolution integral can be computed efficiently in reciprocal space:

$$\hat{\mathcal{F}}\left[\int K(|\mathbf{r} - \mathbf{r}'|) h(\mathbf{r}') d\mathbf{r}'\right] = \hat{\mathcal{F}}\left[K(r)\right] \hat{\mathcal{F}}\left[h(\mathbf{r})\right] \tag{4.14}$$

Starting from $K(r)$ and $h(\mathbf{r})$ in real space, three Fast Fourier Transforms (FFTs) are required to evaluate the convolution: two forward transformations, and one reverse transformation. The drawback of computing the kinetic energy functional with FFTs is that this approach maps the problem to a *periodic tiling* of the system of interest. This periodicity can have consequences on the resulting physics that in many cases are undesirable and can lead to erroneous results.

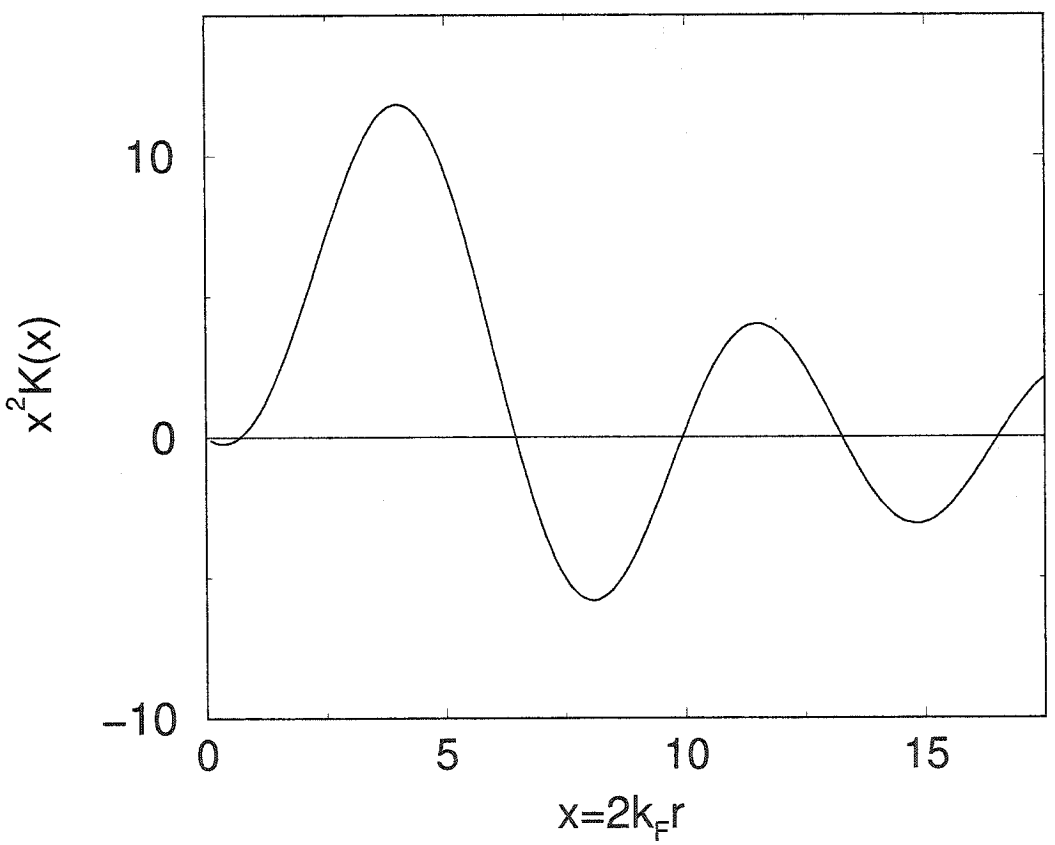

Figure 4.1: Kinetic energy functional kernel $K(x)$, where $x = 2k_F r$, multiplied by x^2, showing its long-ranged nature.

The problem with evaluating the kinetic energy functionals described above in real space lies in the efficient evaluation of convolution integrals. Convolution integrals with a long-ranged kernel can sometimes be evaluated efficiently in real space in an indirect way. For instance, the electrostatic potential of a charge distribution $\rho(\mathbf{r})$ is the convolution of $\rho(\mathbf{r})$ with the very long-ranged $1/r$:

$$\Phi(\mathbf{r}) = -\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \tag{4.15}$$

The integral in this expression can be computed efficiently in real space by solving the Poisson equation:

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}), \tag{4.16}$$

In general, as an alternative to evaluating the convolution

$$V(\mathbf{r}) = \int G(\mathbf{r}-\mathbf{r}')f(\mathbf{r}')d\mathbf{r}', \tag{4.17}$$

one may equivalently solve the integral equation:

$$\begin{aligned} \int H(\mathbf{r}-\mathbf{r}')V(\mathbf{r}')d\mathbf{r}' &= f(\mathbf{r}'), \\ H(\mathbf{r}) &= \hat{\mathcal{F}}^{-1}\left[\frac{1}{\hat{\mathcal{F}}[G(\mathbf{r})]}\right]. \end{aligned} \tag{4.18}$$

In discretized form, this integral equation is a standard linear problem:

$$H_{ij}V_j = f_i \tag{4.19}$$

which can be solved by an iterative linear solver[60]; the efficiency of such approaches depends on the efficiency of multiplying an arbitrary vector by a matrix. For the case of the Poisson equation, the matrix H_{ij} becomes a discrete representation of

the well-localized Laplacian operator, and hence matrix-vector multiplications are efficient. However, when the integral equation that corresponds to the convolution of Eq.(4.5) is constructed, the resulting function $H_K(r)$ is found to be long-ranged. The long-ranged nature of both $K(r)$ and $H_K(r)$ is due to the logarithmic divergence of the slope of $\chi_{Lind}(q)$ at $q = 1$, which causes long-ranged oscillations to appear in its Fourier transform. Thus, matrix-vector multiplications by $H_{ij} = H_K(|\mathbf{r}_j - \mathbf{r}_j|)$, and the solution of the integral Eq.(4.18), are inefficient.

We are therefore interested in developing a method that circumvents these difficulties. As a first step toward this goal, we note that the kernel appearing in the class of kinetic energy functionals under consideration, $\hat{K}(q)$, can be fit well by the rational function:

$$\tilde{K}(q) = \frac{N_2 q^2 + \cdots + N_{2m} q^{2m}}{D_0 + D_2 q^2 + \cdots + D_{2m} q^{2m}}, \tag{4.20}$$

with appropriate choices of the real coefficients N_i and D_i. The odd powers of q are omitted because in the Taylor expansions of $\hat{K}(q)$ about 0 and ∞ only even powers of q appear. The quality of the fit is shown in Fig. 4.2 for several values of m. Next we note that $\tilde{K}(q)$ can be separated into terms of the form:

$$\tilde{K}(q) = \sum_{j=1}^{m} \frac{P_j q^2}{q^2 + Q_j}, \tag{4.21}$$

where the P_j and Q_j are now complex numbers. With this expression, the convolution of a function $f(\mathbf{r})$ with $\tilde{K}(r)$ in reciprocal space becomes:

$$V(\mathbf{q}) \equiv \tilde{K}(q) f(\mathbf{q}) = V_1(\mathbf{q}) + \cdots + V_m(\mathbf{q}), \tag{4.22}$$

where

$$V_j(\mathbf{q}) = \frac{P_j q^2}{q^2 + Q_j} f(\mathbf{q}). \tag{4.23}$$

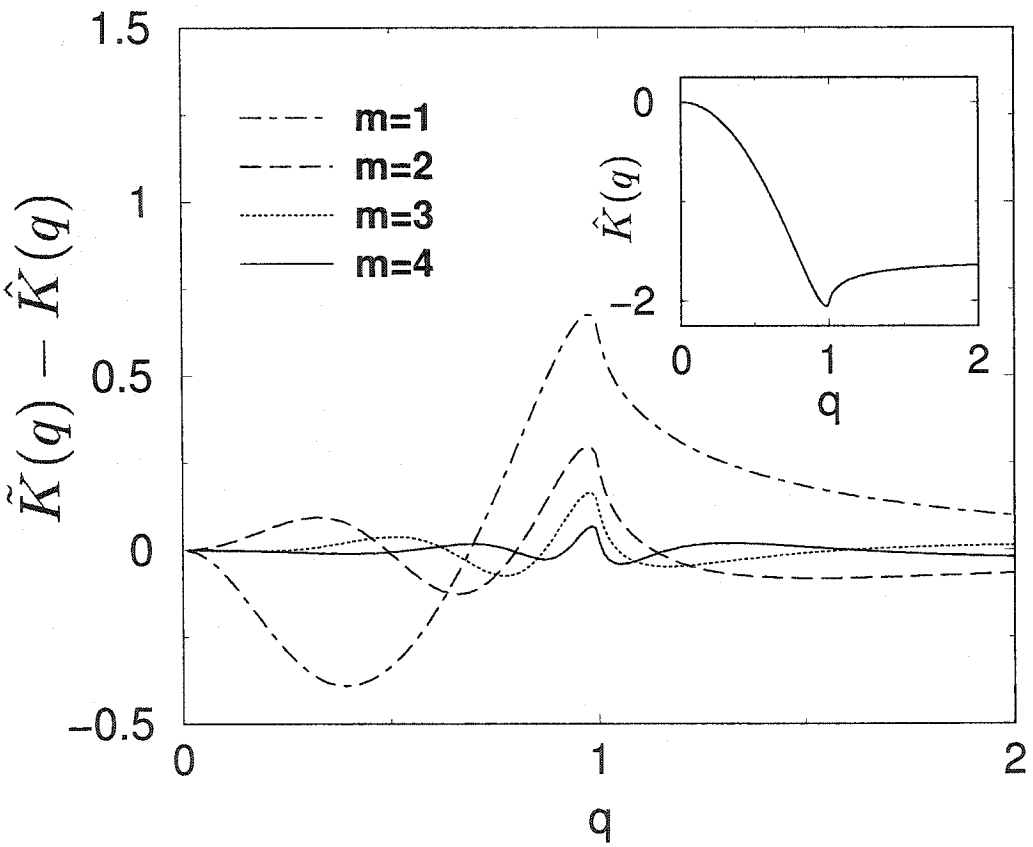

Figure 4.2: The difference between the rational fitting functions $\tilde{K}(q)$ (Eq.(4.20)) and $\hat{K}(q)$ (Eq.(4.9)) with the number of terms ranging from $m = 1$ to $m = 4$. The inset shows $\hat{K}(q)$.

The $V_i(\mathbf{r})$ can be computed efficiently in real space by solving:

$$\begin{aligned} (q^2 + Q_j)V_j(\mathbf{q}) &= P_j q^2 f(\mathbf{q}) \\ \Rightarrow \left[-\frac{1}{(2k_F)^2}\nabla^2 + Q_j\right] V_j(\mathbf{r}) &= -\frac{P_j}{(2k_F)^2}\nabla^2 f(\mathbf{r}) \end{aligned} \tag{4.24}$$

that is, $V_j(\mathbf{r})$ is the solution of a complex Helmholtz equation. A shortcut in computing the $V_j(\mathbf{r})$ results from the fact that $\tilde{K}(q)$, and thus the sum of the $V_j(\mathbf{r})$, is purely real. For every pair of coefficients $\{P_j, Q_j\}$, another pair $\{P_k, Q_k\} = \{P_j^*, Q_j^*\}$ must also appear in the expansion. It follows that $V_k(\mathbf{r}) = V_j^*(\mathbf{r})$, and thus only half of the $V_j(\mathbf{r})$ need be computed. A generalization of the kinetic energy functional with the form of Eq.(4.5) has been developed by Wang et al.[3]. These functionals can also be treated in real space with the present method, as discussed in section 4.1.

The following issue regarding the form of the approximation of Eq.(4.20) deserves further discussion: an important feature of the Lindhard response function lies in the logarithmic singularity in its slope at $q = 1$. As discussed above, this singularity manifests itself mathematically in the long-ranged nature of $K(r)$ and $H_K(r)$. Not surprisingly, this singularity also has important physical consequences, such as Friedel oscillations and the Kohn effect[61]. The approximate kernel $\tilde{K}(q)$ does not exhibit the singularity. It may seem then that from a physical standpoint, $\tilde{K}(q)$ may not adequately describe the kinetic energy of the electron gas. However, it should be noted that in a discrete representation of the problem, the exact singularity at $q = 1$ will not be seen. Furthermore, at non-zero electronic temperatures, however small, the singularity in $\chi_{Lind}(q)$ disappears. Thus, one could think of the fitting form $\tilde{K}(q)$ as representing $K(q)$ for a small but finite electronic temperature. The use of a fictitious, finite electronic temperature is a trick routinely employed to aid numerical

convergence in standard DFT calculations of metals.

We have therefore reduced the problem to solving the complex Helmholtz equation, which by itself provides a special challenge due to the fact that the operator is non-Hermitian. Typical iterative methods for solving linear systems, like the conjugate-gradient algorithm, fail for non-Hermitian matrices. The complex Helmholtz equation is an important problem, arising frequently in the context of electrodynamics. Several iterative methods have been developed for the special class of complex symmetric matrices, into which complex Helmholtz operators fall[62, 63, 64]. For the present tests, the biconjugate-gradient algorithm, specialized to complex symmetric matrices[62], has been employed to solve Eq.(4.24).

Solving the discretized version of Eq.(4.24) in real space suffers from another source of inaccuracy, beyond that introduced by the fitting of $\hat{K}(q)$ by $\tilde{K}(q)$. A discretized version of the Laplacian operator must be employed, which, in reciprocal space, deviates from the exact q^2 behavior. The two sources of error can be easily separated for the purposes of numerical tests. To measure the error due to the approximation of $\hat{K}(q)$ by $\tilde{K}(q)$ alone, the kernel energy $T_K[\rho]$ can be computed with the reciprocal space convolution method, Eq.(4.14), but using $\tilde{K}(q)$ instead of $\hat{K}(q)$. The error due to the use of the discretized Laplacian, in addition to the fitting error, is present when the full real space evaluation method is used. We found that the error due to a fourth-order discrete Laplacian operator was negligible compared to the error introduced by the fitting of $\hat{K}$, and throughout the numerical tests this fourth-order Laplacian was employed. The kinetic energy evaluated with the present real-space method is denoted by $T_{\tilde{K}}[\rho]$, while the kinetic energy evaluated in reciprocal space

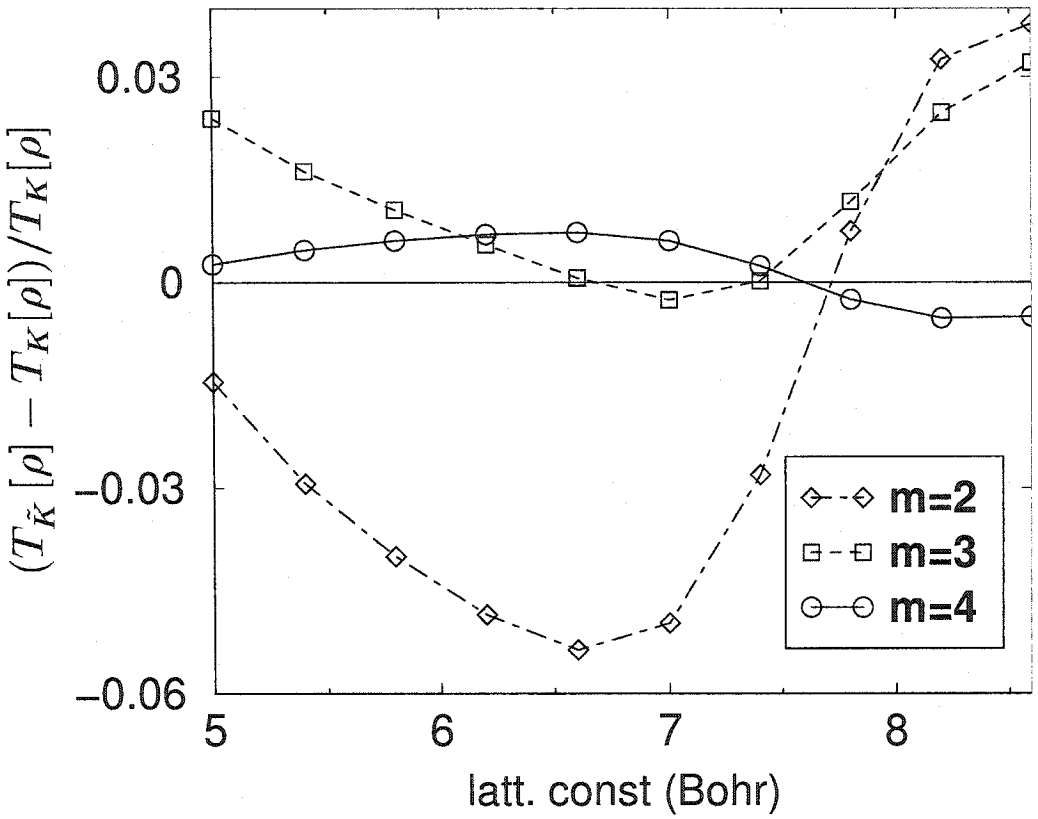

Figure 4.3: The fractional deviation of $T_{\tilde{K}}[\rho]$ from $T_K[\rho]$ for densities $\rho(\mathbf{r})$ obtained from bulk Al calculations (see text).

with $\hat{K}(q)$ is denoted by $T_K[\rho]$. The kinetic energies $T_K[\rho]$ and $T_{\tilde{K}}[\rho]$ are evaluated and compared for a realistic set of charge densities $\rho(\mathbf{r})$ in Fig. 4.3. The $T_K[\rho]$ used is due to Wang et al.[3], and has parameters $\{\alpha, \beta\} = \frac{5}{6} \pm \frac{\sqrt{5}}{6}$, and for $T_{\tilde{K}}[\rho]$, successive fitting orders $m = 2$, 3, and 4 were tried. The fitting coefficients for $m = 4$ are given in Table 4.1. The charge densities considered are generated by minimizing the total energy:

$$E_{tot} = T_{TF}[\rho] + T_{vW}[\rho] + T_K[\rho] + E_H[\rho] + E_{ion}[\rho] + E_{XC}[\rho] \qquad (4.25)$$

(where E_H, E_{XC}, and E_{ion} are the Hartree, exchange-correlation, and electron-ion

Table 4.1: Optimized fitting parameters for $\tilde{K}(q)$ with an order-eight ($m = 4$) rational function of Eq.(4.21). The parameters with even indices $j = 2, 4$ are complex conjugates of the ones given: $P_2 = P_1^*$, $P_4 = P_3^*$, $Q_2 = Q_1^*$, and $Q_4 = Q_3^*$.

	$j = 1$	$j = 3$
P_j	$0.026696 + i0.145493$	$-0.826696 + i0.691930$
Q_j	$-0.818245 - i0.370856$	$0.343051 - i0.689646$

interaction terms), for a bulk fcc aluminum system, with a wide range of lattice constants. Aluminum was represented by the Goodwin-Needs-Heine local pseudopotential[2], and exchange and correlation were treated with the LDA[49]. At each lattice constant, after minimizing the electronic energy with the kernel energy represented by $T_K[\rho]$, the kernel energy is also computed with $T_{\tilde{K}}[\rho]$ and compared to $T_K[\rho]$. As can be seen from Fig. 4.3, the use of more terms in the fitting (higher m) results in a smaller deviation, and at $m = 4$ the accuracy is quite satisfactory (less than 1% deviation for the entire range of lattice constants considered.)

We have performed a different set of tests, in which at each value of the lattice constant the total energy with the kernel energy represented by either $T_K[\rho]$ or $T_{\tilde{K}}[\rho]$ is minimized with respect to the density. In the test discussed earlier the density was fixed to that obtained from minimization of the total energy using $T_K[\rho]$. The equilibrium lattice constant and the bulk modulus obtained from these calculations are given in Table 4.2. As seen from this Table, $m = 4$ offers an acceptable level of accuracy: the values obtained with this approximation to the true kernel are essentially the same as those from the exact kernel, differing only by 0.1% for the lattice constant and by 0.5% for the bulk modulus.

In conclusion, the convolution integrals that appear in the class of kinetic energy

Table 4.2: Lattice constant a_0, in Å, and bulk modulus B, in GPa, for bulk fcc aluminum with the WGC kinetic energy functional[3] with a density-independent kernel, and with $\{\alpha, \beta\} = \frac{5}{6} \pm \frac{\sqrt{5}}{6}$, compared to those values determined with the present real-space method with $m = 2$, 3, and 4. Also shown is a_0 and B from the Kohn-Sham (KS) calculation.

	KS	reciprocal space	$m = 4$	$m = 3$	$m = 2$
a_0(Å)	4.027	4.035	4.030	4.045	4.063
B (GPa)	68.5	71.9	72.3	68.6	68.0

functionals under consideration in this paper cannot be efficiently evaluated directly in real space, and the corresponding inverse integral equations cannot be efficiently solved. The kernels of the convolutions can be approximated by a sum of sub-kernels. This approach makes it possible to evaluate the convolution efficiently in real space by solving the inverse integral equations that correspond to the sub-kernels, which are complex Helmholtz equations. We have demonstrated that this method yields excellent results in numerical tests, in the sense that the error introduced by the real space method is negligible compared to the error inherent in the approximate kinetic energy functionals.

4.1 Real-space evaluation of kinetic energy functionals with density-dependent kernels

Wang, Govind, and Carter[3] (WGC in the following) have developed a class of kinetic energy functionals that include the Thomas-Fermi and von Weizsäcker terms, as well as a term analogous to Eq.(4.5), but with a density-dependent kernel in the

convolution:

$$T_K[\rho] = C_{TF} \times \int f\left(\rho(\mathbf{r})\right) K(\rho(\mathbf{r}), \rho(\mathbf{r}'); |\mathbf{r} - \mathbf{r}'|) g(\rho(\mathbf{r}')) d\mathbf{r} d\mathbf{r}' \tag{4.26}$$

where

$$K(\rho(\mathbf{r}), \rho(\mathbf{r}'); |\mathbf{r} - \mathbf{r}'|) = K(\rho(\mathbf{r}'), \rho(\mathbf{r}); |\mathbf{r} - \mathbf{r}'|).$$

Numerical tests indicate that these kinetic energy functionals are more transferable to systems that deviate significantly from the bulk, like surfaces.

Even by utilizing FFTs, a straightforward evaluation of the convolution in Eq. (4.26) would require $O(N^2)$ operations, where N is proportional to the size of the system. WGC have demonstrated how this convolution can be efficiently, but approximately, evaluated. By Taylor expanding $K(\rho(\mathbf{r}), \rho(\mathbf{r}'); |\mathbf{r} - \mathbf{r}'|)$ with respect to $\rho(\mathbf{r})$ about some chosen average density $\bar{\rho}$, one obtains:

$$\begin{aligned} K(\rho(\mathbf{r}), \rho(\mathbf{r}'); |\mathbf{r} - \mathbf{r}'|) &= K_0(|\mathbf{r} - \mathbf{r}'|) \\ &+ K_1(|\mathbf{r} - \mathbf{r}'|) \left[\Delta\rho(\mathbf{r}) + \Delta\rho(\mathbf{r}')\right] \\ &+ \frac{1}{2} K_{11}(|\mathbf{r} - \mathbf{r}'|) \left[\Delta\rho^2(\mathbf{r}) + \Delta\rho^2(\mathbf{r}')\right] \\ &+ K_{12}(|\mathbf{r} - \mathbf{r}'|) \Delta\rho(\mathbf{r}) \Delta\rho(\mathbf{r}') + \cdots \end{aligned} \tag{4.27}$$

where $\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \bar{\rho}$, and

$$\begin{aligned} K_0(|\mathbf{r}-\mathbf{r}'|) &= K(\bar{\rho},\bar{\rho};|\mathbf{r}-\mathbf{r}'|), \\ K_1(|\mathbf{r}-\mathbf{r}'|) &= \left.\frac{\partial K(\rho(\mathbf{r}),\rho(\mathbf{r}');|\mathbf{r}-\mathbf{r}'|)}{\partial\rho(\mathbf{r})}\right|_{\bar{\rho}}, \\ K_{11}(|\mathbf{r}-\mathbf{r}'|) &= \left.\frac{\partial^2 K(\rho(\mathbf{r}),\rho(\mathbf{r}');|\mathbf{r}-\mathbf{r}'|)}{\partial\rho^2(\mathbf{r})}\right|_{\bar{\rho}}, \\ K_{12}(|\mathbf{r}-\mathbf{r}'|) &= \left.\frac{\partial^2 K(\rho(\mathbf{r}),\rho(\mathbf{r}');|\mathbf{r}-\mathbf{r}'|)}{\partial\rho(\mathbf{r})\rho(\mathbf{r}')}\right|_{\bar{\rho}}, \cdots \end{aligned} \tag{4.28}$$

Then Eq.(4.26) can be evaluated as a sum of separate convolutions with kernels K_0, K_1, etc. WGC also demonstrated that only a few terms of the expansion in Eq.(4.27) are necessary to evaluate the convolution accurately for physical systems.

The present real space method is clearly applicable to such kinetic energy functionals, provided $K_0(\bar{q})$, $K_1(\bar{q})$, etc., where $\bar{q} = k/(2\bar{k}_F)$ and $\bar{k}_F = (3\pi^2\bar{\rho})^{1/3}$, can be fit well by functions of the form of Eq.(4.20). For the kernels other than K_0, a fitting form slightly different than Eq.(4.20) is necessary, because for $\bar{q} \to \infty$, $K_0(\bar{q})$ approaches a constant, as does the rational form Eq. (4.20), but the higher order kernels decay as $\bar{q}^{-2}$. Thus they are fit with the modified form:

$$\frac{N_2\bar{q}^2 + \cdots + N_{2(m-1)}\bar{q}^{2(m-1)}}{D_0 + D_2\bar{q}^2 + \cdots + D_{2m}\bar{q}^{2m}}. \tag{4.29}$$

Functions of this form decompose into the following terms:

$$\tilde{K}(\bar{q}) = \sum_{j=1}^{m} \frac{R_j}{\bar{q}^2 + S_j}. \tag{4.30}$$

Now the $V_j(\mathbf{r})$ are obtained by solving a modified form of Eq. (4.24):

$$\left[-\frac{1}{(2\bar{k}_F)^2}\nabla^2 + S_j\right] V_j(\mathbf{r}) = R_j f(\mathbf{r}), \tag{4.31}$$

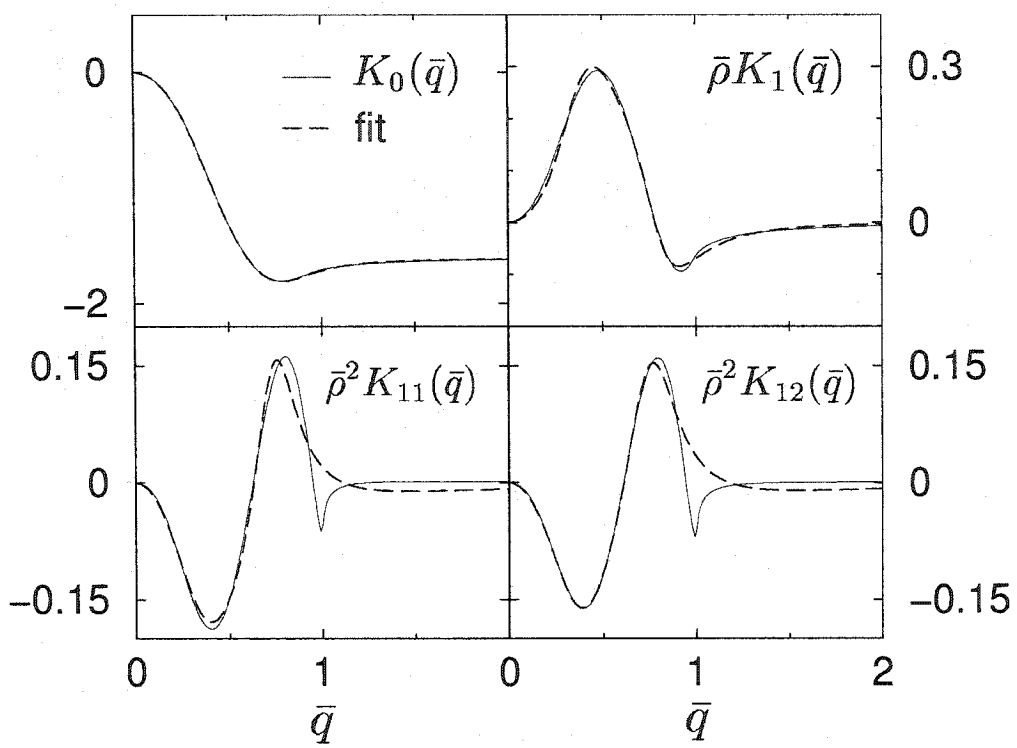

Figure 4.4: The fitting of the various density-dependent kernels of WGC[3] by the forms of Eqs. (4.20) and (4.29) with $m = 4$. Notice the vastly increased scale for K_1, K_{11}, and K_{12}, necessary to show their features, since they are negligible on the scale of K_0.

which is still a complex Helmholtz equation. The kinetic energy functionals of WGC have three parameters, α, β, and γ. Presently, only the case of $\{\alpha, \beta\} = \frac{5}{6} \pm \frac{\sqrt{5}}{6}$, $\gamma = 2.7$, (suggested by favorable numerical tests) is considered. In Fig. 4.4, the best fit to these kernels with a rational function of order eight ($m = 4$) is shown. The quality of the fit is excellent for $K_0(\bar{q})$ and $K_1(\bar{q})$, and reasonable for the second order kernels $K_{11}(\bar{q}), K_{12}(\bar{q})$. WGC have shown that second order terms contribute much smaller parts of $T_K[\rho]$ (which is already a small part of the total energy) than the zeroth and first order terms, and thus a higher fraction of error can be tolerated in these terms. The P_j, Q_j of the decompositions of the fits to K_0 (Eq.(4.21)) and the

Table 4.3: Optimized fitting parameters P_j, Q_j, R_j, and S_j of Eqs. (4.21) and (4.30) for fits to the kernels $K_0(\bar{q})$, $K_1(\bar{q})$, $K_{11}(\bar{q})$, and $K_{12}(\bar{q})$ of the WGC density-dependent kinetic energy functional with $\{\alpha, \beta\} = \frac{5}{6} \pm \frac{\sqrt{5}}{6}$ and $\gamma = 2.7$. The parameters with even indices, $j = 2, 4$, are complex conjugates of the ones given: $X_2 = X_1^*$, and $X_4 = X_3^*$, where $X = P$, Q, R, or S.

		$j=1$	$j=3$
K_0	P_j	$0.108403 + i0.079657$	$-0.908403 + i0.439708$
	Q_j	$-0.470923 - i0.465392$	$0.066051 - i0.259678$
$\bar{\rho}K_1$	R_j	$-0.030515 + i0.015027$	$0.028915 - i0.008817$
	S_j	$-0.597793 - i0.294130$	$-0.087917 - i0.164937$
$\bar{\rho}^2 K_{11}$	R_j	$0.008907 - i0.032841$	$-0.034974 + i0.009116$
	S_j	$-0.537986 - i0.233840$	$-0.041565 - i0.196662$
$\bar{\rho}^2 K_{12}$	R_j	$0.012423 - i0.034421$	$-0.031907 + i0.007392$
	S_j	$-0.511699 - i0.266195$	$-0.034031 - i0.188927$

R_j, S_j of the fits to K_1, K_{11}, and K_{12} (Eq.(4.29)) are given in Table 4.3.

Chapter 5

Efficient computation of forces in electronic structure calculations

A wealth of efficient methods have recently been developed [65, 10, 66, 67, 14, 68, 69, 70, 71, 72, 15, 73, 74, 75, 3] for calculating electronic properties of an extended physical system that require an amount of computation that scales linearly with the size of the system N. This size can be defined to be the number of atoms or the number of valence electrons, or the volume of the system, all of which are linearly related for large condensed systems. In this chapter we present a quasi-linear-scaling ($O(N \log N)$) method for computing the ionic potential, the ionic forces, and the stress tensor in electronic structure calculations. Atomic forces are necessary for the calculation of many physical properties of a system, including the determination of the optimal structure and simulation at a finite temperature. Some linear-scaling methods achieve linear computational scaling for the computation of the energy but not for the forces on all of the ions[15]. Other linear-scaling methods achieve efficient

force calculations by working with a basis of localized functions[70, 73, 74], which are less efficient at representing delocalized electronic states found for example in metallic systems. The present method applies to calculations performed in a periodic parallelepiped supercell, not necessarily orthogonal, in which the electronic charge density $\rho(\mathbf{r})$ is represented on a uniform grid.

As a result of the Hellmann-Feynman[76, 77] theorem, the force on the pth ion (within the Born-Oppenheimer approximation) is given by the sum of the partial derivatives of the ion-ion energy (the Ewald energy) and the electron-ion energy with respect to the atomic coordinates:

$$\mathbf{F}_p = \mathbf{F}_p^{Ewald} + \mathbf{F}_p^{e-i} = -\frac{\partial E^{Ewald}}{\partial \mathbf{t}_p} - \frac{\partial E^{e-i}}{\partial \mathbf{t}_p} \tag{5.1}$$

where

$$E^{e-i} \equiv \int_{cell} \rho(\mathbf{r}) V^{ion}(\mathbf{r}) d\mathbf{r} \tag{5.2}$$

and the ionic potential is defined as:

$$V^{ion}(\mathbf{r}) = \sum_{\mathbf{R}} \sum_{p=1}^{N_{at}} V^{psp}(\mathbf{r} - \mathbf{t}_p - \mathbf{R}) \tag{5.3}$$

where here the $\mathbf{t}_p$ are the atomic positions within the unit cell; $V^{psp}(r)$ is the pseudopotential representing the ions; and the outer sum over $\mathbf{R}$ is over all lattice translation vectors $\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ for all integers n_i, where the $\mathbf{a}_i$ are the lattice vectors defining the unit cell. For simplicity of presentation, we will consider systems that involve only one type of pseudopotential, but the methods presented here generalize readily to systems with multiple pseudopotentials. Furthermore, non-local pseudopotentials can be split into a short-ranged non-local part and a long-ranged local part;

only the long-ranged local part of V^{psp} will be considered here. It should be noted that the present methods make only part of the electronic structure calculation efficient, and in the context of the Kohn-Sham method, the overall electronic structure calculation will still scale as N^3 due to the need to orthogonalize N wavefunctions or to diagonalize an $N \times N$ matrix. Thus the present methods are particularly relevant to the orbital-free density functional methods[75, 3], which deal only with local pseudopotentials and can achieve $O(N)$ scaling for the entire calculation.

The Smooth Particle Mesh Ewald Method[78] (SPME) is an efficient scheme ($O(N \log N)$) for computing the Ewald energy E^{Ewald} and its derivatives with respect to the atomic coordinates, $\partial E^{Ewald}/\partial \mathbf{t}_p$. Here we show how similar ideas can be used to yield efficient methods (also $O(N \log N)$) for determining the ionic potential $V^{ion}(\mathbf{r})$, and the other component of the Born-Oppenheimer forces, $\partial E^{e-i}/\partial \mathbf{t}_p$.

We begin presentation of our method by expressing $V^{ion}(\mathbf{r})$ in terms of the structure factor, $S(\mathbf{r})$:

$$V^{ion}(\mathbf{r}) = \frac{1}{\Omega} \int V^{psp}(\mathbf{r} - \mathbf{r}') S(\mathbf{r}') d\mathbf{r}' \tag{5.4}$$

where

$$S(\mathbf{r}) = \Omega \sum_{\mathbf{R}} \sum_{p=1}^{N_{at}} \delta(\mathbf{r} - \mathbf{t}_p - \mathbf{R}) \tag{5.5}$$

and $\Omega = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$ is the unit cell volume.

In reciprocal space, the expression for V^{ion} becomes:

$$\begin{aligned} \tilde{V}^{ion}_{\mathbf{Q}} &\equiv \frac{1}{\Omega} \int_{cell} V^{ion}(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} d\mathbf{r} \\ &= \frac{1}{\Omega} \tilde{V}^{psp}(\mathbf{Q}) \tilde{S}_{\mathbf{Q}} \end{aligned} \tag{5.6}$$

where

$$\tilde{V}^{psp}(\mathbf{Q}) \equiv \int V^{psp}(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} d\mathbf{r} \tag{5.7}$$

and, using Eq. (5.5),

$$\begin{aligned} \tilde{S}_{\mathbf{Q}} &\equiv \frac{1}{\Omega}\int_{cell} S(\mathbf{r}) e^{i\mathbf{Q}\cdot\mathbf{r}} d\mathbf{r} \\ &= \sum_{p=1}^{N_{at}} e^{i\mathbf{Q}\cdot\mathbf{t}_p} \end{aligned} \tag{5.8}$$

Inversely, we have:

$$V^{ion}(\mathbf{r}) = \sum_{\mathbf{Q}} \tilde{V}^{ion}_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{r}} \tag{5.9}$$

where the $\mathbf{Q}$-sum ranges over all integer multiples of the reciprocal lattice vectors $\mathbf{b}_i$ defined by $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$.

For the purposes of the present work, we assume that $\rho(\mathbf{r})$ is represented on a $N_1 \times N_2 \times N_3$ grid of points $\mathbf{r}_{l_1 l_2 l_3} = \frac{l_1}{N_1}\mathbf{a}_1 + \frac{l_2}{N_2}\mathbf{a}_2 + \frac{l_3}{N_3}\mathbf{a}_3$, with $l_i = 0, ..., N_i - 1$. We will denote quantities such as $\rho(\mathbf{r}_{l_1 l_2 l_3})$ as $\rho(l_1, l_2, l_3)$. The Fourier transform of $\rho(\mathbf{r})$ is approximated by:

$$\tilde{\rho}(m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3) \simeq \mathcal{F}[\rho(l_1, l_2, l_3)] \tag{5.10}$$

where $\mathcal{F}$ is the discrete Fourier transform:

$$\begin{aligned} \mathcal{F}[f(l_1, l_2, l_3)] &= \frac{1}{N}\sum_{l_1=0}^{N_1-1}\sum_{l_2=0}^{N_2-1}\sum_{l_3=0}^{N_3-1} f(l_1, l_2, l_3) \\ &\times \; e^{2\pi i\left(\frac{l_1 m_1}{N_1} + \frac{l_2 m_2}{N_2} + \frac{l_3 m_3}{N_3}\right)} \end{aligned} \tag{5.11}$$

where $N \equiv N_1 N_2 N_3$. Its inverse, $\mathcal{F}^{-1}$, is given by:

$$\begin{aligned} \mathcal{F}^{-1}[\tilde{f}(m_1, m_2, m_3)] &= \sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1}\sum_{m_3=0}^{N_3-1} \tilde{f}(m_1, m_2, m_3) \\ &\times \; e^{-2\pi i\left(\frac{l_1 m_1}{N_1} + \frac{l_2 m_2}{N_2} + \frac{l_3 m_3}{N_3}\right)} \end{aligned} \tag{5.12}$$

Using Eq. (5.6), and transforming back to the real space grid, we can determine $V^{ion}(\mathbf{r})$ via:

$$V^{ion}(l_1, l_2, l_3) = \frac{1}{\Omega}\mathcal{F}^{-1}\left[\tilde{P}(m_1, m_2, m_3)\tilde{S}(m_1, m_2, m_3)\right] \tag{5.13}$$

where $\tilde{P}(m_1, m_2, m_3)$ is the array given by:

$$\tilde{P}(m_1, m_2, m_3) \equiv \tilde{V}^{psp}(m_1'\mathbf{b}_1 + m_2'\mathbf{b}_2 + m_3'\mathbf{b}_3) \tag{5.14}$$

and $m_i' = m_i$ for $0 \leq m_i \leq N_i/2$ and $m_i' = m_i - N_i$ otherwise. The array $\tilde{S}(m_1, m_2, m_3)$ is given by

$$\tilde{S}(m_1, m_2, m_3) \equiv \tilde{S}(m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3) \tag{5.15}$$

Calculation of the structure factor via Eq. (5.8) will scale with the square of the system size, because it needs to be computed at every reciprocal space grid point (m_1, m_2, m_3), and at each of these points a sum must be performed over each atom in the system. However, Essmann et al., as part of their efficient Smooth Particle Mesh Ewald method[78], provide an elegant method for computing the structure factor efficiently (albeit approximately), requiring an amount of computation that scales as only $N \log N$. By incorporating this algorithm to compute the structure factor, the present methods for computing both the ionic potential $V^{ion}(\mathbf{r})$ and the electron-ion contribution to the forces $\mathbf{F}_p^{e-i}$ achieve the same $O(N \log N)$ quasi-linear scaling.

5.1 The smooth particle mesh Ewald method

Here we summarize the method of Essmann et al. for efficiently computing the structure factor, but refer the reader to the original reference [78] for full details. The crux of the algorithm lies in the approximation of the exponential in Eq. (5.8) with (complex) cardinal B-splines. The nth order cardinal B-spline function, $M_n(x)$, is defined as follows: $M_2(x) = 1 - |x - 1|$ for $0 \leq x < 2$, and $M_2(x) = 0$ otherwise; and the higher-order cardinal B-splines are defined recursively:

$$M_n(x) = \frac{x}{n-1} M_{n-1}(x) + \frac{n-x}{n-1} M_{n-1}(x-1) \tag{5.16}$$

We define the grid coordinates of the pth atom as $u_{ip} \equiv N_i \mathbf{t}_p \cdot \mathbf{b}_i$, or equivalently $\mathbf{t}_p = \frac{u_{1p}}{N_1}\mathbf{a}_1 + \frac{u_{2p}}{N_2}\mathbf{a}_2 + \frac{u_{3p}}{N_3}\mathbf{a}_3$. The structure factor, expressed in terms of the grid coordinates, is:

$$\begin{aligned} \tilde{S}(m_1, m_2, m_3) = &\sum_{p=1}^{N_{at}} \exp\left(2\pi i \frac{m_1}{N_1} u_{1p}\right) \\ &\times \exp\left(2\pi i \frac{m_2}{N_2} u_{2p}\right) \exp\left(2\pi i \frac{m_3}{N_3} u_{3p}\right) \end{aligned} \tag{5.17}$$

The exponentials can be approximated by nth order cardinal B-splines, where n is even, as:

$$\begin{aligned} \exp\left(2\pi i \frac{m_j}{N_j} u_{jp}\right) \simeq b_j(m_j) &\sum_{k=-\infty}^{\infty} M_n(u_{jp} - k) \\ &\times \exp\left(2\pi i \frac{m_j}{N_j} k\right) \end{aligned} \tag{5.18}$$

where $b_j(m_j)$ is:

$$b_j(m_j) = \frac{\exp(2\pi i (n-1) m_j / N_j)}{\left[\sum_{k=0}^{n-2} M_n(k+1) \exp(2\pi i \frac{m_j}{N_j} k)\right]} \tag{5.19}$$

When this B-spline approximation of the exponential is used in Eq. (5.17), it becomes a discrete Fourier transform:

$$\begin{aligned}\tilde{S}(m_1,m_2,m_3) &\simeq \tilde{B}(m_1,m_2,m_3) \\ &\times \sum_{p=1}^{N_{at}} \sum_{k_1,k_2,k_3=-\infty}^{\infty} M_n(u_{1p}-k_1)M_n(u_{2p}-k_2) \\ &\times M_n(u_{3p}-k_3) e^{2\pi i \left(\frac{m_1 k_1}{N_1}+\frac{m_2 k_2}{N_2}+\frac{m_3 k_3}{N_3}\right)} \\ &= N\tilde{B}(m_1,m_2,m_3)\mathcal{F}[Q(l_1,l_2,l_3)] \end{aligned} \tag{5.20}$$

where

$$\tilde{B}(m_1,m_2,m_3) \equiv b_1(m_1)b_2(m_2)b_3(m_3) \tag{5.21}$$

and:

$$\begin{aligned}Q(l_1,l_2,l_3) = \sum_{p=1}^{N_{at}} \sum_{c_1,c_2,c_3=-\infty}^{\infty} & M_n(u_{1p}-l_1-c_1N_1) \\ \times M_n(u_{2p}-l_2-c_2N_2) & M_n(u_{3p}-l_3-c_3N_3)\end{aligned} \tag{5.22}$$

The array $Q(l_1,l_2,l_3)$ can be computed quickly ($O(N^{at})$), because it is only non-zero for sub-cubes of dimension $n\times n\times n$ located near each atom. It is because $\mathcal{F}[Q]$ can be computed with the fast Fourier transform (FFT) (which is performed in $O(N\log N)$ operations) that the structure factor itself can be computed with $O(N\log N)$ operations.

5.2 The present methods

It is now easily seen how the structure factor algorithm provided by the SPME method can be used to yield an efficient method for computing $V^{ion}(l_1,l_2,l_3)$. By

substituting Eq. (5.20) into Eq. (5.13), we obtain:

$$V^{ion}(l_1,l_2,l_3) = \frac{N}{\Omega}\mathcal{F}^{-1}\Big[\quad \tilde{P}(m_1,m_2,m_3)\tilde{B}(m_1,m_2,m_3) \times\mathcal{F}[Q(l_1,l_2,l_3)]\Big] \tag{5.23}$$

In practice, Eq. (5.23) could be used to efficiently compute $V^{ion}(l_1,l_2,l_3)$ with the following algorithm. First $Q(l_1,l_2,l_3)$ is computed via Eq. (5.22). Then $\mathcal{F}[Q]$ is obtained via the FFT. $\mathcal{F}[Q]$ is then multiplied by $\tilde{B}$ and $\tilde{P}$, defined by Eqs. (5.21) and (5.14). Then the inverse FFT of this product is computed and multiplied by N/Ω, yielding the array $V^{ion}(l_1,l_2,l_3)$.

The calculation of the electron-ion contribution to the ionic force, $\partial E^{e-i}/\partial \mathbf{t}_p$ can also be made efficient, and again this comes from expressing $\partial E^{e-i}/\partial \mathbf{t}_p$ in terms of the structure factor. In the discrete variable representation, the expression for the pseudopotential energy, Eq. (5.2), becomes:

$$E^{e-i} \simeq \frac{\Omega}{N}\sum_{l_1,l_2,l_3}\rho(l_1,l_2,l_3)V^{ion}(l_1,l_2,l_3) \tag{5.24}$$

which can be viewed as a dot product of ρ and V^{ion}; we can also evaluate this dot product in reciprocal space, and taking into account that ρ is an array of real numbers, the expression becomes:

$$E^{e-i} \simeq \Omega \sum_{m_1,m_2,m_3}\mathcal{F}\left[\rho(l_1,l_2,l_3)\right]^*\mathcal{F}\left[V^{ion}(l_1,l_2,l_3)\right] \tag{5.25}$$

then, using Eq. (5.23) for V^{ion}, this becomes:

$$E^{e-i} \simeq N \sum_{m_1,m_2,m_3}\mathcal{F}\left[\rho(l_1,l_2,l_3)\right]^*\tilde{P}(m_1,m_2,m_3) \times\tilde{B}(m_1,m_2,m_3)\mathcal{F}[Q(l_1,l_2,l_3)] \tag{5.26}$$

Following this substitution, we now express the dot product in real space again:

$$E^{e-i} \simeq \sum_{l_1,l_2,l_3} Q(l_1,l_2,l_3)\mathcal{F}^{-1}\Big[\mathcal{F}\left[\rho(l_1,l_2,l_3)\right] \times \tilde{P}(m_1,m_2,m_3)^*\tilde{B}(m_1,m_2,m_3)^*\Big] \tag{5.27}$$

Because the only factor that depends on the atomic positions is Q, we can readily differentiate with respect to the atomic positions:

$$\frac{\partial E^{e-i}}{\partial \mathbf{t}_{p\alpha}} = \sum_{l_1,l_2,l_3} \frac{\partial Q}{\partial \mathbf{t}_{p\alpha}}\mathcal{F}^{-1}\Big[\mathcal{F}\left[\rho(l_1,l_2,l_3)\right] \times \tilde{P}(m_1,m_2,m_3)^*\tilde{B}(m_1,m_2,m_3)^*\Big] \tag{5.28}$$

where $\alpha = 1,2,3$ is the vector component of the force. Eqs. (5.23) and (5.28) (and the expression for the stress tensor, Eq. (A.3)) constitute the central results of this chapter.

The partial derivatives $\partial Q/\partial \mathbf{t}_{p\alpha}$ can be evaluated readily with the definition of Q, Eq. (5.22), and with the aid of the following B-spline identity:

$$\frac{d}{dx}M_n(x) = M_{n-1}(x) - M_{n-1}(x-1) \tag{5.29}$$

Computing the partial derivatives of Eq. (5.28) for all of the atoms in the system requires an amount of computation that scales as $N \log N$. As with Q, the partial derivative array $\partial Q/\partial \mathbf{t}_{p\alpha}(l_1,l_2,l_3)$ is only non-zero in sub-cubes near the atomic positions, so computing it requires $O(N)$ computation. The fast Fourier transforms scale as $N \log N$.

The application of Eq. (5.28) for rapid computation of the electron-ion forces $\mathbf{F}^{e-i}$ can proceed algorithmically as follows. The Fourier transform of $\rho(l_1,l_2,l_3)$ is computed; then $\mathcal{F}[\rho]$ is multiplied by $\tilde{P}^*$ and $\tilde{B}^*$. Then the inverse Fourier transform of

this product is found. Then by utilizing Eq. (5.29), the derivatives $\partial Q/\partial \mathbf{t}_{p\alpha}(l_1, l_2, l_3)$ are computed during the summing over the l_i's, yielding $\partial E^{e-i}/\partial \mathbf{t}_{p\alpha}$.

Similar ideas can yield an efficient expression for the computation of the stress tensor. The details of the efficient stress tensor method are covered in section A.

Although these methods have been presented assuming that only one type of pseudopotential (and hence only one type of ion) is present in the system, multiple types are readily treated. Since the ionic potential is a linear function of the pseudopotentials, then with multiple types of pseudopotentials the total ionic potential is a sum of ionic potentials of the different types,

$$V_{tot}^{ion}(l_1, l_2, l_3) = \sum_{\tau} V_{\tau}^{ion}(l_1, l_2, l_3) \tag{5.30}$$

where τ indexes the pseudopotential type. The individual V_{τ}^{ion} are each computed with Eq. (5.23) using the Q-array associated with this set of ions, and the $\tilde{P}$-array associated with this pseudopotential type. Likewise to compute the ionic forces, Eq. (5.28) is used to compute the forces on all ions of a given pseudopotential type, using the Q-array associated with this set of ions, and the $\tilde{P}$-array associated with this pseudopotential type.

Several tests have been performed to examine the accuracy of these methods. The accuracy of the approximate structure factor, Eq. (5.20), increases when the number of grid points N increases, and when the B-spline order n is increased. However, in an electronic structure calculation, one is not simply free to choose the number of grid points for computing $\tilde{S}$. It is clear from Eqs. (5.23) and (5.28) that $\tilde{S}$ and the electronic charge density ρ must exist on the same grid, and typically energy convergence considerations dictate a minimum grid density on which ρ is represented.

So it must be established that for a given grid density, the error incurred by using the approximate $\tilde{S}$ when generating V^{ion} and calculating the forces (i.e. the present methods, Eqs. (5.23) and (5.28)) is not much more than the error that would be present using the same finite number of grid points and the *exact* structure factor S when generating V^{ion} and calculating the forces. In other words, it must be shown that the error in the total energy and forces that comes from using an approximate $\tilde{S}$ is smaller than the error due to using a grid of finite density. Furthermore, in order to establish that these methods are indeed (quasi)linear scaling, it must be demonstrated that for a fixed grid density N/Ω and fixed B-spline order n, the error per atom due to the approximate $\tilde{S}$ does not increase when the system size is increased.

All tests here have been done on systems of aluminum atoms simulated with orbital-free density functional theory[75, 3]. The present methods for generating $V^{ion}(\mathbf{r}_i)$ (i.e. the method of Eq. (5.23)) and $\mathbf{F}^{e-i}$ (Eq. (5.28)) were tested as follows: first, in order to test the error due to the approximate $\tilde{S}$ compared to the error from the rest of the calculation, 32 Al atoms were placed in a cubic box 8.08Å on a side, and displaced randomly by about 0.5Å from their fcc crystalline positions. Then V^{ion} was generated using the exact $\tilde{S}$ (Eq. (5.8)), and the present method (Eq. (5.23)) with B-spline orders $n = 6, 8, 10, 12$, and separate electronic relaxations were done in each of these V^{ion}'s, yielding corresponding total energies. After electronic relaxation, electron-ion forces $\mathbf{F}^{e-i}$ were calculated using derivatives of the exact $\tilde{S}$, and with the present method (Eq. (5.28)). This was done for successively higher grid densities. The results are shown in Fig. 5.1. The relative error between the energy and the

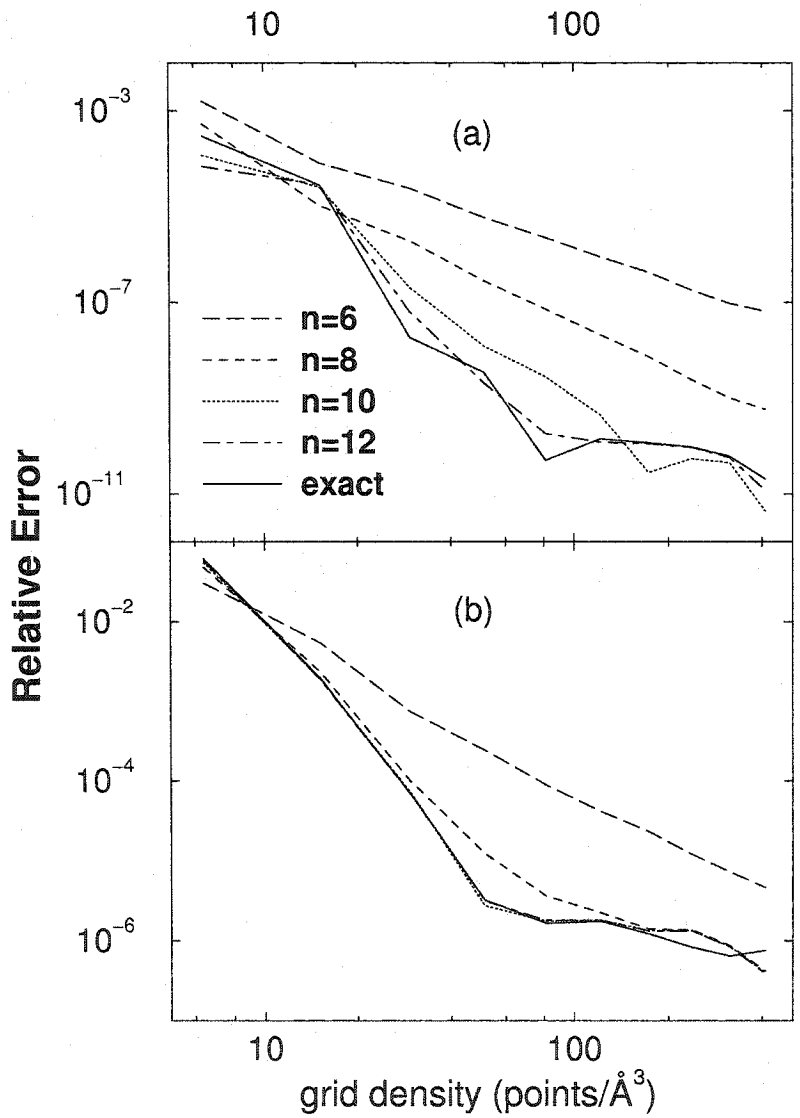

Figure 5.1: (a) Relative error in the energy compared to infinite grid density limit, Eq. (5.31). (b) Relative error in $\mathbf{F}^{e-i}$ compared to the infinite grid limit, Eq. (5.32).

converged energy, as measured by:

$$\frac{|E - E^{\infty}|}{|E^{\infty}|} \tag{5.31}$$

(where E^{∞} is the energy in the limit of infinite grid density) is plotted as a function of grid density for different choices of V^{ion}: V^{ion} generated with the exact method and the present method. E^{∞} is taken to be the total energy evaluated with a grid density of 800 points/$\mathring{A}^3$, a considerably higher density than plotted in Fig. 5.1; thus the deviation from this grid's energy from the true E^{∞} is of a smaller order of magnitude than the energy deviations found at the grid densities explored in Fig. 5.1, making it a suitable energy to use as E^{∞}. It is clear that with a B-spline order of $n = 10$ the error due to the use of the approximate V^{ion} is negligible compared to the error due to the finite grid density. Also plotted is the relative error in the calculated electron-ion forces $\mathbf{F}^{e-i}$. The method used for calculating the force corresponded to the method used to calculate V^{ion}; e.g. the data for the forces calculated with the present method and a 6th order B-spline were done with charge densities relaxed in a V^{ion} generated with the present method with a 6th order B-spline. Thus errors in the forces that were calculated with the present method have some error contribution from using the present method for generating V^{ion}. The relative force error was measured as the fractional root-mean-square deviation of each force component on each atom:

$$\left[\frac{\sum_{i=1}^{N_{at}}\sum_{\alpha}(F_{i\alpha} - F_{i\alpha}^{\infty})^2}{\sum_{i=1}^{N_{at}}\sum_{\alpha}(F_{i\alpha}^{\infty})^2}\right]^{1/2} \tag{5.32}$$

where the $\mathbf{F}_i^{\infty}$'s are the forces in the limit of infinite grid density, in the same sense as expained above for E^{∞}. In this case, already at a B-spline order of $n = 8$ the forces are almost as accurate as those calculated with the exact method, and for $n = 10, 12$ the forces are indistinguishable.

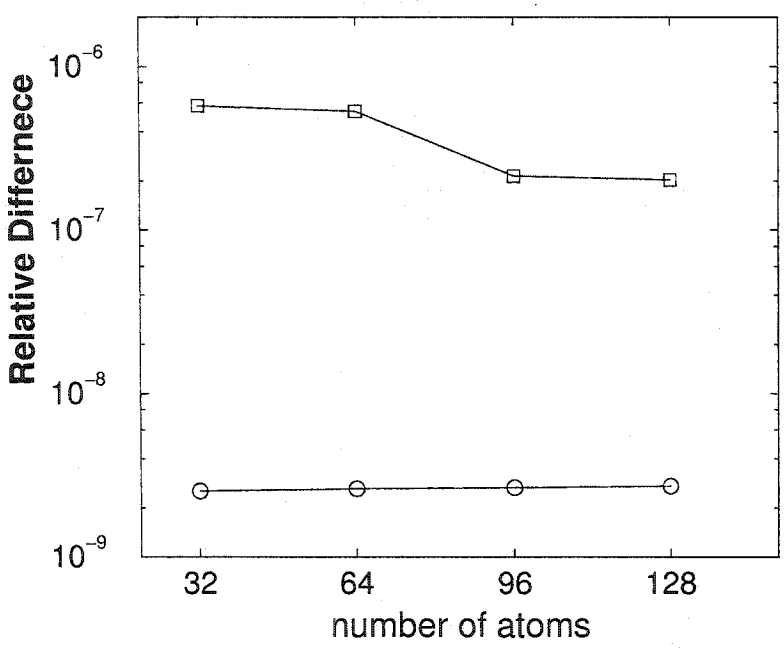

Figure 5.2: The relative difference between the total energy (circles) and forces (squares) calculated with the present methods and with the exact structure factor methods, as a function of system size (see Eqs. (5.31) and (5.32)).

Second, in order to verify the scaling of these methods for different system sizes, four different systems were considered. For systems of 32, 64, 96, and 128 Al atoms, displaced from crystalline positions as before, V^{ion} was generated with the exact $\tilde{S}$, and with an approximate $\tilde{S}$ generated with 8th-order B-splines. The grid density in each case was 237 points/Å^3. After electronic relaxation, forces were calculated using the exact $\tilde{S}$ or with the present method, again corresponding to the method used to generate V^{ion}. The relative error between these energies and forces as a function of system size is plotted in Fig. 5.2. The relative errors are measured as before, with expressions like Eqs. (5.31) and (5.32), but instead of comparing to the energy and forces of the infinite grid density limit, the energy and forces are compared to those calculated using the same grid density and the exact structure factor of Eq. (5.8).

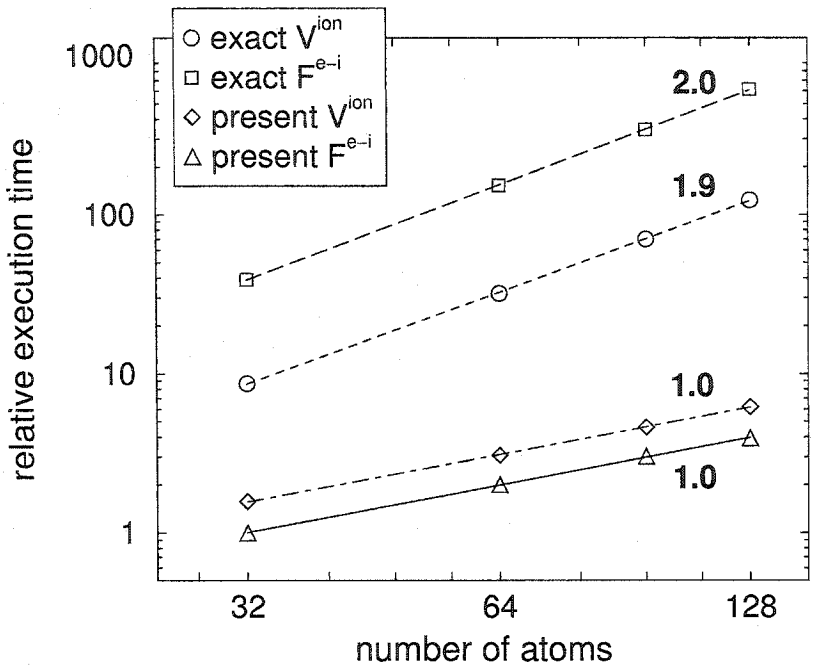

Figure 5.3: The computation time required to generate V^{ion} and evaluate the $\mathbf{F}^{e-i}$'s as a function of number of atoms, for the traditional and present methods. The numbers on each curve represent the exponent of a power law fit.

The relative error in the energy is seen to be constant as a function of system size, and the relative error in the forces is actually seen to decrease slightly with increasing system size. Thus sufficient accuracy can be achieved with the present methods for calculating V^{ion} and the $\mathbf{F}^{e-i}$ using a fixed B-spline order and grid density, confirming that these methods will scale quasi-linearly with system size.

Finally, the time required to generate V^{ion} with the exact and present methods, and to evaluate the electron-ion forces with the exact and the present methods is plotted in Fig. 5.3. One unit of time in the figure corresponded to 1.38 seconds of execution time on a 450 MHz Pentium III processor. The present methods are far superior even for modest numbers of atoms. It is also noted that the present methods

are readily parallelized.

In conclusion, we present accurate and efficient methods for computing the ionic potential and the Born-Oppenheimer forces on atoms by utilizing an approximate form of the structure factor. It has been demonstrated that the errors due to the present methods do not increase with increasing system size. As is typically the case with approximate numerical methods, there is a trade-off between accuracy and efficiency with the present methods. Accuracy can be systematically improved by increasing the grid density or by increasing the B-spline order n, both at the expense of more computing time. However, it was demonstrated that the errors introduced by using the present methods at a B-spline order of $n = 10$ are small compared to errors in other parts of the electronic structure calculation that arise due to the use of a finite grid density. Thus a B-spline order of 10 is recommended as the optimal compromise for simple metals like Al, which was used here as a test case.

Chapter 6

Multiscale simulations in simple metals: methodology

6.1 Introduction

In performing computer simulations of complex physical systems, a premium is placed on accuracy and efficiency. Typically, one of these qualities can be improved at the expense of the other. In recent years, a new approach has emerged that addresses a class of problems in which important small-length-scale phenomena are confined to a small region of the system but can have an impact on the behavior over a much larger scale. A typical case is the tip of a crack, where localized chemical reactions may affect the strength of interatomic bonding, which in turn can influence in a dramatic way the macroscopic mechanical properties of the solid. Such problems fall under the rubric of "multiscale" phenomena, requiring a treatment that addresses important aspects at each scale. The novel feature of this type of simulation is to

use an accurate but computationally demanding method to treat the region of the system in which small-length-scale degrees of freedom are important, and a faster but less accurate method with the small-length-scale physics "coarse-grained", to treat the rest of the system.

Multiscale approaches rely on successfully coupling the two (or more) regions involved, which is referred to as *seamless coupling.* There is no single notion as to what constitutes a seamless coupling, but generally the coupling should be accomplished in such a way that the fictitious boundary between the two regions, which only exists in the coupled simulation and not in the real system, does not introduce any physical consequences. For instance, recently several papers have dealt with the issue of ensuring that phonons are not reflected by the boundary between the two coupling methods[79, 80]. The consummate multiscale method would be one for which the resulting energetics or dynamics is indistinguishable from those that would result from a calculation in which the accurate method was used for the *entire* system. This ideal would be achieved only if the two simulation methods involved were seamlessly matched at the boundary, and further, only if the part of the system treated by the faster, less accurate method was indeed free of important small-scale physics. Another important but obvious characteristic of a good multiscale method is that the computational overhead of performing a coupled simulation is not significant. More specifically, the computation time for the coupled simulation should be on the order of computation time required for the accurate method to treat the small detailed region, since the time required for the less accurate method to treat the rest of the system is typically several orders smaller; nothing is gained if the coupling is so costly that

the coupled method takes as long as using the accurate method to treat the whole system. When the approach requires coupling a quantum mechanical method to a classical method, additional complications arise because of the presence of electronic degrees of freedom in the quantum mechanical region; thus boundary conditions on the electron wavefunctions must be imposed at the interface between the regions. Density functional theory (DFT) provides a significant simplification over more direct quantum mechanical methods in that the calculation of ground state energies and forces requires the minimization of a functional of the electron density $\rho(\mathbf{r})$ only[5]. Thus, in principle boundary conditions need only be imposed on $\rho(\mathbf{r})$. This statement only applies to the formulation of the problem that does not invoke the explicit calculation of electronic wavefunctions (the most common way of implementing DFT actually does involve individual electronic wavefunctions, the so-called Kohn-Sham orbitals[6]). Coupling an approximate DFT calculation that is based on the electronic density alone to a classical interatomic potential should be more straightforward than coupling an orbital-based quantum mechanical method to a classical method.

The present chapter describes a formalism for concurrently coupling a system consisting of two regions, one treated with density functional theory (without invoking electronic wavefunctions) and the other with a classical interatomic potential. Due to the type of approximations involved, the present methods are particularly applicable to simple metallic systems. In section 6.2, other methodologies for coupling multiple simulation approaches and their relevance to the present methods are discussed. In Section 6.3, the formalism of the present class of coupling methods is established. In Section 6.4, details of implementing the methods and achieving efficiency are pre-

sented and some tests of the method are reported in Section 6.5. Finally we conclude in Section 6.6 with a discussion of the results of the tests.

6.2 Background

A large number of concurrent multiscale methods[81, 82, 83] approach the problem of coupling two different simulation methods by writing the energy of the whole system as

$$E[I + II] = E_1[I] + E_2[II] + E^{\text{int}}[I, II] \tag{6.1}$$

where here and throughout the chapter, I refers to the small region where detailed physics are relevant, and II refers to the rest of the system. $E_1[I]$ represents the energy of region I in the absence of region II, $E_2[II]$ represents the energy of region II in the absence of region I, and $E[I+II]$ represents the total energy of the combined system. Eq. (6.1) expresses the total energy of the system is expressed as the energy of region I evaluated with the accurate simulation method 1, plus the energy of region II evaluated with the faster simulation method 2, plus an energy of interaction between the two subsystems. Typically, the crux of a multiscale method lies in its handling of E^{int}. Although tautological, Eq. (6.1) can be rearranged to yield an expression for the interaction energy:

$$E^{\text{int}}[I, II] = E[I + II] - E_1[I] - E_2[II] \tag{6.2}$$

This expression contains no new content, and merely serves to define $E^{\text{int}}[I, II]$, but nevertheless provides direction towards its calculation.

The QM/MM methods are designed to achieve a goal similar to that of the present

method, namely the coupling of a quantum-mechanical simulation with classical potentials, but in the context of covalently bonded organic molecules. In such systems, bonds are localized and typically can be associated with two atoms at either end. The strategy often employed in QM/MM methods to couple quantum mechanics to molecular mechanics is as follows[84]: the system is divided into QM and MM regions with a boundary that cuts across covalent bonds; E_{QM} is evaluated for the QM subsystem, plus additional 'link atoms' placed on the MM side of the severed covalent bonds to mimic the system removed from the QM region; E_{MM} is evaluated for the MM subsystem without the link atoms; and $E_{QM/MM}$ consists of energy terms such as bond-bending terms that are left out of $E_{QM} + E_{MM}$. A similar methodology was developed by Broughton et al.[83] for quantum-classical coupling in silicon, again a covalently-bonded material. Such approaches rely on a somewhat artificial partitioning of the total energy (e.g. into bond-bending and bond-stretching terms), and hence lack a definition that could be readily generalized. But due to the locality of physics in covalently-bonded systems for which QM/MM methods are appropriate, errors introduced at the QM/MM boundary typically do not manifest themselves throughout the system.

In metallic systems, however, the situation is quite different. Bonds are not localized or associated with a distinct pair of atoms. The embedded-atom picture[85, 86] provides a more apt description of the situation. In the embedded atom picture of a simple metallic system, the density of the system is approximately the sum of charge densities of isolated atoms, and the energetic contribution of an individual atom to the system energy is approximately the embedding energy of the atom in a homoge-

neous electron gas. This picture, in various forms[86, 87, 88], has been used to great effect to create classical pair functionals for metals. The success of these potentials in capturing the energetics of simple metallic systems, combined with their foundation on density functional theory arguments, make them ideal candidates for evaluating $E_2[II]$ in the present formalism.

The notions of the embedded-atom method can be extended to describe the energetics of a metallic region (region I) within another metallic region (region II); region I is embedded within region II. The exact nature of the embedding can be formally written in the manner of Eq. (6.1) with density functional theory arguments. We first decide which ions will be associated with region I and which will be in region II, and we will denote those sets of nuclear coordinates by $\mathbf{R}^I$ and $\mathbf{R}^{II}$. We denote the set of all nuclei with $\mathbf{R}^{\text{tot}} \equiv \mathbf{R}^I \cup \mathbf{R}^{II}$. According to the Hohenberg-Kohn theorem, the total system energy, within the Born-Oppenheimer approximation, is given by minimizing a functional of the total charge density:

$$E[\mathbf{R}^{\text{tot}}] = \min_{\rho} E_{\text{DFT}}[\rho^{\text{tot}}; \mathbf{R}^{\text{tot}}] \tag{6.3}$$

In order to be explicit, by $E_{\text{DFT}}[\rho; \{\mathbf{R}\}]$ we mean:

$$\begin{aligned} E_{\text{DFT}}[\rho; \mathbf{R}] \equiv\ & T_{\text{s}}[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] \\ & + \sum_i \int \rho(\mathbf{r}) V_{\text{psp}}(\mathbf{r} - \mathbf{R}_i) \mathrm{d}\mathbf{r} + \sum_{i<j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \end{aligned} \tag{6.4}$$

where T_{s} is the non-interacting kinetic energy functional, E_{H} is the Hartree energy, E_{xc} is the exchange-correlation energy, and V_{psp} is the ionic pseudopotential. Thus E_{DFT} represents the combined electronic and ion-ion (Madelung) energy.

If ρ^{tot} is partitioned into two sub-densities, ρ^I and ρ^{II}, such that $\rho^{\text{tot}} = \rho^I + \rho^{II}$,

then the $E_{\rm DFT}$ can be partitioned:

$$E_{\rm DFT}[\rho^{\rm tot};\mathbf{R}^{\rm tot}] = E_{\rm DFT}[\rho^{I};\mathbf{R}^{I}] + E_{\rm DFT}[\rho^{II};\mathbf{R}^{II}] + E^{\rm int}[\rho^{I},\rho^{II};\mathbf{R}^{I},\mathbf{R}^{II}] \tag{6.5}$$

where $E^{\rm int}$ is defined as in Eq. (6.2). By varying the total energy with respect to ρ^I, one finds that the potential felt by ρ^I is equal to the sum of the potential from region I alone, plus an embedding potential $V_{\rm emb}(\mathbf{r})$ that completely represents the effect of region II upon region I:

$$\begin{aligned} \frac{\delta E_{\rm DFT}[\rho^{\rm tot};\mathbf{R}^{\rm tot}]}{\delta\rho^{I}} &= \frac{\delta E_{\rm DFT}[\rho^{I};\mathbf{R}^{I}]}{\delta\rho^{I}} + V_{\rm emb}(\mathbf{r}) \\ V_{\rm emb}(\mathbf{r}) &\equiv \frac{\delta E^{\rm int}[\rho^{I},\rho^{II};\mathbf{R}^{I},\mathbf{R}^{II}]}{\delta\rho^{I}} \end{aligned} \tag{6.6}$$

By using different approximations for the terms in Eq. (6.6), different coupled methods are obtained. Wesolowski and Warshel[89], building on the formalism of Cortona [90], used this partitioning to describe an efficient DFT method. In their scheme $E[I]$ and $E[II]$ are treated with Kohn-Sham DFT, but $E^{\rm int}$ is evaluated with "orbital-free" density functional theory (OF-DFT), i.e. pure density functional theory in which the Kohn-Sham orbitals are not used and the non-interacting kinetic energy is approximated with an explicit functional of the density[52, 54, 51, 58, 3]. This allows $E[I]$ and $E[II]$ to be alternately minimized in the embedding potential of the other. Govind et al.[82] utilized the partitioning of Eq. (6.5) to obtain a quantum chemistry(QC)/DFT coupled method. There $E_1[I]$ was calculated with QC, $E_2[II]$ with DFT, and again $E^{\rm int}$ was based on OF-DFT. They used this method to explore the electronic structure of molecules adsorbed on metal surfaces. Recently Klüner et al.[91] have extended this formalism to treat adsorbed molecules in their excited

state.

6.3 Formalism

The present method follows in the same vein as the last few examples, to achieve a DFT/classical potential coupling. The general idea of the present methods is as follows. $E_1[I]$ is to be calculated with DFT. $E_2[II]$ is calculated via a classical potential. A choice can be made for the calculation of $E^{\rm int}$, which results in distinct coupling methods, which we examine in detail below.

6.3.1 Classical interaction energy

$E^{\rm int}$ can be calculated using the classical potential:

$$E^{\rm int}[I, II] = E_{\rm cl}[I + II] - E_{\rm cl}[I] - E_{\rm cl}[II] \tag{6.7}$$

Although this interaction energy is intended to represent the same DFT interaction energy that appears in Eq. (6.5), it is not contradictory to use the classical potential to evaluate it, since the classical potential energy, evaluated for a given ionic configuration $\mathbf{R}$, can be viewed as an approximation to the DFT functional that has been minimized with respect to the density; that is:

$$E_{\rm cl}[\mathbf{R}] \simeq \min_{\rho} E_{\rm DFT}[\rho, \mathbf{R}]. \tag{6.8}$$

This choice of interaction energy results in a total energy of:

$$E[\mathbf{R}^{\rm tot}] = E_{\rm cl}[\mathbf{R}^{\rm tot}] - E_{\rm cl}[\mathbf{R}^I] + \min_{\rho^I} E_{\rm DFT}[\rho^I, \mathbf{R}^I] \tag{6.9}$$

In this scheme, the forces on all atoms in region II are identical to forces on corresponding atoms if the classical potential were used for the entire system; i.e. the DFT region has no effect on these atoms. If the cutoff length of classical potential is r_c, then atoms that lie within region I and are farther than r_c from the boundary will experience a force entirely from $E_{\mathrm{DFT}}[I]$; these atoms feel a force no different than corresponding atoms in a DFT calculation of an isolated region I. The force on atoms in region I that are within r_c of the boundary do not come entirely from $E_{\mathrm{DFT}}[I]$, but also have contributions from $E_{\mathrm{cl}}[I+II] - E_{\mathrm{cl}}[I]$. These contributions should in principle be corrections to the surface forces experienced by these atoms from $E_{\mathrm{DFT}}[I]$. Classical potentials have been developed to mimic the energetics, forces, and geometries obtained from DFT calculations of various configurations, including surfaces[92]; such potentials should be particularly apt for the present coupling scheme.

The implementation of this method demands nothing beyond what is required for a DFT calculation and a classical potential calculation. It should be noted, however, that the DFT calculation, $E_{\mathrm{DFT}}[I]$, is a non-periodic calculation, and if OF-DFT is to be used, special considerations may need to be made for the calculation of non-periodic systems[93].

6.3.2 Quantum interaction energy

Alternatively E^{int} can be calculated more accurately with a quantum mechanical method. Although we only represent region II by the coordinates of the ions of region II atoms and calculate the energetics with a classical potential, there is an implicit charge density ρ^{II} associated with $E_{\mathrm{cl}}[\mathbf{R}^{II}]$ via Eq. (6.8). Because of this, we can

consider a more sophisticated coupling scheme where the interaction energy is based on density functional theory. However, in order to compute the interaction energy via DFT when all we know about region II is an approximation of its charge density, the traditional Kohn-Sham scheme of DFT is not suitable. In the Kohn-Sham scheme, we start with a potential and obtain the density and energy of electrons in this potential. Instead, we need a means of calculating the energy of a system of electrons given their density. OF-DFT allows us to do this. Thus we can write down the interaction energy in terms of OF-DFT energy functionals:

$$E^{\text{int}}[I, II] = E_{\text{OF}}[I + II] - E_{\text{OF}}[I] - E_{\text{OF}}[II] \tag{6.10}$$

At first glance this seems like a useless scheme, because if DFT is used to calculate $E^{\text{int}}[I, II]$, we may as well use DFT to calculate $E[I+II]$, and thus no computational expense is saved with the coupled method. But because of the nature of many of the useful OF-DFT functionals, this will turn out not to be the case. If $E^{\text{int}}[I, II]$ is calculated with OF-DFT, for typical approximate kinetic energy functionals the computation in Eq. (6.10) will require a computation time that is on the order of the computation time required to compute $E_{\text{OF}}[I]$, the small subsystem, rather than the time required to compute $E_{\text{OF}}[I + II]$. This is because significant cancellation is implicit in $E_{\text{OF}}[I + II] - E_{\text{OF}}[II]$.

Various approximate kinetic energy functionals exist, differing in accuracy and computational efficiency, and different choices of functional can be made for the evaluation of $E[I]$ and E^{int}, which further increases the number of possible coupling methods. This possibility is important, because the degree to which the computation of E^{int} can be made efficient depends on the choice of kinetic energy functional; and

the functionals that will most efficiently treat E^{int} are not accurate enough to treat $E[I]$.

Regardless of the choice of kinetic energy functional, the evaluation of $E^{\text{int}}[I, II]$ within this coupling scheme requires knowing the electronic density of region II, $\rho^{II}(\mathbf{r})$. In the present method, $\rho^{II}(\mathbf{r})$ is approximated as the sum of atomic charge densities $\rho^{\text{at}}(\mathbf{r})$ centered at the region II nuclei:

$$\rho^{II}(\mathbf{r}) = \sum_i \rho^{\text{at}}(\mathbf{r} - \mathbf{R}_i^{II}) \tag{6.11}$$

This approximation is supported by the embedded-atom picture of simple metallic systems. In principle, $\rho^{\text{at}}(\mathbf{r})$ could be a non-spherically symmetric density. For example, if the arrangement of the region II atoms is always close to the bulk lattice arrangement, then a non-spherically-symmetric charge density that reproduces the bulk charge density when periodically tiled may be more appropriate. However in this chapter $\rho^{\text{at}}(\mathbf{r})$ is always taken to be spherically symmetric.

The density in region II is never explicitly represented in the calculation, but is given a precise form via Eq. (6.11). Thus region II is entirely described by the ionic coordinates $\mathbf{R}^{II}$, and ρ^{II}, the form of which is needed to evaluate E^{int}, is implicitly determined by $\mathbf{R}^{II}$.

The second coupling method is summarized by the expression for the energy as a function of nuclear coordinates within the method:

$$\begin{aligned} E[\mathbf{R}^{\text{tot}}] = E_{\text{cl}}[\mathbf{R}^{II}] + \min_{\rho^I} \Big[& E_{\text{OF}}[\rho^{\text{tot}}; \mathbf{R}^{\text{tot}}] \\ & - E_{\text{OF}}[\rho^{II}; \mathbf{R}^{II}] - E_{\text{OF}}[\rho^I; \mathbf{R}^I] + E_{\text{DFT}}[\rho^I; \mathbf{R}^I] \Big] \end{aligned} \tag{6.12}$$

The last term, $E_{\rm DFT}[\rho^I;\mathbf{R}^I]$, is written as such (and not as $E_{\rm OF}[I]$) to emphasize that we could choose to compute it either with a Kohn-Sham-type calculation or with OF-DFT, but utilizing a more accurate kinetic energy functional than the other OF-DFT terms. This would allow for three distinct levels of accuracy in the calculation: Kohn-Sham accuracy within region I, OF-DFT accuracy for the coupling between regions I and II, and the accuracy of the classical potential in region II. In this case, ρ^I would consist of a set of Kohn-Sham orbitals, $\rho^I(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$, and we would minimize over the ψ_i:

$$\begin{aligned} E[\mathbf{R}^{\rm tot}] &= E_{\rm cl}[\mathbf{R}^{II}] + \min_{\psi_i}\Big[E_{\rm OF}[\rho^{\rm tot};\mathbf{R}^{\rm tot}] \\ - E_{\rm OF}[\rho^{II};\mathbf{R}^{II}] &- E_{\rm OF}[\rho^I;\mathbf{R}^I] + E_{\rm KS}[\psi_i;\mathbf{R}^I]\Big] \qquad (6.13) \\ \rho^{\rm tot} &= \sum\nolimits_i |\psi_i|^2 + \rho^{II}, \qquad (6.14) \\ \rho^I &= \sum\nolimits_i |\psi_i|^2 \qquad (6.15) \end{aligned}$$

However, this interesting possibility is not explored presently; instead we use the same type of OF-DFT calculation for the last four terms of Eq. (6.12). It should be noted that in this case the last two terms cancel, and then the total energy is given by:

$$\begin{aligned} &E[\mathbf{R}^{\rm tot}] = E_{\rm cl}[\mathbf{R}^{II}] + \\ &\min_{\rho^I}\Big[\quad E_{\rm OF}[\rho^{\rm tot};\mathbf{R}^{\rm tot}] - E_{\rm OF}[\rho^{II};\mathbf{R}^{II}]\Big] \qquad (6.16) \end{aligned}$$

6.3.3 Orbital-free DFT and approximate kinetic energy functionals

Orbital-free DFT is a necessary part of the second coupling method, because the electronic structure of region II is represented only in terms of its density via Eq.

(6.11); thus in order to utilize that information, E^{int} must be based only on the charge density and the ionic coordinates. Here we describe some key ideas of OF-DFT.

Hohenberg and Kohn[5] showed that the ground state energy of a system of electrons moving in an external potential is given by minimizing a density functional. Kohn and Sham[6] wrote a useful partitioning of this energy functional:

$$E[\rho] = T_s[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] + \int V_{\text{e-i}}(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \tag{6.17}$$

where T_{s} is the non-interacting kinetic energy functional, E_{H} is the Hartree energy, E_{xc} is the exchange-correlation energy, and $V_{\text{e-i}}$ is the ionic potential. By introducing a set of fictitious non-interacting particles, we can obtain a set of single-particle equations, the Kohn-Sham equations, that allow for the evaluation of $E[\rho]$ with an approximate E_{xc}. The Kohn-Sham method results in an exact evaluation of $T_{\text{s}}[\rho_0]$ for the density ρ_0 that minimizes $E[\rho]$, but the method does not provide a means of evaluating $T_s[\rho]$ for an arbitrary density ρ.

The Kohn-Sham partitioning of the energy, Eq. (6.17), has turned out to be useful beyond the Kohn-Sham method. Because a number of limits of the exact non-interacting kinetic energy functional $T_{\text{s}}[\rho]$ are known[94], $T_{\text{s}}[\rho]$ has been approximated by explicit density functionals constructed to satisfy one or more of these known limits. The orbital-free DFT methods are based on minimizing $E[\rho]$ with T_{s} replaced with an approximate kinetic energy functional.

OF-DFT methods are typically more computationally efficient than the Kohn-Sham method. If the approximate T_{s} can be evaluated with an amount of computation that scales linearly with the system size, then minimizing $E_{\text{OF}}[\rho]$ will require an amount of computation linear in the system size. Since within OF-DFT all terms of

the energy are explicit functionals of the density, there is no need for the introduction of fictitous orbitals, and the density ρ is the only represented variable. Thus the OF-DFT methods also use less memory than the Kohn-Sham method.

Moreover, unlike the Kohn-Sham method, the total energy functional $E_{\mathrm{OF}}[\rho]$ can be evaluated for a given $\rho(\mathbf{r})$. This property makes OF-DFT a suitable candidate for computing $E^{\mathrm{int}}[\rho^I, \rho^{II}]$ in the second coupling method discussed in the previous subsection.

The number of available approximate kinetic energy functionals is sizeable, and the choice of functional is made based on considerations of efficiency and the types of systems to be treated. Because the systems to be considered are simple metals with free-electron-like charge densities, an important property that should be included in the approximate kinetic energy functional is the correct linear response around uniform densities:

$$\hat{\mathcal{F}}\left[\left.\frac{\delta^2 T_{\mathrm{s}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\right|_{\rho_0}\right] = -\frac{1}{\chi_{\mathrm{Lind}}(k)} \tag{6.18}$$

where $\hat{\mathcal{F}}$ is the Fourier transform, and $\chi_{\mathrm{Lind}}(k)$ is the Lindhard response function:

$$\chi_{\mathrm{Lind}}(k) = -\frac{k_F}{\pi^2}\left[\frac{1}{2} + \frac{1-x^2}{4x}\ln\left|\frac{1+x}{1-x}\right|\right] \tag{6.19}$$

with $k_F = (3\pi^2\rho_0)^{1/3}$ and $x = k/2k_F$.

A significant number of efficient functionals have been developed that satisfy the linear response limit for a particular chosen average density[52, 54, 51, 58, 3]. Such functionals often consist of several terms that are local or localized functionals, such as the Thomas-Fermi energy and the von Weizsäcker functionals, plus one or more

convolution terms:

$$T_K[\rho] = \int f(\rho(\mathbf{r}))K(|\mathbf{r}-\mathbf{r}'|)g(\rho(\mathbf{r}'))\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{6.20}$$

By choosing the kernel $K(r)$ properly, the approximate functional can be made to satisfy the correct linear response, Eq. (6.18), around some chosen uniform density ρ_0. Numerical tests indicate that among the current available efficient kinetic energy functionals, the ones of this form are most suitable for simple metallic systems.

However, kinetic energy functionals that contain a convolution part with a long-ranged kernel make the efficient evaluation of $E^{\mathrm{int}}[I, II]$ more difficult; the consequences of this will be discussed in the following section.

6.4 Implementation of coupling

6.4.1 Classical interaction energy

The calculation of the energetics and ionic forces with the first coupling scheme described above involve only DFT calculations and classical potential calculations. However, if an ionic relaxation is to be done on the whole system, there are several possible techniques, the optimal choice depending on the system being relaxed.

If the partitioning of the system into regions 1 and 2 is such that the time required to calculate $E_{\mathrm{DFT}}[I]$ is comparable with the computation time of $E_{\mathrm{cl}}[I+II]$, then ionic relaxation of the total system may be done by using a gradient-based minimizer such as conjugate gradients methods or quasi-Newton methods like BFGS[29]. However, if the system partitioning is such that the time required to evaluate $E_{\mathrm{DFT}}[I]$

is considerably more than that required for the computation of $E_{\rm cl}[I+II]$, as is often the case, then an alternate relaxation scheme may be more efficient. We may relax the total system by using a gradient-based minimizer on the region I system alone, while fully relaxing the region II ions between each ionic update of region I. Gradient-based minimizers like BFGS are only effective if the gradients involved are indeed gradients of an underlying object function. It is not immediately apparent that such is the case in this alternate-relaxation scheme, but we can demonstrate it as follows.

The energy calculated with the first coupling scheme, as a function of all ionic coordinates, is given in Eq. (6.9). A secondary function that only depends on the region I ionic positions can be defined as:

$$E'[\mathbf{R}^I] \equiv \min_{\mathbf{R}^{II}} E[\mathbf{R}^{\rm tot}] \tag{6.21}$$

The useful aspect of E' is that its gradient with respect to $\mathbf{R}_i^I$ can be easily evaluated:

$$\frac{\partial E'}{\partial \mathbf{R}_i^I} = \frac{\partial E[\mathbf{R}^{\rm tot}]}{\partial \mathbf{R}_i^I} + \sum_j \frac{\partial E[\mathbf{R}^{\rm tot}]}{\partial \mathbf{R}_j^{II}} \frac{\partial \mathbf{R}_{j,\rm min}^{II}}{\partial \mathbf{R}_i^I} \tag{6.22}$$

$$= \frac{\partial E[\mathbf{R}^{\rm tot}]}{\partial \mathbf{R}_i^I} \tag{6.23}$$

where the second term on the right of Eq. (6.22) vanishes because all derivatives are evaluated at the minimum of $E[\mathbf{R}^{\rm tot}]$ with respect to $\mathbf{R}^{II}$. This result is analogous to the Hellmann-Feynman theorem[76]. The introduction of the E' function allows for the following relaxation algorithm:

- minimize $E[\mathbf{R}^{\rm tot}]$ with respect to $\mathbf{R}^{II}$ while holding $\mathbf{R}^I$ fixed. This only involves the classical potential, $E_{\rm cl}[\mathbf{R}^{\rm tot}]$.

- Calculate $\min_{\rho^I} E_{\mathrm{DFT}}[\rho^I; \mathbf{R}^I]$ and $E_{\mathrm{cl}}[\mathbf{R}^I]$, and the forces on the region I ions. Using Eq. (6.23) the gradient of E' is obtained.
- Perform a step of a gradient-based minimization of E'.
- Repeat until the system is relaxed.

In this manner, the number of DFT calculations performed is greatly reduced, albeit at the expense of more classical potential calculations.

6.4.2 Quantum interaction energy

Implementation details of the second coupling method require more elaboration. One important point is that ρ^I must be confined to lie within a fixed region Ω^I. This region can (and should) overlap with ρ^{II}. But if ρ^I were not confined to lie within a finite volume Ω^I, it could in principle spread throughout the combined system, and during the course of minimizing with respect to ρ^I (Eq. (6.12)), we would essentially be performing a DFT calculation of the whole system. On the other hand, Ω^I should be chosen large enough so that ρ^I is not artificially confined. In the test systems we have examined, we have found that when increasing the size of Ω^I, a point is reached where the results (e.g. the shape of ρ^I and the forces on the ions) change little. The confinement of ρ^I within Ω^I is illustrated in Fig. 6.1.

The second coupling method maintains efficiency due to the cancellation that occurs when $E_{\mathrm{OF}}^{\mathrm{int}}$ is computed. Consider the computation of a local term of E^{int}; such as the exchange-correlation functional within the local density approximation[5]

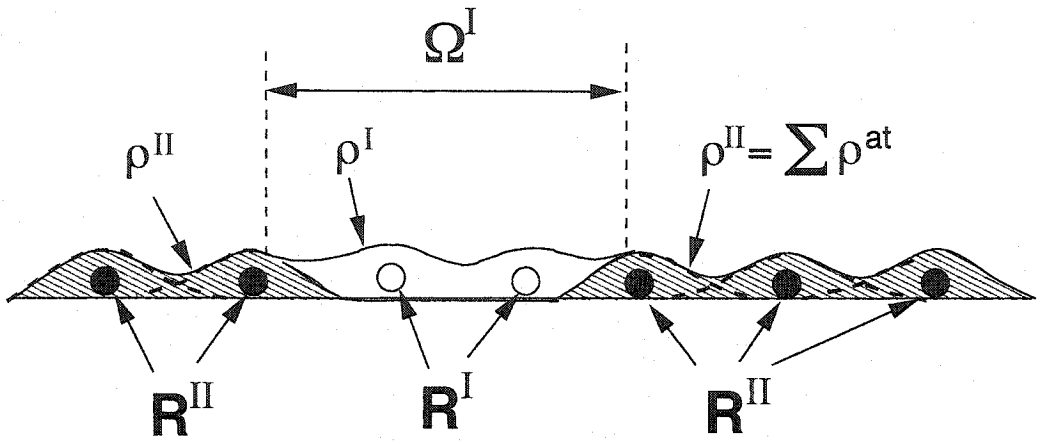

Figure 6.1: An illustration of the partitioning of the system according to the second coupling method.

(LDA):

$$\begin{aligned} E_{\rm xc}^{\rm int} &= \int f_{\rm xc}(\rho^{\rm tot})\mathrm{d}\mathbf{r} - \int_{\Omega^I} f_{\rm xc}(\rho^I)\mathrm{d}\mathbf{r} - \int f_{\rm xc}(\rho^{II})\mathrm{d}\mathbf{r} \\ &= \int_{\Omega^I} \left[f_{\rm xc}(\rho^{\rm tot}) - f_{\rm xc}(\rho^I) - f_{\rm xc}(\rho^{II})\right] \mathrm{d}\mathbf{r} \end{aligned} \tag{6.24}$$

where $f_{\rm xc}(\rho) \equiv \rho\epsilon_{\rm xc}(\rho)$ and we have used the fact that $\rho^{II}(\mathbf{r}) = \rho^{\rm tot}(\mathbf{r})$ for $\mathbf{r} \notin \Omega^I$. Thus calculation of $E_{\rm xc}^{\rm int}$ is an integral over Ω^I and not the entire system, in accord with our efficiency criterion. Any local functional of ρ will obviously be efficient in the same manner. We note that in the case that the same kinetic energy functional is used for the interaction energy and $E_{\rm OF}[I]$ (which is the case for the tests performed in this paper), the cancellation exhibited in Eq. (6.16) occurs. In this case it is wasteful to compute the interaction energy as in Eq. (6.24), and then compute and add on $E_{\rm xc}[\rho^I]$, as it exactly cancels the second term of Eq. (6.24). Instead, we compute:

$$E_{\rm xc}^{\rm int+I} = \int_{\Omega^I} \left[f_{\rm xc}(\rho^{\rm tot}) - f_{\rm xc}(\rho^{II})\right] \mathrm{d}\mathbf{r} \tag{6.25}$$

Similar considerations apply to the other parts of the energy.

We have formulated an efficient method for evaluating the interaction kinetic energy, $T_{\rm s}^{\rm int}$, for more sophisticated functionals with convolution terms such as Eq. (6.20); its derivation is contained in section B.

Attention must be paid to the non-local terms of $E^{\rm int}$. As usual, cancellation occurs between electron-electron, electron-ion, and ion-ion terms that eliminates long-ranged interactions. The Hartree interaction energy is given by:

$$\begin{aligned} E_{\rm H}^{\rm int} &= \frac{1}{2}\int \frac{\rho^{\rm tot}(\mathbf{r})\rho^{\rm tot}(\mathbf{r}') - \rho^{I}(\mathbf{r})\rho^{I}(\mathbf{r}') - \rho^{II}(\mathbf{r})\rho^{II}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}{\rm d}\mathbf{r}{\rm d}\mathbf{r}' \\ &= \int \frac{\rho^{I}(\mathbf{r})\rho^{II}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}{\rm d}\mathbf{r}{\rm d}\mathbf{r}' \\ &= \int_{\Omega^I} \rho^{I}(\mathbf{r}) \sum_i V_{\rm H}^{\rm at}(\mathbf{r}-\mathbf{R}_i^{II}){\rm d}\mathbf{r} \end{aligned} \tag{6.26}$$

where

$$V_{\rm H}^{\rm at}(\mathbf{r}) \equiv \int \frac{\rho^{\rm at}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}{\rm d}\mathbf{r}' \tag{6.27}$$

Similarly the electron-ion interaction energy $E_{\rm e-i}^{\rm int}$ reduces to:

$$\begin{aligned} E_{\rm e-i}^{\rm int} = &\int_{\Omega^I} \rho^{I}(\mathbf{r}) \sum_i V_{\rm psp}(\mathbf{r}-\mathbf{R}_i^{II}){\rm d}\mathbf{r} \\ &+ \int \sum_j \rho^{\rm at}(\mathbf{r}-\mathbf{R}_j^{II}) \sum_i V_{\rm psp}(\mathbf{r}-\mathbf{R}_i^{I}){\rm d}\mathbf{r} \end{aligned} \tag{6.28}$$

where $V_{\rm psp}(r)$ is the pseudopotential representing the ion, and we have used Eq. (6.11) to express $\rho^{II}(\mathbf{r})$ as a sum of $\rho^{\rm at}$. Finally the ion-ion interaction energy is given by:

$$E_{\rm i-i}^{\rm int} = \sum_{i,j} \frac{Z_i Z_j}{|\mathbf{R}_i^I - \mathbf{R}_j^{II}|} \tag{6.29}$$

The combination of all three Coulomb terms can be expessed as:

$$\begin{aligned} E_{\rm H}^{\rm int} + E_{\rm e-i}^{\rm int} + E_{\rm i-i}^{\rm int} &= \int_{\Omega^I} \rho^{I}(\mathbf{r}) \left[\sum_i V_{\rm elec}^{\rm at}(\mathbf{r}-\mathbf{R}_i^{II}) \right] {\rm d}\mathbf{r} \\ &+ \sum_{i,j} \phi_{ij}(\mathbf{R}_i^I - \mathbf{R}_j^{II}) \end{aligned} \tag{6.30}$$

where we have defined:

$$\begin{aligned} V_{\rm elec}^{\rm at}(\mathbf{r}) &\equiv V_{\rm H}^{\rm at}(\mathbf{r}) + V_{\rm psp}(\mathbf{r}), \\ \phi_{ij}(\mathbf{R}_i^I - \mathbf{R}_j^{II}) &\equiv \frac{Z_i Z_j}{|\mathbf{R}_i^I - \mathbf{R}_j^{II}|} \\ &+ \int V_{\rm psp}(\mathbf{r} - \mathbf{R}_i^I)\rho^{\rm at}(\mathbf{r} - \mathbf{R}_j^{II})\mathrm{d}\mathbf{r} \end{aligned} \tag{6.31}$$

Both $V_{\rm elec}^{\rm at}(\mathbf{r})$ and $\phi_{ij}(\mathbf{R})$ are short-ranged functions in which the $1/R$ dependence of the constituent terms cancel.

Within the second coupling method: (1) we minimize the energy with respect to ρ^I, and (2) we calculate the forces on all of the ions and update their position. In order to minimize the energy with respect to ρ^I, the derivative $\delta E^{\rm int}/\delta\rho^I(\mathbf{r})$ needs to be calculated for $\mathbf{r} \in \Omega^I$. This derivative can be evaluated efficiently for the local functionals like $E_{\rm xc}$:

$$\frac{\delta E_{\rm xc}^{\rm int}}{\delta\rho^I(\mathbf{r})} = f'_{\rm xc}(\rho^{\rm tot}) - f'_{\rm xc}(\rho^I) \tag{6.32}$$

where $f'_{\rm xc} = df_{\rm xc}/d\rho$. For the long-ranged Coulombic functionals, the derivative is given by:

$$\frac{\delta}{\delta\rho^I(\mathbf{r})}\left[E_{\rm H}^{\rm int} + E_{\rm e-i}^{\rm int} + E_{\rm i-i}^{\rm int}\right] = \sum_i V_{\rm elec}^{\rm at}(\mathbf{r} - \mathbf{R}_i^{II}) \tag{6.33}$$

Evaluating this combined contribution to $\delta E^{\rm int}/\delta\rho^I$ is a simple matter of evaluating $V_{\rm elec}^{\rm at}$ for region II ions located near the boundary with region I. And so the gradient of the total energy with respect to ρ^I is:

$$\begin{aligned} \frac{\delta E}{\delta\rho^I(\mathbf{r})} = \frac{\delta E_{\rm DFT}[\rho^I; \mathbf{R}^I]}{\delta\rho^I(\mathbf{r})} + \frac{\delta E_{\rm xc}^{\rm int}}{\delta\rho^I(\mathbf{r})} + \frac{\delta T_s^{\rm int}}{\delta\rho^I(\mathbf{r})} \\ + \sum_i V_{\rm elec}^{\rm at}(\mathbf{r} - \mathbf{R}_i^{II}) \end{aligned} \tag{6.34}$$

The calculation of the ionic forces proceeds differently for region I and region II ions. Calculation of the region I ionic forces is facilitated by the Hellmann-Feynman theorem[76]. If we denote the part of the energy (Eq. (6.12)) that is minimized with respect to ρ^I by $G[\rho^I; \mathbf{R}^I, \mathbf{R}^{II}]$:

$$\begin{aligned} G[\rho^I; \mathbf{R}^I, \mathbf{R}^{II}] \equiv & E_{\mathrm{OF}}[I+II] - E_{\mathrm{OF}}[II] \\ & - E_{\mathrm{OF}}[I] + E_{\mathrm{DFT}}[I], \end{aligned} \tag{6.35}$$

then we have, for the second coupling scheme:

$$E[\mathbf{R}^{\mathrm{tot}}] = E^{\mathrm{cl}}[\mathbf{R}^{II}] + \min_{\rho^I} G[\rho^I; \mathbf{R}^I, \mathbf{R}^{II}] \tag{6.36}$$

and when forces on region I ions are computed, the expression simplifies:

$$\frac{\partial E[\mathbf{R}^{\mathrm{tot}}]}{\partial \mathbf{R}_i^I} = \frac{\partial G}{\partial \mathbf{R}_i^I} + \int_{\Omega^I} \frac{\delta G}{\delta \rho^I(\mathbf{r})} \frac{\partial \rho^I_{\mathrm{min}}(\mathbf{r})}{\partial \mathbf{R}_i^I} \mathrm{d}\mathbf{r} \tag{6.37}$$

$$= \frac{\partial G}{\partial \mathbf{R}_i^I} \tag{6.38}$$

where the last term in Eq. (6.37) vanishes because we have minimized G with respect to ρ^I, and so $\delta G/\delta \rho^I|_{\rho^I_{\mathrm{min}}} = 0$. So the forces on the region I ions are determined solely by the terms of G that explicitly depend on $\mathbf{R}^I$; these terms are the electron-ion energy and the ion-ion energy. Using Eq. (6.30), the force on the ith region I ion is given by:

$$\begin{aligned} -\mathbf{F}_i^I &= \frac{\partial E[\mathbf{R}^{\mathrm{tot}}]}{\partial \mathbf{R}_i^I} \\ &= \frac{\partial}{\partial \mathbf{R}_i^I} \left[E_{\mathrm{e-i}}[I] + E_{\mathrm{i-i}}[I] + E^{\mathrm{int}}_{\mathrm{e-i}}[I, II] + E^{\mathrm{int}}_{\mathrm{i-i}}[I, II] \right] \\ &= \frac{\partial}{\partial \mathbf{R}_i^I} \left[E_{\mathrm{e-i}}[I] + E_{\mathrm{i-i}}[I] \right] + \sum_j \nabla \phi_{ij}(\mathbf{R}_i^I - \mathbf{R}_j^{II}) \end{aligned} \tag{6.39}$$

Thus it can be seen from Eq. (6.39) that forces on region I ions are given by the sum of the electron-ion and ion-ion forces present in subsystem 1 alone, and short-ranged interactions with region II ions that are nearby region I.

The forces on the region II ions come mostly from the classical potential, but they have contributions from $E^{\text{int}}[I, II]$ because ρ^{II} is a function of $\mathbf{R}_j^{II}$. Since we have not minimized with respect to ρ^{II}, there is no Hellmann-Feynman simplification when calculating the forces on region II ions, and all terms in the interaction energy contribute. The force on the jth region II ion is given by:

$$\begin{aligned} -\mathbf{F}_j^{II} &= \frac{\partial E[\mathbf{R}^{\text{tot}}]}{\partial \mathbf{R}_j^{II}} \\ &= \frac{\partial E^{\text{cl}}[\mathbf{R}^{II}]}{\partial \mathbf{R}_j^{II}} + \frac{\partial E^{\text{int}}[I, II]}{\partial \mathbf{R}_j^{II}} \end{aligned} \tag{6.40}$$

Local functional parts of E^{int} such as the XC energy will have a contribution to the force given by:

$$\begin{aligned} &\frac{\partial E_{\text{xc}}^{\text{int}}}{\partial \mathbf{R}_j^{II}} = \\ &\quad -\int_{\Omega^I} \nabla \rho^{\text{at}}(\mathbf{r} - \mathbf{R}_j^{II})[f'_{\text{xc}}(\rho^{\text{tot}}) - f'_{\text{xc}}(\rho^{II})]\mathrm{d}\mathbf{r} \end{aligned} \tag{6.41}$$

with analogous expressions for other local contributions that may exist in the kinetic energy functional such as the Thomas-Fermi energy. These local force contributions are only non-zero for region II ions with an atomic density that extends into Ω^I. It is also worth noting that the integral in Eq. (6.41) need not be carried out over all of Ω^I, but only over the intersection of Ω^I with $\rho^{\text{at}}(\mathbf{r} - \mathbf{R}_j^{II})$.

The Coulomb contributions to the region II ionic forces are given by:

$$\frac{\partial}{\partial \mathbf{R}_j^{II}}\left[E_{\mathrm{e-i}}^{\mathrm{int}}+E_{\mathrm{i-i}}^{\mathrm{int}}+E_{\mathrm{H}}^{\mathrm{int}}\right]=-\sum_i \nabla\phi_{ij}(\mathbf{R}_i^I-\mathbf{R}_j^{II}) - \int_{\Omega^I}\rho^I(\mathbf{r})\nabla V_{\mathrm{elec}}^{\mathrm{at}}(\mathbf{r}-\mathbf{R}_j^{II})\mathrm{d}\mathbf{r} \tag{6.42}$$

This contribution also is non-zero only for region II ions near the boundary of Ω^I.

If a more sophisticated kinetic energy functional with a convolution term like Eq. (6.20) is used in E^{int}, then such a term adds considerable complication to the calculation of the forces on region II ions, but these contributions nonetheless die off as we move farther from region I. Thus within the framework of this coupling scheme, the forces on region II ions take the intuitively satisfying form of being equal to the force that arises from the classical potential, plus a correction force for ions near the boundary of Ω^I.

If the partitioning of the system between parts I and II is such that region I requires much longer computation than region II, the second coupling method, like the first, allows for an efficient algorithm for ionic relaxation. We define a different partitioning of the ions as follows: we denote by $\mathbf{R}^{I'}$ the set of region I ions plus all region II ions $\mathbf{R}_j^{II}$ for which the interaction force $\partial E^{\mathrm{int}}/\partial \mathbf{R}_j^{II}$ is not negligible, and we denote by $\mathbf{R}^{II'}$ the rest of the $\mathbf{R}^{II}$ ions. The point is that the forces on the $\mathbf{R}^{II'}$ ions only depend on the classical potential (as seen from Eq. (6.40)), and also that ρ^I does not depend on the $\mathbf{R}^{II'}$ ions (as seen from Eq. (6.34)). Thus the same algorithm for relaxing the system in the first coupling scheme can be used, but with $\mathbf{R}^I$ replaced with $\mathbf{R}^{I'}$, and $\mathbf{R}^{II}$ replaced with $\mathbf{R}^{II'}$. That is, before each relaxation step of the $\mathbf{R}^{I'}$, the $\mathbf{R}^{II'}$ are to be fully relaxed.

6.5 Tests

In order to test the present coupling methods, we have focused on a simple coupled system that is readily analyzed. The system consists of $10 \times 10 \times 10$ cubic unit cells (4 atoms each) of crystalline fcc aluminum. The innermost $2 \times 2 \times 2$ unit cells (32 atoms total) are considered to be region I, and all atoms outside are considered to be in region II. region II, which is treated with the classical potential, is treated as a periodic system in order to remove effects of surfaces from the simulation. So in fact the test system consists of an infinite array of (accurately simulated) 32-atom Al clusters embedded in an Al bulk. Obviously, if there is good coupling between region I and region II, the entire system should simply behave like pure bulk fcc Al, making it easy to evaluate the quality of the coupling. This test system is illustrated in Fig. 6.2.

In all of the present tests, region I is treated with OF-DFT. However, the particular kinetic energy functional used differs among the tests. The Al ions are represented with the Goodwin-Needs-Heine local pseudopotential[2]. For all tests, the classical potential used is the "glue" potential of Ercolessi and Adams[87], which has the embedded-atom method (EAM) form:

$$E[\mathbf{R}_i] = \sum_i F\left[\sum_{i \neq j} \rho^{\mathrm{EAM}}(|\mathbf{R}_{ij}|)\right] + \frac{1}{2}\sum_{i \neq j} \phi(|\mathbf{R}_{ij}|) \tag{6.43}$$

The EAM potential has been scaled both in $\mathbf{r}$ and in energy:

$$\begin{aligned} F[\rho] &\rightarrow \alpha F[\rho], \\ \rho^{\mathrm{EAM}}(\mathbf{R}) &\rightarrow \rho^{\mathrm{EAM}}(\beta\mathbf{R}), \\ \phi(\mathbf{R}) &\rightarrow \alpha\phi(\beta\mathbf{R}) \end{aligned} \tag{6.44}$$

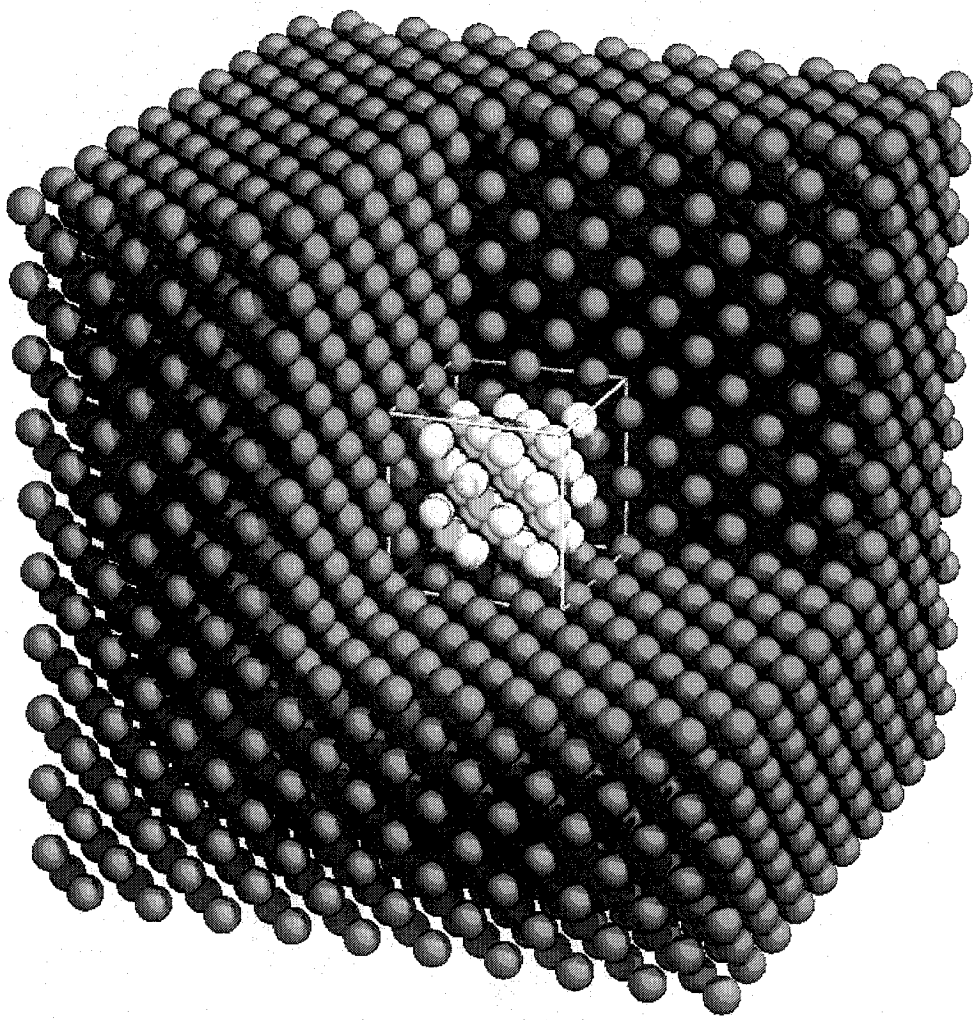

Figure 6.2: A cut-away view of the test system. White atoms belong to region I, and dark atoms belong to region II. For the second method, the region Ω^I is shown by the white cube.

with α and β chosen so that the potential yields precisely the same lattice constant and bulk modulus of fcc Al simulated with OF-DFT employing the particular kinetic energy functional used in that test. This is in accord with the philosophy that the coupled simulation should behave as if the accurate method were used for the entire system. But this procedure is also done so that a "fair" assessment of the coupling itself can be made; we wish to examine errors in the present coupling methods themselves and the approximations involved in them, but not the errors coming from a trivial incompatibility between energy methods. To make the classical potential even more compatible with the OF-DFT method, we could re-determine the form of the classical potential using the method that Ercolessi and Adams originally used to develop their potential[92]: perform a large number of reference energetic calculations of Al using OF-DFT, and find the EAM-type potential that best reproduces these results. This would be a rather involved procedure, so we have chosen to simply scale the potential.

6.5.1 Test of classical ineraction energy method

In the first coupling method (Eq. (6.9)), the energetics of region I was treated with OF-DFT, and the kinetic energy functional used was one developed by Wang et al.[3], with a density-dependent kernel, and parameters $\{\alpha, \beta, \gamma, \rho_*\} = \{5/6 + \sqrt{5}/6, 5/6 - \sqrt{5}/6, 2.7, 0.183\ \text{Å}^{-3}\}$ (in the notation of Ref. [3]), as suggested by their numerical experiments. This functional has six convolution terms of the form of Eq. (6.20). The classical potential was scaled to match the lattice constant (4.033 Å) and bulk modulus (55.7 GPa) of fcc Al obtained by this kinetic energy functional.

The system was initially arranged in the perfect fcc lattice configuration The forces on the region II atoms were identically zero, since they come entirely from $E^{\mathrm{cl}}[I+II]$, which is at a minimum in the initial configuration. However, the forces on the region I atoms are not zero, as the OF-DFT and EAM forces do not perfectly cancel each other. The average magnitude of the force difference per atom between the OF-DFT and EAM calculations of region I was 0.34 eV/Å. These initial forces on the region I atoms are shown in Fig. 6.3(a), with the drawn force vectors scaled so that a force of 1 eV/Å would extend one lattice constant. Then the coupled system was relaxed. After relaxation, it was found that the atomic positions deviated from the correct fcc lattice positions by an average of 0.004 Å per atom. The average deviation of just the region I atoms was 0.076 Å per atom. The atomic deviation is shown in Fig. 6.3(b), in which the relaxed atomic positions for the region I and region II atoms are drawn as white and black balls, respectively, and the correct lattice positions are drawn as gray balls of very slightly smaller radius. Note that only the four (100) layers that include region I are shown. From this diagram it can be seen that in general the relaxed atomic positions deviate from the the perfect lattice positions by bulging out from region I slightly, with the deviation decreasing with increasing distance from region I.

6.5.2 Test of coupling method with quantum interaction energy

The second coupling method was applied to the same simple test system. Here the kinetic energy functional used was different. The functional employed was again from

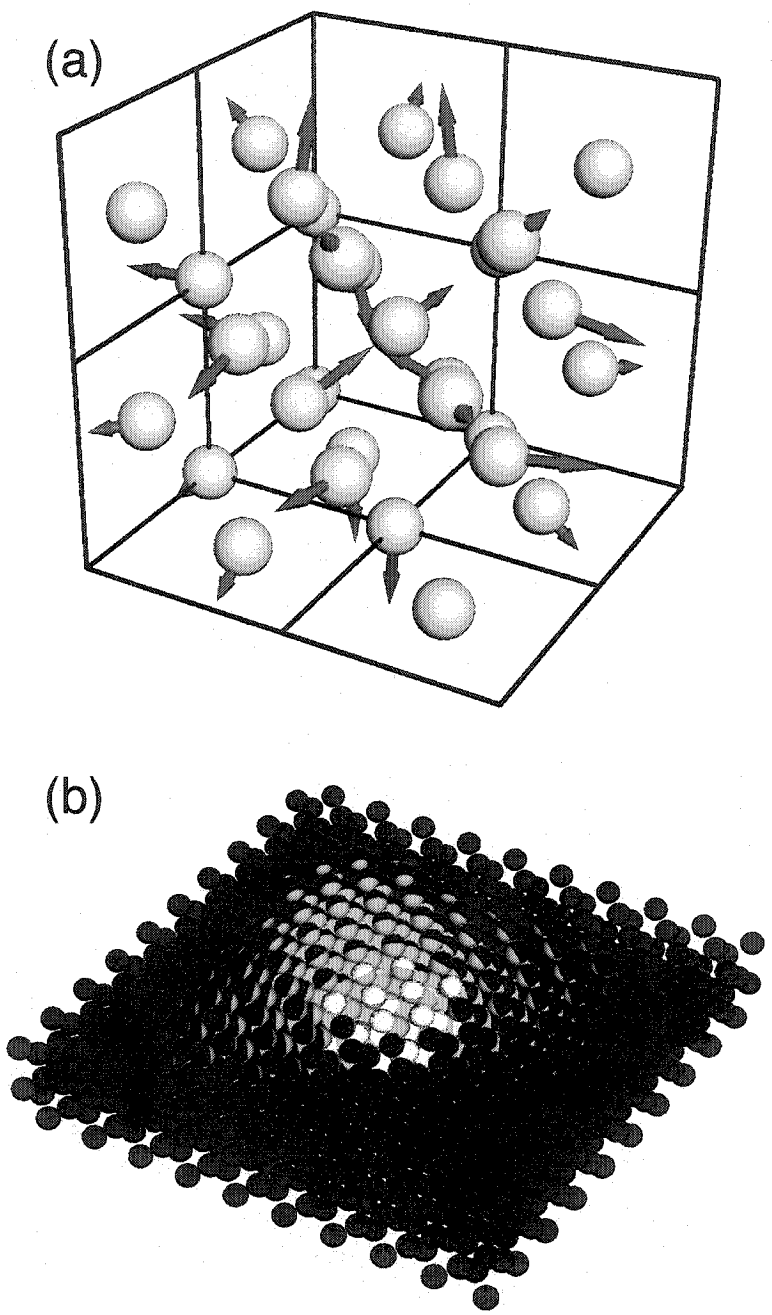

Figure 6.3: Test of coupling method with classical interaction energy. (a) The forces on the region I atoms when the atomic positions are at the perfect lattice positions. The force factors are scaled so that a vector of one lattice constant is 1 eV/Å. (b) The relaxed region I and 2 atomic positions shown in white and black, respectively. The perfect lattice sites are drawn as gray spheres of slightly smaller radius.

Ref. ([3]), but in this case it was a functional with a density-independent kernel, with parameters $\{\alpha, \beta\} = \{5/6 \pm \sqrt{5}/6\}$. A different functional was chosen for this test because of its simpler form: it contains only one convolution term of the form of Eq. (6.20), while the functional used in the test of the first coupling method has six. This makes the evaluation of the kinetic interaction energy, $T_{\mathrm{s}}^{\mathrm{int}}$, simpler. Furthermore, this functional performs well for structures that do not deviate much from the bulk system. This functional was found to be inapplicable to the test of the first coupling method, because in the first coupling method, the $E_{\mathrm{OF}}[I]$ calculation is essentially a cluster calculation; unlike with the second coupling method, there is no embedding potential from region II, and it turns out that the first coupling method, using this simpler functional, the minimization with respect to ρ^I does not converge. For bulk fcc Al, however, this functional performs well, yielding a lattice constant of 4.035 Å and a bulk modulus of 71.9 GPa.

Another aspect of the second coupling method is the choice of atomic density functions ρ^{at} representing ρ^{II} through Eq. (6.11). Two different choices of ρ^{at} were tried. One choice, $\rho^{\mathrm{at,gas}}$, was the valence density from a Kohn-Sham calculation of an isolated Al atom, represented with the same pseudopotential[2]. The other choice, $\rho^{\mathrm{at,cryst}}$ was again a spherically symmetric charge density chosen such that the charge density that results from periodically superposing it on an fcc lattice most closely matches the charge density coming from an OF-DFT calculation of bulk fcc Al. The desired spherically symmetric charge density $\rho^{\mathrm{at,cryst}})(r)$ minimizes:

$$\int_{\Omega} \left[(\delta^{\mathrm{fcc}} * \rho^{\mathrm{at,cryst}})(\mathbf{r}) - \rho^{\mathrm{fcc}}(\mathbf{r}) \right]^2 \mathrm{d}\mathbf{r} \tag{6.45}$$

where $\delta^{\mathrm{fcc}}(\mathbf{r})$ is an infinite fcc lattice of delta functions, $*$ denotes convolution, and Ω

is one unit cell; $\rho^{\rm fcc}(\mathbf{r})$ is the valence charge density of the fcc crystal. In reciprocal space, this becomes:

$$\Omega \sum_{\mathbf{q}} \left| \tilde{S}_{\mathbf{q}} \tilde{\rho}^{\rm at,cryst}(Q) - \tilde{\rho}^{\rm fcc}_{\mathbf{q}} \right|^2 \tag{6.46}$$

where $\tilde{S}_{\mathbf{q}}$ is the structure factor. This is minimized by requiring:

$$\tilde{\rho}^{\rm at,cryst}(Q) = \frac{\langle \tilde{\rho}^{\rm fcc}_{\mathbf{q}} \rangle_Q}{\langle \tilde{S}_{\mathbf{q}} \rangle_Q} \tag{6.47}$$

where $\langle \cdots \rangle_Q$ denotes an averaging over reciprocal lattice vectors $\mathbf{q}$ of length Q. This $\tilde{\rho}^{\rm at,cryst}(Q)$ was then used to construct a radial charge density $\rho^{\rm at,cryst}(r)$ that was commensurate with $\tilde{\rho}^{\rm at,cryst}(Q)$ at the values of Q where the latter was defined. The two charge density choices are shown in Fig. 6.4.

The second coupling method was tested on the same system used to test the first coupling method. With the second method, however, we must choose the form of the atomic density representing ρ^{II}, and the extent of the region Ω^I. With respect to the choice of Ω^I, we have found the following general behavior: if Ω^I is chosen to be too small, then after minimization with respect to ρ^I, there is an excess buildup of charge near the boundary of Ω^I. This in turn results in a net attraction of the region I ions toward the boundary of Ω^I. This is remedied by an increase in the size of Ω^I. When Ω^I is increased further still, the results (the ionic forces, and ρ^I) are found to change only very slightly. We note that regardless of the size of Ω^I, the total density is always found to be continuous at the boundary due to the high energy that the kinetic energy functional assigns to a discontinuity in the density. In Fig. (6.5) we have plotted the total charge density after minimizing with respect to ρ^I, with ρ^{II} given by (a) a superposition of $\rho^{\rm at,gas}$, and (b) a superposition of $\rho^{\rm at,cryst}$.

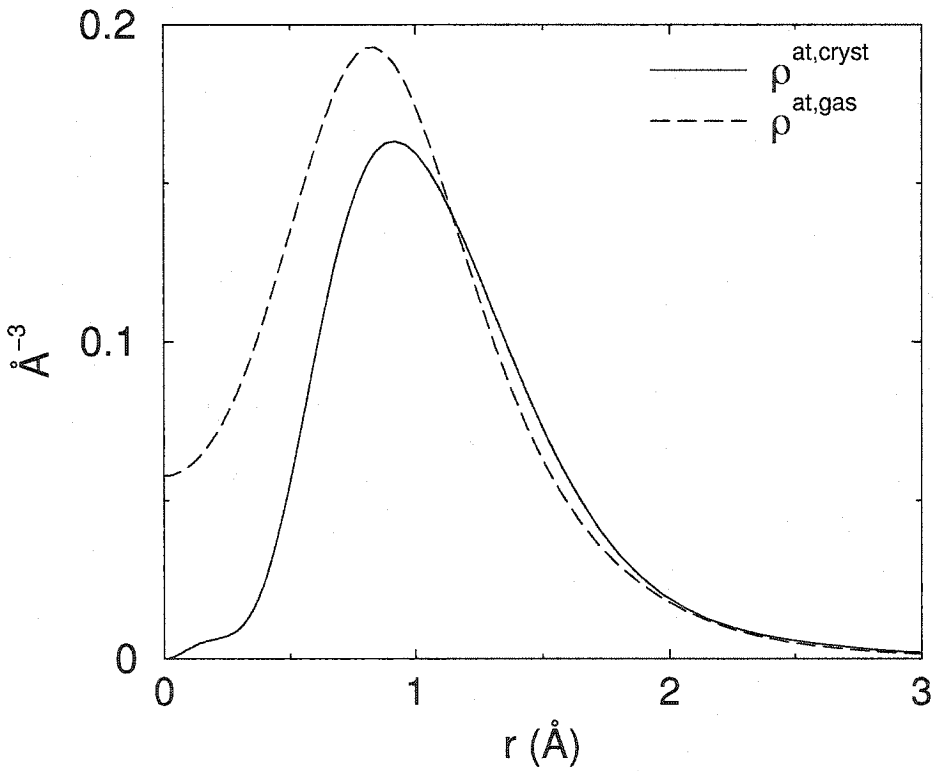

Figure 6.4: The two atomic densities used to represent ρ^{II}.

Table 6.1: Summary of the performance of the two coupling methods, and the two choices for ρ^{at} in the second method. Shown are the maximum force on region I atoms ($\mathbf{F}^I_{\text{max}}$), the maximum force on region II atoms ($\mathbf{F}^{II}_{\text{max}}$), and the maximum displacement from the perfect lattice positions upon relaxation for region I (d^I_{max}) and region II atoms (d^{II}_{max}).

	F^I_{max} (eV/Å)	F^{II}_{max} (eV/Å)	d^I_{max} (Å)	d^{II}_{max} (Å)
method 1	0.45	0	0.12	0.048
method 2, $\rho^{\text{at,gas}}$	0.12	0.27	0.12	0.15
method 2, $\rho^{\text{at,cryst}}$	0.094	0.37	0.217	0.265

The particular slice of the charge density is a (100) plane that passes through one of the central atomic planes of the region I cluster. In (c) and (d) we have plotted the difference between these coupled charge densities and the density of this system when computed entirely with OF-DFT. In general, using $\rho^{\text{at,cryst}}$ results in a better total charge density. From (c) and (d) several observations can be noted. It is clear that the superposition of $\rho^{\text{at,cryst}}$ reproduces the pure OF-DFT crystal charge density better than $\rho^{\text{at,gas}}$. Although for both choices of ρ^{at} the total density is quite close to the OF-DFT charge density for points $\mathbf{r}$ well within Ω^I, the situation is better when $\rho^{\text{at}} = \rho^{\text{at,cryst}}$; the density is matched better at the boundary of Ω^I.

It turns out, however, that the forces on the ions, for both choices of ρ^{at}, are comparable. Also comparable is the amount of deviation from the perfect lattice positions upon atomic relaxation for both choices. The exact numbers for these quantities for both coupling methods and the two choices of ρ^{at} are summarized in Table 6.1.

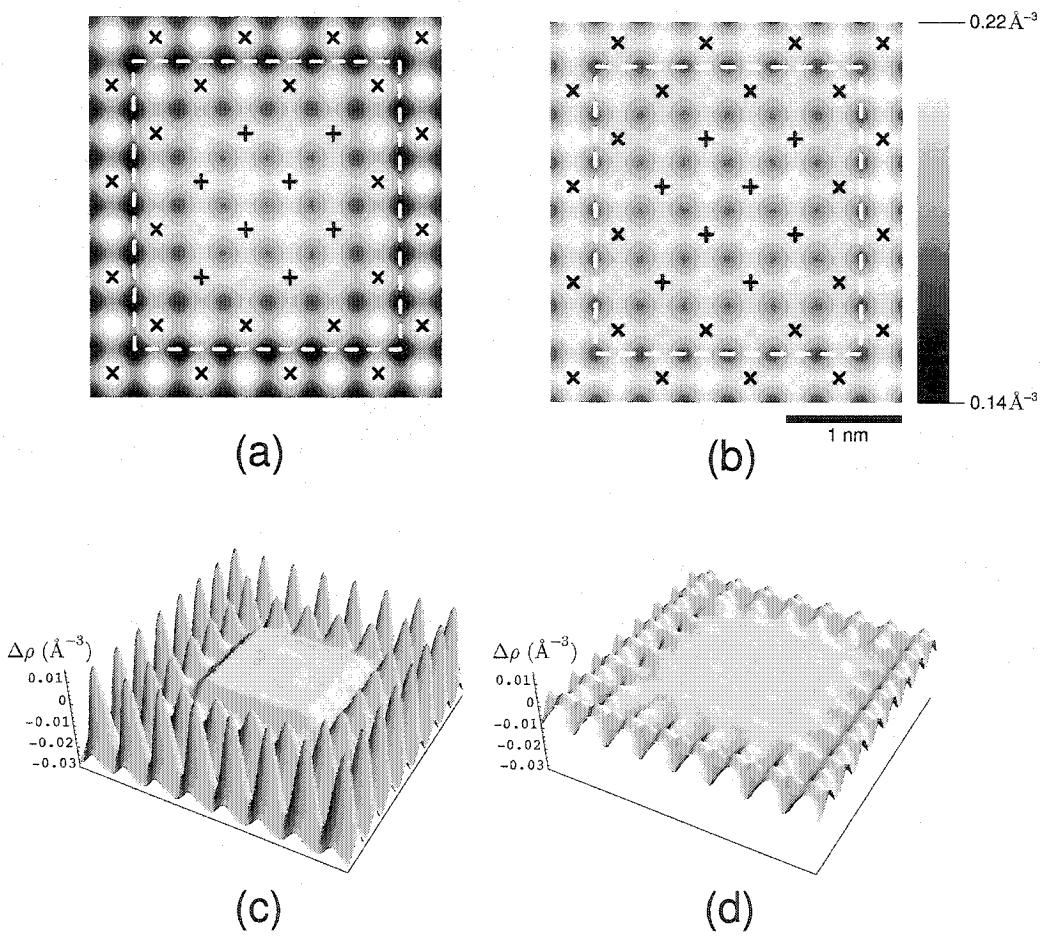

Figure 6.5: Results for the quantum interaction energy method. (a) and (b) are contour plots of ρ^{tot} with ρ^{II} given by superpositions of $\rho^{\text{at,gas}}$ and $\rho^{\text{at,cryst}}$, respectively. The boundary of Ω^I is shown with a dashed line, and the positions of the region I atoms lying in this plane are indicated by $(+)$, and region II atoms by $(\times)$. (c) and (d) show the difference between these two densities and the "correct" density coming from a purely OF-DFT calculation of the whole system.

6.6 Conclusions

We have presented two methods for combining classical and quantum mechanical simulation of simple metals. Both are based on a similar partitioning of the energy of the system, but they differ in how the energy of interaction between the classical and quantum mechanical parts of the system are treated. We have presented numerical schemes that allow both coupling methods to be efficient.

Within the first coupling method, in which the interaction energy is determined from the classical potential, forces in the classical region are fully determined by the classical potential. Forces in the quantum region are determined by both classical and quantum energetics, the quantum energetics dominating well within the quantum region. It is worth noting that if region I contains many atomic species, one need not have a classical potential for each species and their interactions; if these atoms are well within region I, then the potential representing them does not matter at all.

Within the second coupling method, in which the interaction energy is determined via OF-DFT, forces in the classical region are determined mostly by the classical potential, with quantum contributions to atomic forces near the boundary of the regions. Forces in the quantum region are determined fully by quantum energetics. Within the quantum region, the charge density accurately reproduces the correct charge density, and smoothly joins with the implicit density (a sum of atomic densities) of the classical region.

Test results indicate that the second coupling method yields better forces on the atoms in the quantum region than the first method, but that the first method yields better forces for the atoms in the classical region. This may be due to some extent

to the less-accurate OF-DFT method used in the test of the second coupling method. The first coupling method also yielded a better relaxed structure, probably due to its better treatment of forces on the classical atoms. However, unlike the first method, the second coupling method results in an accurate charge density within the quantum mechanical region, allowing for an accurate treatment of physical problems such as the introduction of impurities, where the background density in which the impurities are introduced is important. We also find that a superposition of atomic charge densities can reproduce the actual charge density well for a simple metallic system, given an appropriate choice for the atomic charge density; this allows for a gentle density transition at the boundary between regions.

Clearly an important issue affecting the coupling quality for both methods is the agreement between forces from the DFT methods; within both methods there are atoms whose forces are determined by a combination of quantum and classical energetics, and the more closely the two energetics agree, the better the coupling will be. An improvement in the quality of the coupling might be seen if a classical potential is optimized to reproduce the DFT energetics; this is also in accord with the multiscale philosophy that a coupled simulation should strive to act as if the most accurate method involved is used to simulate the entire system.

The coupling of classical and quantum simulation in metals involves a set of challenges quite different than those that arise in the coupling of covalently-bonded materials and molecules, and hence requires a different set of approaches. We have presented several possible methods for accomplishing this.

Chapter 7

Multiscale simulations in simple metals: applications

In this chapter we apply the coupling method of chapter 6 to several physical systems. This serves to both demonstrate the capabilities and realm of applicability of the method, but additionally it is an exploration of the interesting physics of dislocation-impurity interactions.

7.1 Isolated dislocations in fcc aluminum

The simulation of an isolated dislocation presents a surprisingly difficult challenge. Most of the important properties of a dislocation are related to the structure of its core. The dislocation dissociation, the Peierls barrier, and the interaction of the dislocation with impurities are all related to the core structure. The core of the dislocation produces a long-ranged strain field, and conversely the core is strongly

affected by any modification of its natural strain field. For example, Olmsted *et al*[95] report that the Peierls stress of a dislocation in aluminum is overestimated by a factor of two, due to finite-size effects, when the dislocation is simulated in a cylinder of a large but finite diameter of 180 Å. This example is typical for dislocations, where the property of interest (in this example, the Peierls stress) is of a very small magnitude, but the elastic energy of the long-ranged field is large (infinite!), and does not decrease quickly with distance from the core.

The long-ranged strain field of the dislocation, although important in that it can strongly affect the core, does not carry much information. In fact, one only has to move a short distance from the core before the strain field can be accurately described by (anisotropic) continuum mechanics. As such, dislocations are a prime candidate for study via multiscale simulation.

In fact, a multiscale simulation of dislocations in which the core was treated with a quantum method has been performed before. Woodward and Rao[96] present calculations of isolated dislocations in bcc Mo and Ta. Their method proceeds as follows. First the lattice Green's function (the linear map from small atomic forces on atom i to small displacements of neighboring atoms j) is determined for the material and crystal structure in question. This is found by simulating a small piece of the crystal with DFT, and applying a small force on an atom and observing the relaxed positions of nearby atoms. Then a periodic supercell containing a dislocation is established, and divided into three concentric cylindrical regions centered on the dislocation line, region 1 being the innermost region containing the dislocation, which is surrounded successively by regions 2 then 3. The supercell is simulated with DFT,

and the region 1 atoms are relaxed while holding the remaining atoms fixed. This produces "incompatibility forces" in the region 2 and 3 atoms. The incompatibility forces on the region 2 atoms are used to find displacements via the lattice Green's function for all of the atoms. Then the region 1 must be relaxed again, and so forth, until convergence is achieved. In this manner, region 1 atoms are essentially relaxed under forces from a quantum calculation, while the region 2 atoms respond to the displacements of the region 1 atoms the way the infinite lattice simulated with DFT would, via the lattice Green's function. Stated another way, region 2 acts as a "DFT infinite lattice" boundary condition to the dislocation.

Woodward and Rao's method is an elegant solution to the isolated dislocation problem, but has several conceptual disadvantages. For one, there is still an issue of dealing with the boundary of their quantum mechanical region. They explore two possibilities: 1) terminating region 3 with vacuum, and 2) having the atoms of region 3 extend right up to their periodic counterparts. In either case, there is an artificiality added, and Woodward and Rao simply hope (but convincingly demonstrate, for the systems they study) that it does not affect the regions of more physical significance, regions 1 and 2. Although their method can yield an accurate core structure and Peierls stress, it is not clear how it could yield total energetic properties, such as the Peierls energy. Finally, their method is limited to treating systems that are coupled to a near-perfect lattice, for otherwise the lattice Green's function would not be a translationally invariant, simple object, and the effort of calculating it would overwhelm the effort of simply performing a DFT calculation of the whole coupled system to be ultimately studied.

These issues do not arise in the coupling methods developed in chapter 6. These methods deal very explicitly with the termination of the quantum mechanical region. Furthermore, they are total energy methods, and as such can be used to calculate energetic quantities such as the Peierls energy. And they are not restricted to physical systems in which the classical region is very close to crystalline.

The second of these coupling methods, in which the interaction energy is computed with a quantum mechanical method, is employed to simulate isolated screw and edge dislocations in fcc aluminum. The method is readily generalized to the treatment of systems that are periodic in one direction, as is the case with the line dislocations. The kinetic energy functional of Wang et al.[3], with a density-independent kernel, and parameters $\{\alpha, \beta\} = \{5/6+\sqrt{5}/6, 5/6-\sqrt{5}/6\}$ was employed. The energy of region I and the coupling energy are both computed with orbital-free density functional theory using the density-independent kinetic energy functional described in Sec. (6.5.2). Region II is treated with the Ercolessi-Adams glue potential[92], scaled such that the lattice constant and bulk modulus match that of aluminum simulated with OF-DFT. The Goodwin-Needs-Heine local pseudopotential[2] was used to represent Al in the OF-DFT parts of the calculations.

7.1.1 Screw dislocation

The screw dislocation line lies in the [110]-direction. The simulation consists of a minimal two (110)-planes. The dimensions of the region Ω^I to which the density ρ^I was confined has dimensions of $31.85 \times 31.85 \times 2.85$ Å in the $[1\bar{1}2]$, $[1\bar{1}\bar{1}]$, and [110] directions, respectively. This extends slightly beyond the region I atoms and into the

region II atoms (see Fig. 7.1). The EAM region (region II) extends to a radius of 95 Å, and the atoms between a radius of 85 and 95 Å are fixed in positions determined by the anisotropic elastic strain field of a screw dislocation in a continuous anisotropic material with elastic moduli of OF-DFT-simulated aluminum.

The detailed structure of a dislocation is best displayed with a differential displacement map (DDM). In the case of the screw dislocation, the DDM displays a circle at the position of every column of atoms that extends in the [110]-direction in the undistorted crystal. The circles are placed at the undistorted positions of the column of atoms. Between each neighboring column of atoms is drawn an arrow, the magnitude and direction of which is given by the differential displacement of the two columns of atoms. The differential displacement of two atomic atomic columns is the difference in the displacement of the two columns along the [110]-Burgers direction; and since each column of atoms can be translated by the Burgers vector $\mathbf{b}$ without change, the differential displacement can always be considered to be in the interval $[-b/2, b/2]$. The arrow is scaled such that a differential displacement of $\pm b/2$ extends all the way between the two circles representing the atomic columns.

From Fig. 7.1 it can be seen that the overall structure of the screw dislocation simulated with purely the EAM interatomic potential is quite similar to that of the screw dislocation simulated with the coupling method. Both simulations show a similar dissociation of the dislocation on the $[1\bar{1}\bar{1}]$ plane. The overall similarity is partially due to the closeness in energetics between the EAM potential and OF-DFT over the relatively "safe" realm of configuration space explored by a screw dislocation: the neighborhoods of most atoms in the system are fairly close to the bulk structure.

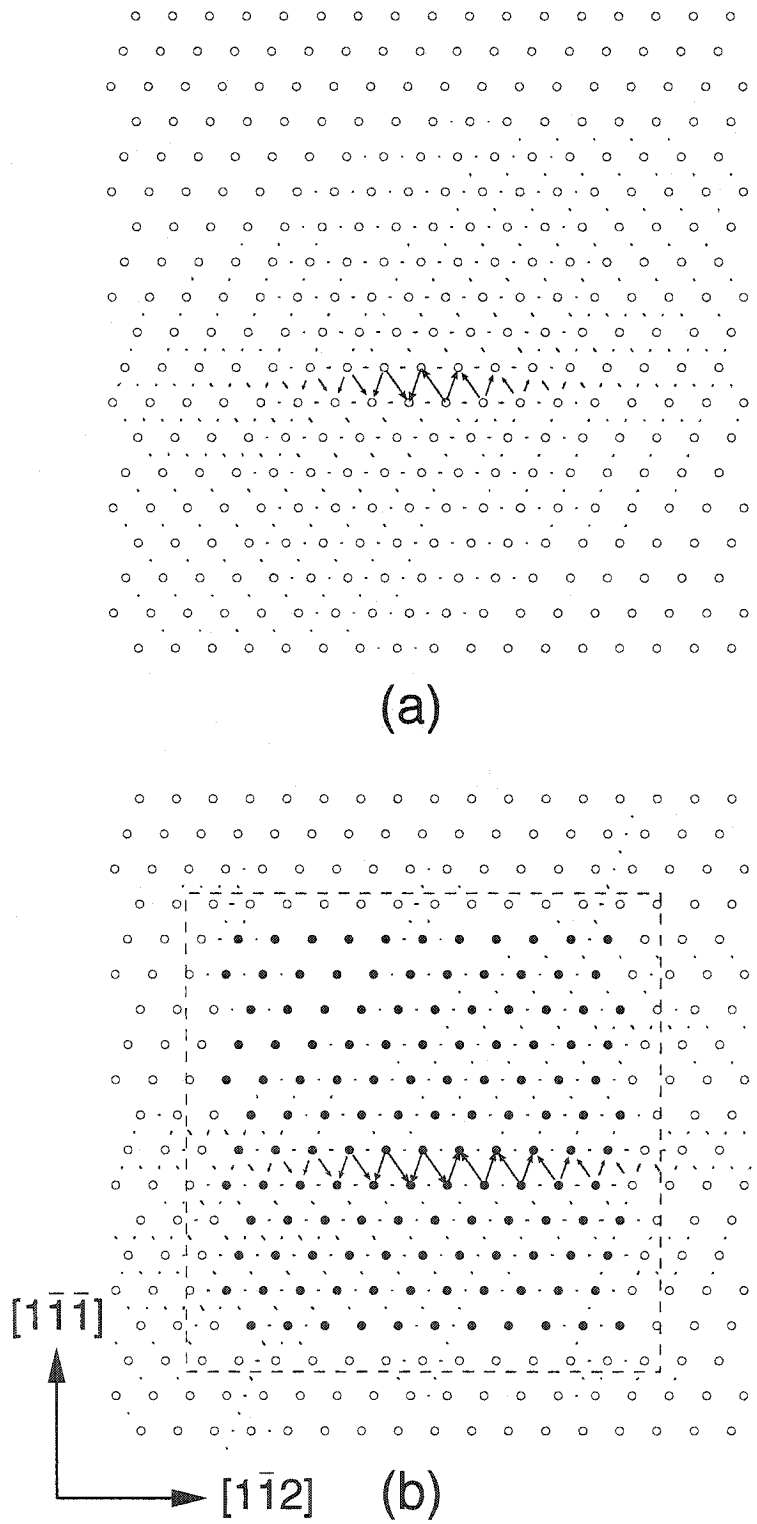

Figure 7.1: Differential displacement maps of a screw dislocation in aluminum (a) simulated with the Ercolessi-Adams interatomic potential, and (b) the quantum-interaction-energy coupling method of chapter 6. The region Ω^I to which the density ρ^I is confined is shown with a dashed line, and the $\mathbf{R}^I$ atoms are displayed as black circles. Note that only a portion of the simulated system is shown.

But also implicit in the similarity is the fact that the disturbance introduced by the coupling itself is fairly small.

Differences between the structures do exist, though. These are more easily revealed by plotting the relative [110]-displacement of the two $(1\bar{1}\bar{1})$ planes that sandwich the dislocation. This is shown in Fig. 7.2. From this figure several key features can be noticed. Most prominently, the screw dislocation is more spread-out in the coupled simulation. If the spread of the dislocation is defined as the distance between points where the relative slip equals 1/4 and 3/4 of the full Burgers-vector slip, then the spread of the dislocation simulated with the EAM potential is 10.1 Å, while the spread of the dislocation simulated with the coupling method is considerably more at 17.6 Å. It can also be seen from the figure that there is a somewhat distinct splitting of the dislocation into two partial dislocations when simulated with the coupling method: slip proceeds quickly across the partials, but between the partials slip occurs less rapidly. Finally, there is a small but strange feature at around -20 Å. This is probably a manifestation of the coupling. It should be noted, however, that the features of Fig. 7.2 should not be considered in too much detail, because the full curve is an interpolation from the fairly coarse displacement data that is available only at each column of atoms.

7.1.2 Edge dislocation

We also simulated the edge dislocation with the quantum interaction coupling method. The edge dislocation line lies in the $[1\bar{1}2]$-direction, with a Burgers vector in the [110] direction. The simulation was periodic along the $[1\bar{1}2]$-direction, with a

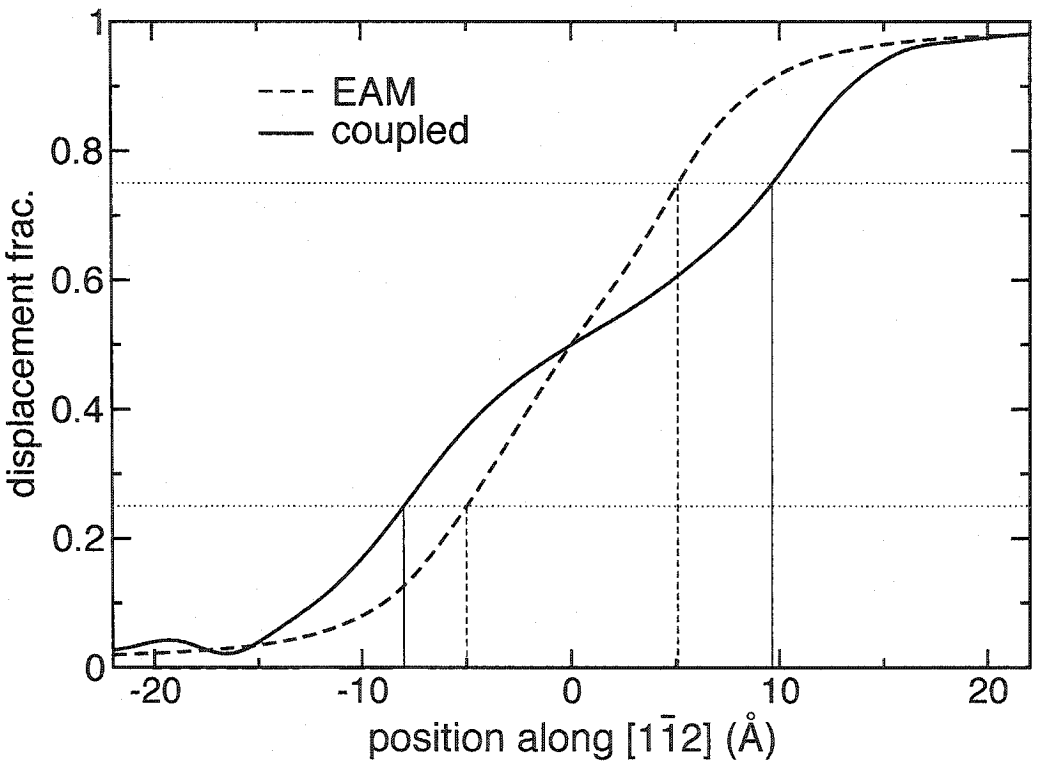

Figure 7.2: The relative displacement along the [110]-direction of the two $(1\bar{1}\bar{1})$-planes that sandwich the screw dislocation, for both the EAM simulation and the OF-DFT/EAM coupled simulation. The locations where the relative displacement equals 1/4 and 3/4 of the full Burgers-vector slip is indicated.

periodicity of $a\sqrt{3/2}$ where a is the lattice constant. The edge dislocation dissociates further than the region Ω^I that contains the density ρ^I had dimensions of $40.48 \times 23.88 \times 4.94$ Å in the [110], $[1\bar{1}\bar{1}]$, and $[1\bar{1}2]$ directions, respectively. Region II again extends to a radius of 95 Å, and the atoms with a radius between 85 and 95 Å are fixed at positions specified by anisotropic elasticity theory.

The displacement maps of the edge dislocation simulated with the EAM potential and the coupling method are shown in Fig. 7.3. In the case of the edge dislocation, the arrows between neighboring $[1\bar{1}2]$-columns of atoms indicate the difference in the displacement along the [110] direction (the x-direction in the figure) between the two columns.

As with the screw dislocation, the differences between the EAM and the coupled simulations of the edge dislocation are more apparent when the relative displacement of the two $(1\bar{1}\bar{1})$-planes above and below the slip plane are plotted (see Fig. 7.4). As with the screw dislocation, the edge dislocation is seen to be more spread out in the coupled simulation compared to the EAM simulation. And again, the dislocation in the coupled simulation shows more characteristics of two distinct partial dislocations when compared to the EAM simulation. The dislocation spreads, measured as with the screw dislocation, are 16.4 Å for the EAM simulation, and 24.7 Å for the coupled simulation. The slight noise in the coupled simulation curve that occurs near ± 20 Å is at the edge of region I, and again is probably a slight manifestation of imperfect coupling between the two regions.

A feature common to both dislocations is that the dissociation is seen to be wider in the coupled simulation compared to the EAM simulation. This suggests a dif-

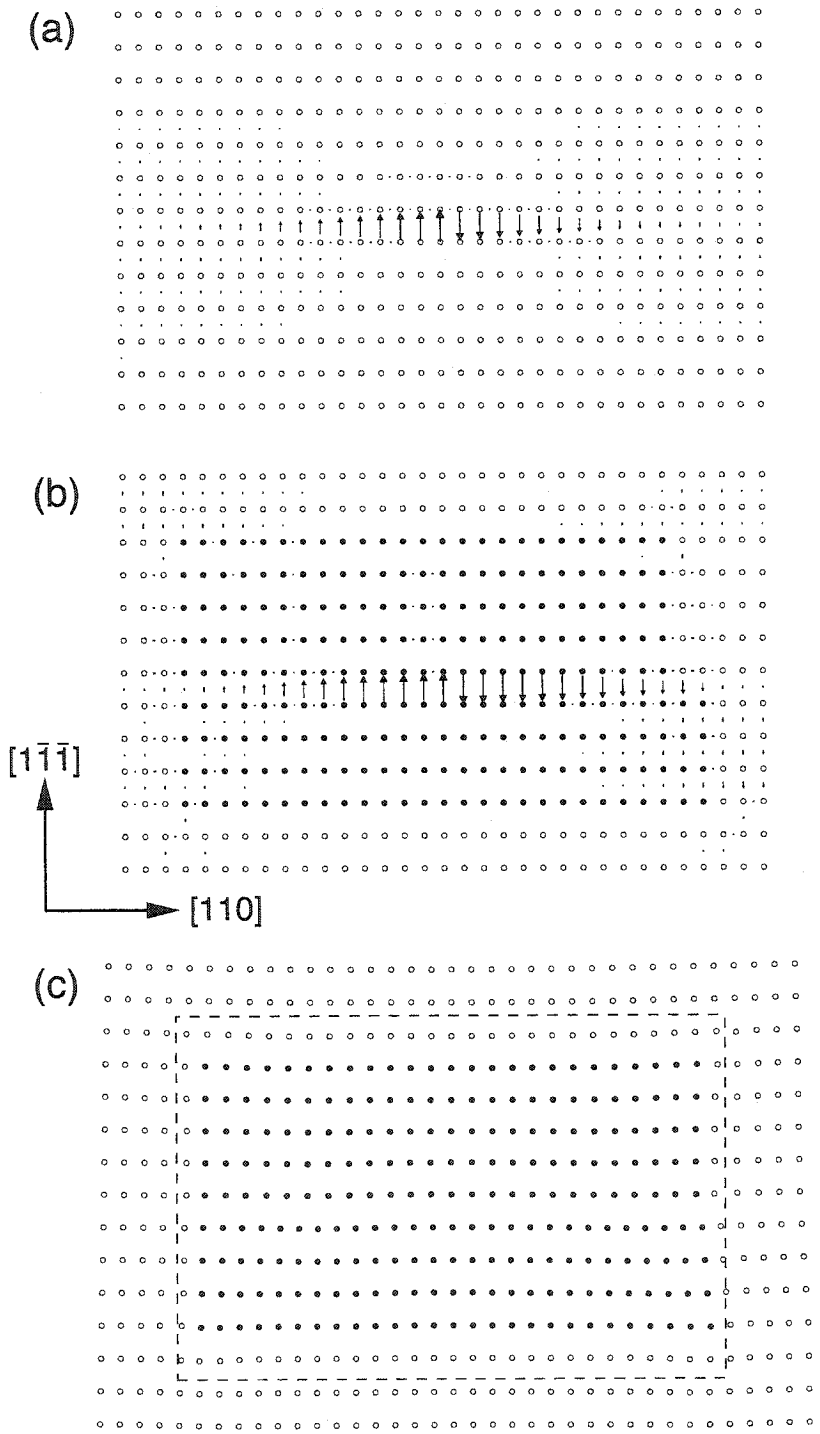

Figure 7.3: Differential displacement maps of an edge dislocation in aluminum (a) simulated with the (unscaled) Ercolessi-Adams interatomic potential, and (b) the quantum-interaction-energy coupling method of chapter 6. The relaxed positions of the atoms in the $(1\bar{1}2)$ plane are shown in (c). The region Ω^I to which the density ρ^I is confined is shown with a dashed line, and the $\mathbf{R}^I$ atoms are displayed as black circles. Note that only a portion of the simulated system is shown.

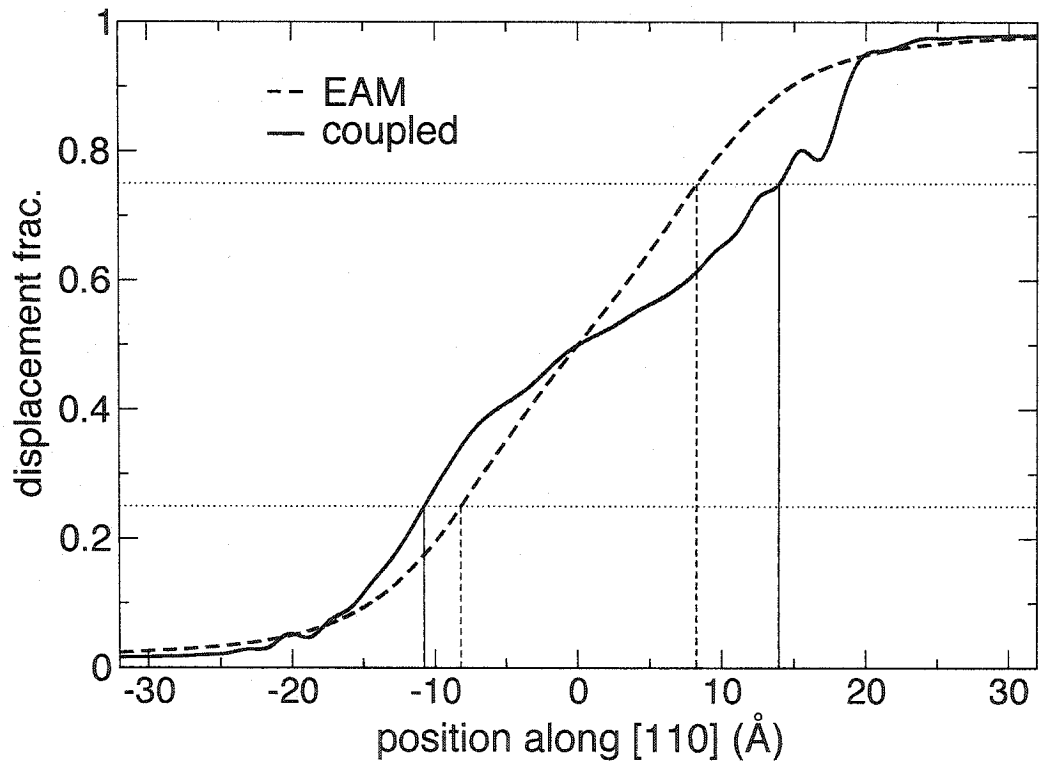

Figure 7.4: The relative displacement along the [110]-direction of the two $(1\bar{1}\bar{1})$-planes that sandwich the edge dislocation, for both the EAM simulation and the OF-DFT/EAM coupled simulation. The locations where the relative displacement equals 1/4 and 3/4 of the full Burgers-vector slip is indicated.

ference in energetics between the EAM potential and OF-DFT using the particular kinetic energy functional. This possibility was explored by performing a stacking fault calculation. A periodic simulation cell consisting of six (111)-planes, each with one aluminum atom, was established. The lattice translation vectors were $\mathbf{a}_1 = \frac{1}{2}[10\bar{1}]$, $\mathbf{a}_2 = \frac{1}{2}[1\bar{1}0]$, and $\mathbf{a}_3 = 2[111] + \frac{\alpha}{2}[2\bar{1}\bar{1}]$, where α ranges from 0 to 1 and induces a stacking fault in the fcc system. When $\alpha = 0$ or 1, the cell is a perfect fcc lattice. A value of $\alpha = \frac{1}{3}$ introduces an intrinsic stacking fault, as the standard $ABCABC\cdots$ fcc stacking becomes $|ABCABC|BCABCA|CAB\cdots$, with the '|' representing the stacking fault that occurs between the simulation cell and its periodic neighbor in the $\mathbf{a}_3$ direction. The six atoms in the unit cell were then allowed to relax in the [111] direction. This was done using Kohn-Sham DFT, then OF-DFT, and finally the EAM potential. The results are shown in Fig. 7.5.

The Kohn-Sham stacking fault had an energy of 108 mJ/m^2. The EAM stacking fault had a nearly identical energy of 105 mJ/m^2, which is not surprising, since the reproduction of the stacking fault energy was one of the constraints in the construction of the EAM potential[92]. The OF-DFT stacking fault had a much lower energy of 43 mJ/m^2. This provides an explanation of the observed wide dislocation dissociation in the coupled calculations. The width of dissociation is determined by the interplay of elastic and stacking fault energetics. A wider dislocation has a lower elastic energy, but has more stacking fault area. OF-DFT exhibits a comparable bulk modulus but a much smaller stacking fault energy than the EAM potential, thus resulting in wider dislocations.

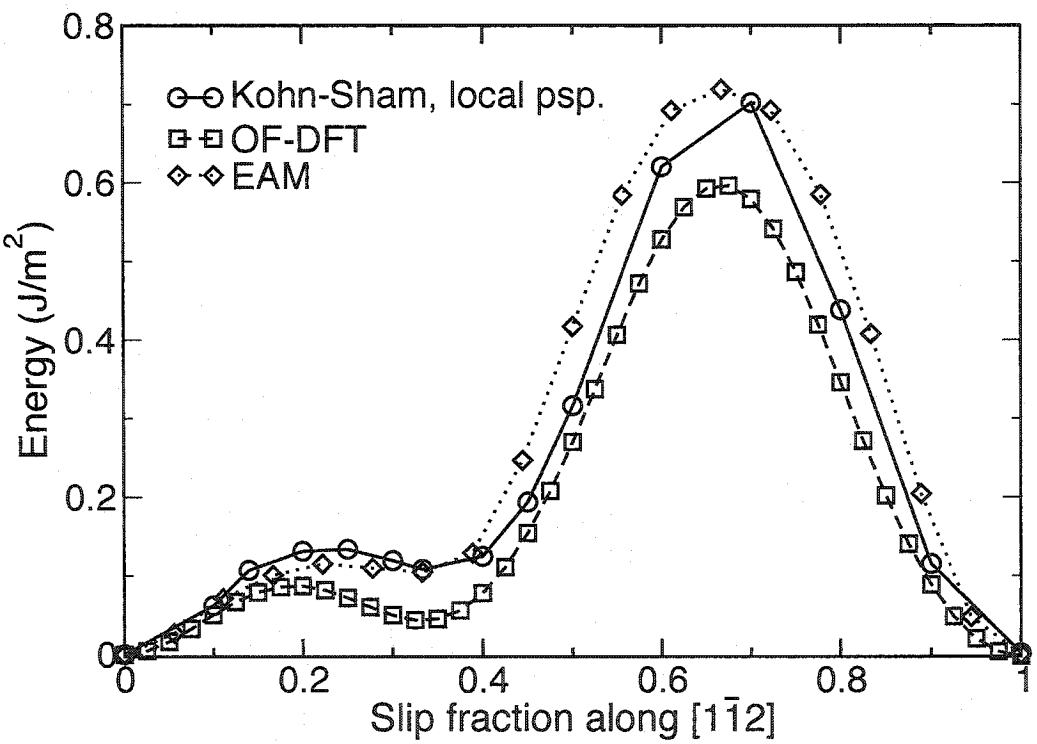

Figure 7.5: The stacking fault energy curves as a function of the slip parameter α, for Kohn-Sham DFT, OF-DFT, and the EAM potential.

7.2 Dislocation-hydrogen interaction

As a means of demonstrating the capabilities of the coupling method, we have also studied the interaction of hydrogen with the edge dislocation in aluminum. Dislocation-hydrogen interaction is of considerable technological interest, as it is a component problem of the more general issue of the embrittlement of metals in the presence of hydrogen. Decades of experiments and theories have not definitively elucidated the dominant mechanisms of H embrittlement, but have in general indicated that the real answer must be rather complex[97].

7.2.1 The physics of the hydrogen-dislocation interaction

The jellium model of the hydrogen impurity[98] often provides a simple but reasonable description of the energetics of hydrogen in metals. In this model, the energy of interaction between hydrogen and a material is approximated to be equal to the energy of interaction between hydrogen and a homogeneous electron gas of a density equal to the electron density at the point (or region) where the hydrogen is placed (before the hydrogen is placed there). The hydrogen-homogeneous electron gas interaction energy has been calculated[99] with DFT, and the result is that hydrogen binds most strongly to a homogeneous electron gas of density 0.0026 electrons/Bohr^3 with a strength of 1.8 eV. Note that the average electron density in aluminum is much higher at 0.027 electrons/Bohr^3. These facts lead to a picture of hydrogen being attracted to regions of relatively lower charge density in aluminum.

The jellium model would suggest that hydrogen prefers to occupy the octahedral sites of the fcc lattice over the tetrahedral sites. To test this, we performed an

OF-DFT calculation using a kinetic energy functional developed by Wang et al.[3], with a density-dependent kernel, and parameters $\{\alpha, \beta, \gamma, \rho_*\} = \{5/6 + \sqrt{5}/6,\ 5/6 - \sqrt{5}/6, 2.7, 0.08\ \text{Bohr}^{-3}\}$. A simulation cell was constructed which consisted of 32 Al atoms arranged in $2 \times 2 \times 2$ 4-atom cubic unit cells. A hydrogen atom was placed at a tetrahedral site in one calculation, and at an octahedral site in another, and the ionic forces were relaxed. The system with the hydrogen residing in the octahedral site was lower in energy than the system with the hydrogen in the tetrahedral site by 0.1 eV, in agreement with the jellium picture. We note, however, that Lu et al.[100] found the tetrahedral site to be lower in energy than the octahedral site by 0.07 eV when simulated with the Kohn-Sham method.

The jellium model suggests that dislocations affect the distribution of hydrogens in a metal. Because hydrogen has a lower energy when embedded in a region relatively lower electronic density in the metallic host, hydrogen should be attracted to regions under tensile strain. The strain field of the screw dislocation is mostly of shear character and has little hydrostatic component, and should not affect hydrogens much. But the edge dislocation puts the host material into compressive strain on one side of the slip plane and into tensile strain on the other side, and thus should, according to the jellium model, attract hydrogen to the side with tensile strain.

In order to get an idea of the strength with which strain affects the embedding energy of a hydrogen, we can perform some simple simulations. We can define the embedding energy of hydrogen in a metallic system that has undergone a homogeneous deformation with a strain tensor ϵ_{ij} as:

$$E_{\text{emb}}[\epsilon_{ij}] \equiv E[\text{metal}, \epsilon_{ij} + \text{H}] - E[\text{metal}, \epsilon_{ij}] - E[\text{isolated H}] \tag{7.1}$$

We can then define a useful quantity, the "embedding coefficient of compression," which we define as:

$$K_{\text{emb}} \equiv V \left.\frac{\partial E_{\text{emb}}}{\partial V}\right|_{V_0} = \frac{1}{3}\sum_{i=1}^{3} \left.\frac{\partial E_{\text{emb}}}{\partial \epsilon_{ii}}\right|_{\epsilon_{ii}\to 0} \tag{7.2}$$

This coefficient measures the change in hydrogen binding energy per unit change of hydrostatic strain. We performed four OF-DFT calculations using the 32-atom supercell used above in order to determine K_{emb}. Calculations were performed with and without hydrogen occupying an octahedral site, and with and without a small strain applied. By computing the numerical derivative of the energies, we found a K_{emb} of -1.47 eV. The sign of this quantity indicates that as strain increases (tensile strain), the embedding energy of hydrogen becomes more negative (becomes stronger), in accord with the jellium picture.

However, the jellium model is only the simplest picture. It has been shown, for instance, that this picture undergoes significant quantitative alteration when the effects of hybridization between the hydrogen's s electrons and the host metal's d electrons are taken into account, however such considerations are not relevant to the case of aluminum, which has no d electrons. Further, it does not take into account the effects and energetics of the relaxation of the lattice around the hydrogen. The picture of hydrogen-dislocation presented above certainly breaks down at the core of the dislocation, where the precise atomic arrangement is important. Multiscale simulation is one means of studying the problem more precisely.

7.2.2 Multiscale simulation of hydrogen-edge dislocation interaction

We have studied the energetics of hydrogen in the same edge dislocation system studied in Sec. 7.1.2. However, we have used a different kinetic energy functional than that used to examine the edge dislocation core structure. Rather, we have used the one described in Sec. 7.2.1. This functional, with its particular choice of parameters, was chosen to deal with the problem at hand. Unfortunately, this functional exhibits a stacking fault energy even smaller (and further from the Kohn-Sham calculation) than the functional used above for the isolated dislocation calculations. However, it was found that the functionals with density-independent kernels, like that used in Sec. 7.1, were not suitable for simulating aluminum systems that contained hydrogen: when the systems were ionically relaxed, inevitably the hydrogen would be drawn to one of the nearby aluminum atoms and the two atoms would collapse to a point, with the forces and energy diverging. This problem is rectified by using other kinetic energy functionals from the same work[3] that have density-dependent kernels. This latter variety of functional, although more computationally demanding, is intended to be more accurate over a wider range of densities, and were shown to be transferrable to a wider range of atomic environments, including surfaces and vacancies. However, switching to functionals with density-dependent kernels was not sufficient; the ρ_*-parameter needed to be increased as well. This parameter represents the density around which the density-dependent functionals most accurately reflect the proper linear response, and as such, for bulk-like systems we have usually set this parameter to the average electron density in the system. However, we have found that increasing

this average density tends to alleviate the kind of collapse problems described above. This seems to be a shortcoming of these kinetic energy functionals. We speculate that typical electronic relaxations performed with these functionals that result in a physical-looking charge density may actually be finding some local energy minimum of the functional, and that the global minimum energy may be $-\infty$, with the charge density collapsed to a point.

Using the the edge dislocation structure found in Sec. 7.1.2 as an initial starting point, we relaxed the structure using the functional with the density-dependent kernel. Then using this relaxed structure, we introduced a hydrogen atom at various sites around the dislocation, and relaxed the system again. The sites where the hydrogen atoms were introduced were chosen by selecting a number of octahedral sites in the un-dislocated crystal, and identifying the six Al atoms that make up the octahedral cage around the site. The location for introducing the hydrogen in the distorted system containing the edge dislocation was then taken to be the average distorted position of the six atoms that made up the octahedral cage in the undistorted crystal.

In Fig. 7.6 we display the positions at which we have placed hydrogen and computed the relaxed energy of the system. For each hydrogen position, we have calculated the embedding energy of the hydrogen atom as the energy of the dislocation system plus hydrogen, minus the energy of the dislocation system without hydrogen (and, technically, minus the energy of an isolated hydrogen atom, but since this term appears in all hydrogen embedding energies, and throughout this chapter we are only interested in the difference between embedding energies, the isolated hydrogen energy has not been calculated). We compare these embedding energies to the embedding

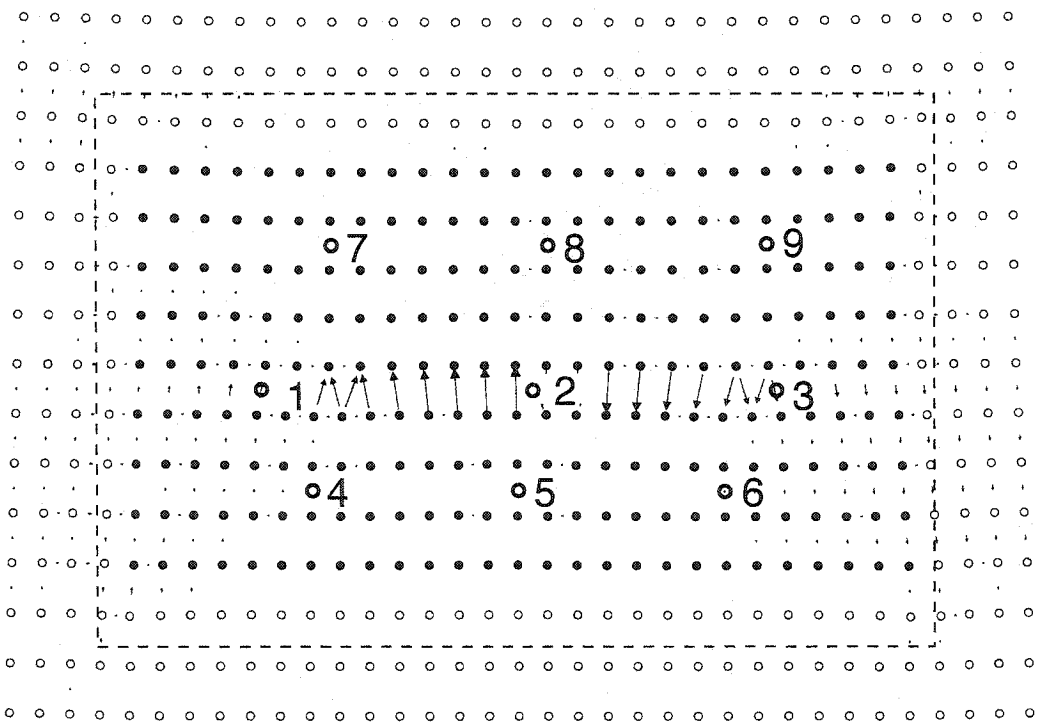

Figure 7.6: The selected sites of the hydrogens around the edge dislocation are shown in bold circles. The differential displacement of neighboring atomic columns are shown for reference.

energy of a hydrogen in an undislocated crystal, which we have also calculated with the coupling method. In principle, we could have used the data from the calculations used to derive the embedding coefficient K_{emb} for the hydrogen-in-crystal embedding energy, which were calculated with a fully periodic OF-DFT calculation. But a difference was found between this and the hydrogen-in-crystal embedding energy calculated with the coupling method, the latter being 0.05 eV higher. This difference is most likely an unfortunate manifestation of the coupling boundary.

In Table 7.1 we list the difference between the hydrogen-in-dislocated-crystal embedding energy and the hydrogen-in-undislocated-crystal embedding energy, for the hydrogen positioned at the nine sites. We also report the percentage change in volume of the octahedral cages around the hydrogen sites between the dislocated crystal (without hydrogen) and the undislocated crystal (without hydrogen). Finally, we report the change in embedding energy at the nine sites as predicted by the jellium theory, which we approximate as:

$$\Delta E_{\text{emb}}^{\text{jellium}} \simeq \frac{V_d - V_0}{V_0} K_{\text{emb}} \tag{7.3}$$

where V_d is the octahedron volume in the dislocated crystal without hydrogen, and V_0 is the octahedron volume in the undislocated crystal without hydrogen.

From the few octahedral sites examined, it is difficult to draw very general conclusions. To a reasonable extent, the physics of the jellium model seem to hold. The two hydrogen locations of highest energy, sites 4 and 6, are both located on the compressed side of the slip plane. And the sites located on the tensile side of the slip plane all have a lowered embedding energy. The fact that the octahedral sites that lie directly on the slip plane (sites 1, 2, and 3) have a change in embedding energy

Table 7.1: The hydrogen embedding energy at each site minus the hydrogen embedding energy in the undistorted lattice (ΔE_{emb}), the fractional change in volume of the octahedral sites going from the undislocated crystal to the dislocated crystal ($\Delta V/V$), and the change in embedding energy that the jellium picture would suggest ($\Delta E_{\text{emb}}^{\text{jellium}}$).

Site #	ΔE_{emb} (eV)	$\Delta V/V$ (%)	$\Delta E_{\text{emb}}^{\text{jellium}}$ (eV)
1	-0.03	+2.4%	-0.035
2	+0.17	+0.8%	-0.012
3	+0.06	+2.5%	-0.036
4	+0.01	-0.6%	+0.009
5	+0.04	+0.3%	-0.004
6	+0.02	-0.1%	+0.001
7	-0.03	+3.1%	-0.046
8	-0.01	+2.0%	-0.030
9	-0.02	+1.9%	-0.027

different than the prediction of the jellium picture is not contradictory: the change in energy suggested by the jellium picture is based on the change in volume of the octahedral site, but in the dislocated system, the octahedra that surround sites 1, 2, and 3 have vertices (Al atoms) that lie on both sides of the slip plane, and hence the octahedra get severely drawn out. The aluminum atoms most relevant to the energetics of the hydrogens at these sites may not be the six atoms that define the octahedral site in the undistorted crystal. The scale of the energetic effect of the dislocation on the hydrogen is on the same order of magnitude as room temperature thermal energetics. Thus at room temperatures and below, dislocations in Al would have a significant effect on the distribution of interstitial hydrogen.

The present set of calculations serve to highlight the kind of problem that the multiscale methods developed in this thesis can uniquely attack. In order to calculate the interaction between hydrogen and a dislocation, it is necessary to use a quantum

mechanical method in order to capture the energetics of how the hydrogen affects the valence electrons of the metal. But in order to allow the dislocation to assume its natural core structure, the correct long-ranged strain field must be imposed. Without a multiscale method, one may be left with only the option of simulating the system with a very large supercell. But even this alternative would entail using a supercell with multiple dislocations (a topological necessity for periodic systems), and the core structure can be strongly affected by dislocation-dislocation interaction (particularly in the case of dislocations in fcc metals, where the dislocation energy scales tend to be much smaller than in bcc metals.)

Chapter 8

Concluding remarks

In general, there is some sort of an inverse relationship between the efficiency of a simulation method and its accuracy or information content. This is rather obvious, actually, since if this relationship were not the case, then no one would have bothered to develop the less accurate methods! But it is the case, and physicists will always continue to search for new methods that achieve a better balance of both qualities. The orbital-free density functional theory methods, a central theme of the present thesis, occupies a position somewhere between Kohn-Sham DFT and tight-binding methods both in accuracy and efficiency in the context of simple metals. The goal of this thesis has been the general improvement of OF-DFT methods.

8.1 Pseudopotentials

If OF-DFT methods are to make use of pseudopotentials that are anything like the potentials of the established methods, then the pseudopotentials used must be local.

Local pseudopotentials are suitable for simple metals, and possibly a few transition metals if one is willing to include some core electrons in the calculation. But local pseudopotentials have not been developed significantly since the inception of non-local norm-conserving pseudopotentials. We have developed a method for generating local pseudopotentials in which the important qualities of norm-conserving potentials are maintained to the greatest possible degree. Specifically, the method involves starting with a good guess for the potential, then searching "local pseudopotential space" for the local potential that yields an optimal combination of scattering properties and smoothness.

We have used this method to produce a local pseudopotential for aluminum, a simple metal, but a non-trivial one both technologically and physically, exhibiting a significant amount of covalent bonding character. Although the present method produces pseudopotentials from a free-atom context, the aluminum potential is found to yield accurate bulk properties. The aluminum potential is a linear combination of two local potentials with shapes optimized to reproduce scattering off of a ground-state Al atom and an Al atom with one electron missing. Individually these two potentials did not exhibit good bulk properties, suggesting that the scattering properties in both atomic occupancies are important in the bulk environment.

8.2 Efficient methods for OF-DFT

The study of OF-DFT is largely tied up with the kinetic energy functional, since it is the only thing that formally distinguishes OF-DFT from the Kohn-Sham methods. The kinetic energy functionals that show promise for application to materials

science problems are designed to correctly reproduce the perturbation response of a homogeneous, non-interacting electron gas. This introduces a long-rangedness into the functionals, making them inefficient to evaluate in real space with direct methods, and thus making the whole OF-DFT method inefficient. We have developed an efficient method for real-space evaluation of these functionals. The functionals are split into separate parts, and each part is reformulated from a convolution integral into an equivalent differential equation. The method is shown to be efficient and accurate. The real space formulation allows for OF-DFT calculation of finite and other systems for which periodicity is undesired or inapplicable. This method turns out to be crucial to the OF-DFT-classical potential coupling method developed in chapter 6.

We have also developed an efficient method for evaluating atomic forces in electronic structure methods. Although not limited to the OF-DFT context, the method is particularly relevant to OF-DFT, for several reasons. First, the method only calculates forces on objects represented by local potentials, which we use if OF-DFT calculations anyway, and second, this method, when used in an OF-DFT context, makes the whole method efficient, while with the Kohn-Sham scheme calculating forces is not the computational bottleneck, so a Kohn-Sham calculation in which the present force method is used will not necessarily be efficient overall. The method is based on an efficient scheme for calculating the structure factor, borrowed from a fast method for calculating the Ewald energy. We demonstrate that the present force method can simultaneously achieve speed and accuracy.

8.3 Coupling OF-DFT and classical potentials in simple metals

The above developments come into play in the quantum-classical coupling method developed in this thesis. Coupling quantum and classical regions is essentially about preserving the correct bonding across the interface between the two regions and disturbing the sea of electrons as little as possible. Given the clear difference between the bonding in simple metallic systems and in covalent insulators or semiconductors, it is not surprising that a good quantum-classical coupling cannot be achieved for both types of system with the same specific methods.

We have developed a family of coupling methods for simple metals that utilizes the "embedding" picture of the metallic bond rather than the "pair-bond" picture that has been previously used to couple covalent solids and molecules. In order to achieve the simplest possible coupling, the quantum formalism chosen was DFT, in which the basic variable, the density, is in many ways the simplest possible variable that can be used. Through the DFT formalism, we show that coupling a quantum and classical region essentially only involves a boundary condition for the density, which in practice is much simpler to work with than, say, wavefunction boundary conditions. Our coupling formalism builds on previous work of others in which the basic DFT problem was partitioned in real space ("subsystem" DFT). In our method, the classical energetics is seen as an approximation to DFT energetics, and the classical subsystem is seen to carry an implicit charge density which we approximate as a sum of charge densities centered on each classical atom. Via the subsystem DFT for-

malism, the classical subsystem generates an effective potential in which the quantum subsystem lives.

An interesting feature of one of the methods is that it explicitly requires orbital-free density functionals. Furthermore, in order for the whole calculation of a coupled quantum-classical system to be efficient, an efficient real-space OF-DFT method is needed, and we employ our own method of chapter 4.

A family of coupling methods results from our formalism due to the freedom to choose the manner in which the energetics of interaction between the classical and quantum subsystems is calculated. This interaction itself can be calculated classically or quantum mechanically. We find that when the interaction is treated classically, the imperfection in coupling is seen in the forces on the atoms in the quantum mechanical region. Conversely, when the interaction is treated quantum mechanically, the forces on the atoms in the classical region are imperfect. The classical interaction energy method has the upper-hand in terms of simplicity. It is very readily implemented in practice by simply creating a "wrapper" layer around a DFT code and a molecular dynamics code. The quantum mechanical interaction method, however, has the advantage that the electronic charge density throughout the quantum region is seen to be very close to the charge density that would result if quantum mechanics were used to treat the entire system. This makes it suitable, for example, for introducing impurities, and in general simulating systems in which an accurate electronic charge density is needed.

8.4 The future

The methods developed in this thesis push forward the applicability and efficieny of orbital-free density functional methods. Yet they are not free of shortcomings. It is worthwhile to look at what might be done in the future to improve the methods and to consider possible applications of them.

We look first at pseudopotentials. While the method we have developed for creating local pseudopotentials seems to work well for Al, it would be interesting to carry out a detailed study to determine how useful local pseudopotentials can be across the periodic table. Unfortunately, our method requires too much human interaction to execute a completely thorough study[1].

An interesting thing to contemplate is a more radical twist of the pseudopotential ideas for use within OF-DFT. Traditional non-local pseudopotentials do not make sense without orbitals, but some sort of non-local "pseudopotential density functional" might take the form:

$$E^{\mathrm{psp}}[\rho] \equiv \int \rho^{\alpha}(\mathbf{r}) V(\mathbf{r},\mathbf{r}') \rho^{\alpha}(\mathbf{r}) d\mathbf{r} \tag{8.1}$$

and might be able to usefully go beyond local pseudopotentials. This form is clearly a generalization of the local potential, to which it reduces if $\alpha = 1/2$ and $V(\mathbf{r},\mathbf{r}') \rightarrow \delta(\mathbf{r}-\mathbf{r}')V^{\mathrm{psp}}(\mathbf{r})$. Another interesting and more radical possibility is that of some sort of environment-dependent pseudopotential. One might concieve of a local pseudopotential with a shape that depends on its atomic environment, or perhaps the amount of electron density in its vicinity.

[1] traditional pseudopotential construction has been considered to be a bit of a "black art," and the present method is not entirely free from that characterization!

The family of coupling methods developed in chapter 6 leave clear room for further exploration and improvement. It would be very interesting to explore some of the other possible choices one could make for the energy methods used in the coupling scheme. In particular it would be worthwhile to study the effectiveness of using the classical coupling scheme (Sec. 6.3.1) and employing the Kohn-Sham method for the DFT energetics.

When we tested the coupling methods on bulk aluminum embedded within bulk aluminum, a physical situation in which all atomic forces should be zero, we find that they are small, but not exactly zero. There is a pattern to these force errors which hints at why the errors are there. When classical energetics is used for computing the subsystem interaction, forces on the atoms in the quantum region are non-zero. We note that with this choice of coupling, the total system energy within our formalism is the classical energy of the whole system, *minus* the classical energy of the quantum atoms, *plus* the quantum energy of the quantum atoms. The forces on the atoms at the boundary of the quantum region are non-zero because the quantum and classical energetics do not match for these atoms. Conversely, when quantum energetics is used to calculate the subsystem interaction, forces on the atoms in the classical region that lie near the boundary of the subsystems are non-zero. And we note that for this choice of coupling, the total system energy is the quantum energy of the whole system, *minus* the quantum energy of the classical atoms, *plus* the classical energy of the classical atoms. And again, the atoms with significantly non-zero forces are the atoms in the classical subsystem that lie near the interface of the two subsystems, because the quantum and classical energetics there do not precisely cancel.

Thus it appears that the imperfections in the coupling method show themselves for atoms that end up being at a surface in some part of the calculation. The quantum and classical methods both define an energy that is a function on configuration space. For configuration-space points that are far from the bulk environment, such as surfaces, it is typically harder to achieve a matching between the quantum and classical energy functions. Because the present coupling methods ultimately both depend on cancellation of energetics of configurations far from the bulk, it is difficult to achieve a perfect coupling.

One would obviously improve the coupling if one had at one's disposal an orbital-free DFT method and a classical potential that gave very similar energetics for surface-like atomic configurations. One might protest that having such a potential would render the whole coupling method superfluous: if the classical energy function is very close to the orbital-free DFT method energy function, why do we need the computationally expensive orbital-free energetics in the first place? To this we respond that requiring good bulk energetic agreement and good surface energetic agreement is a far weaker condition that agreement over all of configuration space. Secondly, much of the potential power and usefulness of the coupling method also comes from the ability to easily introduce impurities and study the effects of chemistry on multiscale systems, something that is not easy to achieve with classical potentials.

Nevertheless the disagreement between classical and DFT energetics for surface configurations is a feature of the present state of affairs of classiscal potentials. These considerations demand that we ask an important question: could there be a coupling formalism that somehow only involves calculations of near-bulk configurations? We

do not know the answer to this. But if the answer is yes, then it seems likely that the method would be both theoretically and practically very satisfactory.

Bibliography

[1] L. Pollack *et al.*, Phys. Rev. B **55**, 15544 (1997).

[2] L. Goodwin, R. J. Needs, and V. Heine, J. Phys.-Condes. Matter **2**, 351 (1990).

[3] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 16350 (1999).

[4] G. B. Bachelet, D. R. Hamann, and M. Schluter, Phys. Rev. B **26**, 4199 (1982).

[5] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, B864 (1964).

[6] W. Kohn and L. J. Sham, Phys Rev **140**, 1133 (1965).

[7] G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996).

[8] M. D. Segall *et al.*, J. Phys.-Condes. Matter **14**, 2717 (2002).

[9] M. Frisch *et al*, *Gaussian 03*, Gaussian, Inc., Pittsburgh, PA, 2003.

[10] X. P. Li, W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).

[11] W. T. Yang, Phys. Rev. A **44**, 7823 (1991).

[12] D. R. Bowler, I. J. Bush, and M. J. Gillan, Int. J. Quantum Chem. **77**, 831 (2000).

[13] D. Sanchez-Portal, P. Ordejon, E. Artacho, and J. M. Soler, Int. J. Quantum Chem. **65**, 453 (1997).

[14] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[15] E. Hernandez, M. J. Gillan, and C. M. Goringe, Phys. Rev. B **53**, 7147 (1996).

[16] J. I. Steinfeld, in *Molecules and Radiation : an Introduction to Modern Molecular Spectroscopy* (MIT Press, Cambridge, MA, 1985), pp. 77–79.

[17] L. H. Thomas, Proc. Camb. Philos. Soc. **23**, 542 (1927).

[18] E. Fermi, Rend. Accad. Naz. Lincei **6**, 602 (1927).

[19] M. Levy, Proc. Natl. Acad. Sci. U. S. A. **76**, 6062 (1979).

[20] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).

[21] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[22] J. Harris and R. O. Jones, J Phys-F-Metal Phys **4**, 1170 (1974).

[23] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[24] D. C. Langreth and J. P. Perdew, Phys. Rev. B **21**, 5469 (1980).

[25] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **6412**, 129901 (2001).

[26] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **6408**, 089903 (2001).

[27] C. von Weizsäcker, Z. Phys. **96**, 431 (1935).

[28] J. Lindhard, Kgl. Danske Videnskab. Selskab. **28**, 8 (1954).

[29] D. C. Liu and J. Nocedal, Math. Program. **45**, 503 (1989).

[30] E. L. Shirley, D. C. Allan, R. M. Martin, and J. D. Joannopoulos, Phys. Rev. B **40**, 3652 (1989).

[31] T. Starkloff and J. Joannopoulos, Phys. Rev. B **16**, 5212 (1977).

[32] C. Fiolhais, J. P. Perdew, S. Q. Armster, and J. M. MacLaren, Phys. Rev. B **51**, 14001 (1995).

[33] C. Fiolhais *et al.*, Phys. Rev. B **53**, 13193 (1996).

[34] F. Nogueira, C. Fiolhais, and J. P. Perdew, Phys. Rev. B **59**, 2570 (1999).

[35] S. Watson, B. J. Jesson, E. A. Carter, and P. A. Madden, Europhys. Lett. **41**, 37 (1998).

[36] M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999).

[37] Y. Wang and R. G. Parr, Phys. Rev. A **47**, R1591 (1993).

[38] A. Nagy and N. H. March, Phys. Rev. A **39**, 5512 (1989).

[39] F. Aryasetiawan and M. J. Stott, Phys. Rev. B **38**, 2974 (1988).

[40] C. O. Almbladh and A. C. Pedroza, Phys. Rev. A **29**, 2322 (1984).

[41] S. H. Werden and E. R. Davidson, in *Local Density Approximations in Quantum Chemistry and Solid State Physics* (Plenum, New York, N.Y., 1987), pp. 117–184.

[42] Q. S. Zhao, R. C. Morrison, and R. G. Parr, Phys. Rev. A **50**, 2138 (1994).

[43] Q. S. Zhao and R. G. Parr, Phys. Rev. A **46**, 2337 (1992).

[44] M. Levy and J. P. Perdew, in *Density Functional Methods in Physics* (Plenum, New York, N. Y., 1985), pp. 11–30.

[45] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods* (Academic Press, New York, N. Y., 1982).

[46] M. Teter, Phys. Rev. B **48**, 5031 (1993).

[47] M. Y. Chou, Phys. Rev. B **45**, 11465 (1992).

[48] A. Filippetti *et al.*, Phys. Rev. B **52**, 11793 (1995).

[49] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[50] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[51] L. W. Wang and M. P. Teter, Phys. Rev. B **45**, 13196 (1992).

[52] E. Chacon, J. E. Alvarellos, and P. Tarazona, Phys. Rev. B **32**, 7868 (1985).

[53] P. Garcia-Gonzalez, J. E. Alvarellos, and E. Chacon, Phys. Rev. B **57**, 4857 (1998).

[54] P. Garcia-Gonzalez, J. E. Alvarellos, and E. Chacon, Phys. Rev. B **53**, 9509 (1996).

[55] F. Perrot, J. Phys.-Condes. Matter **6**, 431 (1994).

[56] E. Smargiassi and P. A. Madden, Phys. Rev. B **49**, 5220 (1994).

[57] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 17162 (1999).

[58] Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **58**, 13465 (1998).

[59] C. Herring, Phys. Rev. A **34**, 2614 (1986).

[60] Y. Saad, *Iterative methods for sparse linear systems* (PWS Pub. Co., Boston, 1996).

[61] W. Harrison, *Solid State Theory* (Dover, New York, N.Y., 1980).

[62] R. W. Freund, Siam J Sci Stat Comput **13**, 425 (1992).

[63] R. W. Freund and N. M. Nachtigal, Numer. Math. **60**, 315 (1991).

[64] H. A. van der Vorst and J. B. M. Melissen, IEEE Trans. Magn. **26**, 706 (1990).

[65] W. T. Yang, Phys. Rev. Lett. **66**, 1438 (1991).

[66] P. Ordejon, D. A. Drabold, M. P. Grumbach, and R. M. Martin, Phys. Rev. B **48**, 14646 (1993).

[67] P. Ordejon, D. A. Drabold, R. M. Martin, and M. P. Grumbac, Phys. Rev. B **51**, 1456 (1995).

[68] P. D. Haynes and M. C. Payne, Solid State Commun. **108**, 737 (1998).

[69] P. D. Haynes and M. C. Payne, Phys. Rev. B **59**, 12173 (1999).

[70] J. M. Soler *et al.*, J. Phys.-Condes. Matter **14**, 2745 (2002).

[71] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

[72] E. Hernandez and M. J. Gillan, Phys. Rev. B **51**, 10157 (1995).

[73] K. N. Kudin and G. E. Scuseria, Phys. Rev. B **61**, 16440 (2000).

[74] Y. H. Shao, C. A. White, and M. Head-Gordon, J. Chem. Phys. **114**, 6572 (2001).

[75] S. C. Watson and E. A. Carter, Comput. Phys. Commun. **128**, 67 (2000).

[76] H. Hellmann, *Einfürung in die Quantenchemie* (Deuticke, Leipzig, 1937).

[77] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[78] U. Essmann *et al.*, J. Chem. Phys. **103**, 8577 (1995).

[79] W. E and Z. Y. Huang, J. Comput. Phys. **182**, 234 (2002).

[80] W. Cai, M. de Koning, V. V. Bulatov, and S. Yip, Phys. Rev. Lett. **85**, 3213 (2000).

[81] J. Gao, *Reviews in Computational Chemistry*, edited by K. B. Libkowitz and D. B. Boyd (VCH, New York, N. Y., 1996), pp. 119–185.

[82] N. Govind, Y. A. Wang, A. J. R. da Silva, and E. A. Carter, Chem. Phys. Lett. **295**, 129 (1998).

[83] J. Q. Broughton, F. F. Abraham, N. Bernstein, and E. Kaxiras, Phys. Rev. B **60**, 2391 (1999).

[84] Q. Cui *et al.*, J. Phys. Chem. B **105**, 569 (2001).

[85] M. S. Daw, Phys. Rev. B **39**, 7441 (1989).

[86] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).

[87] F. Ercolessi, M. Parrinello, and E. Tosatti, Philos. Mag. A-Phys. Condens. Matter Struct. Defect Mech. Prop. **58**, 213 (1988).

[88] K. W. Jacobsen, J. K. Norskov, and M. J. Puska, Phys. Rev. B **35**, 7423 (1987).

[89] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[90] P. Cortona, Phys. Rev. B **44**, 8454 (1991).

[91] T. Kluner, N. Govind, Y. A. Wang, and E. A. Carter, Phys. Rev. Lett. **88**, 209702 (2002).

[92] F. Ercolessi and J. B. Adams, Europhys. Lett. **26**, 583 (1994).

[93] N. Choly and E. Kaxiras, Solid State Commun. **121**, 281 (2002).

[94] Y. A. Wang and E. A. Carter, in *Theoretical Methods in Condensed Phase Chemistry, Progress in Theoretical Chemistry and Physics* (Kluwer, Dordrecht, The Netherlands, 2000), pp. 117–184.

[95] D. Olmsted, K. Hardikar, and R. Phillips, Mod. and Sim. in Mat. Sci. and Eng. **9**, 215 (2001).

[96] C. Woodward and S. Rao, Phys. Rev. Lett. **88**, 216402 (2002).

[97] S. Mysers *et al.*, Rev. Mod. Phys. **64**, 559 (1992).

[98] M. J. Stott and E. Zaremba, Phys. Rev. B **22**, 1564 (1980).

[99] M. J. Puska and R. M. Nieminen, Phys. Rev. B **29**, 5382 (1984).

[100] G. Lu *et al.*, Phys. Rev. B **65**, 064102 (2002).

[101] O. H. Nielsen and R. M. Martin, Phys. Rev. B **32**, 3780 (1985).

Appendix A

The electron-ion contribution to the stress tensor

The methods of Chap. 5 are readily applied to the computation of the electron-ion contribution to the stress tensor. The stress tensor can be calculated with the methods of Nielsen and Martin[101]. One constructs an ionic potential skewed by a strain tensor ϵ, and a density similarly skewed and scaled to preserve normalization:

$$\begin{aligned} V_\epsilon^{ion}(\mathbf{r}) &\equiv V^{ion}((\ \ +\epsilon)^{-1}\mathbf{r}), \\ \rho_\epsilon(\mathbf{r}) &\equiv [\det(\ \ +\epsilon)]^{-1}\rho((\ \ +\epsilon)^{-1}\mathbf{r}) \end{aligned} \tag{A.1}$$

The electron-ion contribution to the stress density tensor is then given by:

$$\Omega\sigma_{\alpha\beta}^{e-i} = \left.\frac{\partial E_\epsilon^{e-i}}{\partial\epsilon_{\alpha\beta}}\right|_{\epsilon\to\mathbf{0}} = \frac{\partial}{\partial\epsilon_{\alpha\beta}}\int_{cell_\epsilon}\rho_\epsilon(\mathbf{r})V_\epsilon^{ion}(\mathbf{r})d\mathbf{r} \tag{A.2}$$

Under the strain transformation, the structure factor is unchanged, and the Fourier components of the density, $\mathcal{F}[\rho(l_1,l_2,l_3)]$, are simply scaled by $[\det(\ \ +\epsilon)]^{-1}$. The reciprocal lattice vectors, to first order in ϵ, become $\mathbf{b}_i \to (\ \ -\epsilon)\mathbf{b}_i$, and hence

$\partial \mathbf{b}_{i\gamma}/\partial \epsilon_{\alpha\beta} = -\delta_{\alpha\gamma}\mathbf{b}_{i\beta}$. Differentiating Eq. (5.26) with respect to $\epsilon_{\alpha\beta}$, one obtains an efficient method for the electron-ion stress density tensor:

$$\sigma_{\alpha\beta}^{e-i} = -\frac{E^{e-i}}{\Omega}\delta_{\alpha\beta} - \frac{N}{\Omega}\sum_{m_1,m_2,m_3}\frac{\mathbf{Q}_\alpha \mathbf{Q}_\beta}{|\mathbf{Q}|}\tilde{P}'(m_1,m_2,m_3)\mathcal{F}[\rho(l_1,l_2,l_3)]^* \times\tilde{B}(m_1,m_2,m_3)\mathcal{F}[Q(l_1,l_2,l_3)] \qquad \text{(A.3)}$$

where $\mathbf{Q} \equiv m_1'\mathbf{b}_1 + m_2'\mathbf{b}_2 + m_3'\mathbf{b}_3$, and $\tilde{P}'(m_1,m_2,m_3) \equiv d\tilde{V}^{psp}(\mathbf{Q})/d|\mathbf{Q}|$. This method has been subjected to simple tests comparing the derivative of the energy with respect to the crystal lattice constant to the components of the stress tensor. These tests indicate an accuracy similar to the present force method. It must be noted, however, that unlike the ionic forces, the cell stress has contributions from all terms in the energy, and $\sigma_{\alpha\beta}^{e-i}$ is merely one of them.

Appendix B

The interaction energy for complex kinetic energy functionals

We describe here how the interaction energy (defined in Chap. 6) can be efficiently calculated when the approximate kinetic energy functional used is of a more complicated form, containing a convolution term of the form of Eq. (6.20). That is, we will describe a method for evaluating:

$$\begin{aligned} T_{\mathrm{K}}^{\mathrm{int}}[\rho^I, \rho^{II}] = & \int f_{12}(\mathbf{r}) K(\mathbf{r}-\mathbf{r}') g_{12}(\mathbf{r}') \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \\ & - \int f_1(\mathbf{r}) K(\mathbf{r}-\mathbf{r}') g_1(\mathbf{r}') \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \\ & - \int f_2(\mathbf{r}) K(\mathbf{r}-\mathbf{r}') g_2(\mathbf{r}') \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}', \end{aligned} \tag{B.1}$$

where we have defined $f_1(\mathbf{r}) \equiv f(\rho^I(\mathbf{r}))$, $f_{12}(\mathbf{r}) \equiv f(\rho^I(\mathbf{r}) + \rho^{II}(\mathbf{r}))$, and so on. Then we define two new functions,

$$\begin{aligned} F(\mathbf{r}) &\equiv f_{12}(\mathbf{r}) - f_2(\mathbf{r}), \\ G(\mathbf{r}) &\equiv g_{12}(\mathbf{r}) - g_2(\mathbf{r}) \end{aligned} \tag{B.2}$$

Note that $F(\mathbf{r})$ and $G(\mathbf{r})$ are zero for points $\mathbf{r} \notin \Omega^I$. Using F and G we can re-express Eq. (B.1) as:

$$\begin{aligned} T_{\mathrm{K}}^{\mathrm{int}} = & \int_{\Omega^I} F(\mathbf{r})(K * G)(\mathbf{r})\mathrm{d}\mathbf{r} - \int_{\Omega^I} f_1(\mathbf{r})(K * g_1)(\mathbf{r})\mathrm{d}\mathbf{r} \\ & + \int_{\Omega^I} F(\mathbf{r})(K * g_2)(\mathbf{r})\mathrm{d}\mathbf{r} + \int_{\Omega^I} G(\mathbf{r})(K * f_2)(\mathbf{r})\mathrm{d}\mathbf{r} \end{aligned} \tag{B.3}$$

where we have now defined:

$$(K * G)(\mathbf{r}) \equiv \int K(\mathbf{r} - \mathbf{r}')G(\mathbf{r}')\mathrm{d}\mathbf{r}', \text{ etc.} \tag{B.4}$$

We point out that if this interaction energy is being calculated in a coupled simulation in which the energetics of region I are calculated using the same kinetic energy (i.e. $E[I]$ and $E^{\mathrm{int}}[I, II]$ being treated at the same level of theory), then the final term of Eq. (B.3) is equal to and will cancel with the corresponding term in $E[I]$.

So with Eq. (B.3) we have expressed $T_{\mathrm{K}}^{\mathrm{int}}$ purely in terms of intergrals over Ω^I; the problem is now reduced to efficiently calculating the functions $(K * f_2)(\mathbf{r})$ and $(K * g_2)(\mathbf{r})$ for points $\mathbf{r}$ within Ω^I. A straightforward integration for each point $\mathbf{r} \in \Omega^I$ is not an option, because $K(r)$ is typically long-ranged, and thus determining $(K * f_2)(\mathbf{r})$ at one single point $\mathbf{r}$ would require an integration over the volume of the whole coupled system, which would be highly inefficient. We now describe a method for determining $(K * f_2)(\mathbf{r})$, and $(K * g_2)(\mathbf{r})$ can be determined with precisely the same method.

In earlier work[93] we have developed a method for efficiently evaluating convolutions like Eq. (B.4) when the convolution kernel $K(r)$ is of the particular form typically encountered in kinetic energy functionals involving convolution terms[52,

54, 51, 58, 3]. We will invoke this method to determine $(K * f_2)$. In this method, the kernel is fit in reciprocal space with the following form:

$$\begin{aligned} \tilde{K}(k) &\simeq \sum_i \tilde{K}_i(k), \\ \tilde{K}_i(k) &= \frac{P_i k^2}{k^2 + Q_i} \end{aligned} \tag{B.5}$$

where P_i, Q_i are complex fitting parameters. The kernels encountered in many kinetic energy functionals are well-fit with this form, with 4 terms. The kernel in real space can be expressed as the sum of the inverse Fourier transform of each term of Eq. (B.5):

$$\begin{aligned} K(\mathbf{r}) &\simeq \sum_i K_i(\mathbf{r}), \\ K_i(\mathbf{r}) &\equiv P_i \delta(\mathbf{r}) - P_i Q_i \frac{e^{-\sqrt{Q_i} r}}{4\pi r} \end{aligned} \tag{B.6}$$

Thus $(K * f_2)$ can be written as the sum of separate convolutions:

$$(K * f_2)(\mathbf{r}) = \sum_i (K_i * f_2)(\mathbf{r}), \tag{B.7}$$

$$(K_i * f_2)(\mathbf{r}) \equiv \int K_i(\mathbf{r} - \mathbf{r}') f_2(\mathbf{r}') \mathrm{d}\mathbf{r}' \tag{B.8}$$

Because in reciprocal space the $(K_i * f_2)$ satisfy:

$$\left[k^2 + Q_i\right] \widetilde{(K_i * f_2)}(\mathbf{k}) = P_i k^2 \tilde{f}_2(\mathbf{k}), \tag{B.9}$$

in real space they satisfy:

$$\left[\nabla^2 - Q_i\right] (K_i * f_2)(\mathbf{r}) = P_i \nabla^2 f_2(\mathbf{r}) \tag{B.10}$$

i.e. they are solutions to (complex) Helmholtz equations which can be solved with conjugate-gradient-based methods[63]; such methods are efficient and only involve op-

erations within Ω^I. The solutions to Eqs. (B.10) are only well-defined when boundary conditions for $(K_i * f_2)(\mathbf{r})$ are supplied.

We propose the use of Dirichlet boundary conditions. The value of $(K_i * f_2)(\mathbf{r})$ for points $\mathbf{r}$ on the boundary of Ω^I can be found by evaluating the convolutions, Eqs. (B.8). Because of the regular nature of $\rho^{II}(\mathbf{r})$, being the sum of atomic densities, an efficient method for evaluating these convolutions exists. The form of the convolution that needs to be evaluated is:

$$(K_i * f_2)(\mathbf{r}) = \int K_i(\mathbf{r}-\mathbf{r}') f\left(\sum_j \rho^{\mathrm{at}}(\mathbf{r}'-\mathbf{R}_j^{II})\right) \mathrm{d}\mathbf{r}' \tag{B.11}$$

If $f(\rho)$ were a linear function, then this would reduce to a sum of pair functions. For many kinetic energy functionals, $f(\rho)$ is not linear, but equal to some power of ρ: $f(\rho) = \rho^\alpha$. This leads us to consider a Taylor expansion of $f(\rho)$ about some average density ρ_0. This Taylor expansion suffers in places where $\rho^{II}(\mathbf{r})$ is near 0, which occurs, for instance, in the center of Ω^I. An expansion that is much more accurate down to small values of ρ is obtained if we Taylor expand the function $h(\rho) \equiv f(\rho)/\rho$ and express $f(\rho^{II})$ in terms of this expansion:

$$f\left(\sum_j \rho^{\mathrm{at}}(\mathbf{r}'-\mathbf{R}_j^{II})\right) \simeq \left(\sum_j \rho^{\mathrm{at}}(\mathbf{r}'-\mathbf{R}_j^{II})\right)$$
$$\times \left[h(\rho_0) + h'(\rho_0)\left[\sum_k \rho^{\mathrm{at}}(\mathbf{r}'-\mathbf{R}_k^{II}) - \rho_0\right] + \cdots\right] \tag{B.12}$$

In Fig. B.1 we illustrate the effectiveness of (B.12) compared to expanding $f(\rho)$ directly when $f(\rho) = \rho^{1.5}$ and $\rho_0 = 1$. Although the h-expansion is taken only to first order, while the f-expansion is taken to second order, the h-expansion is seen to be more accurate at small ρ.

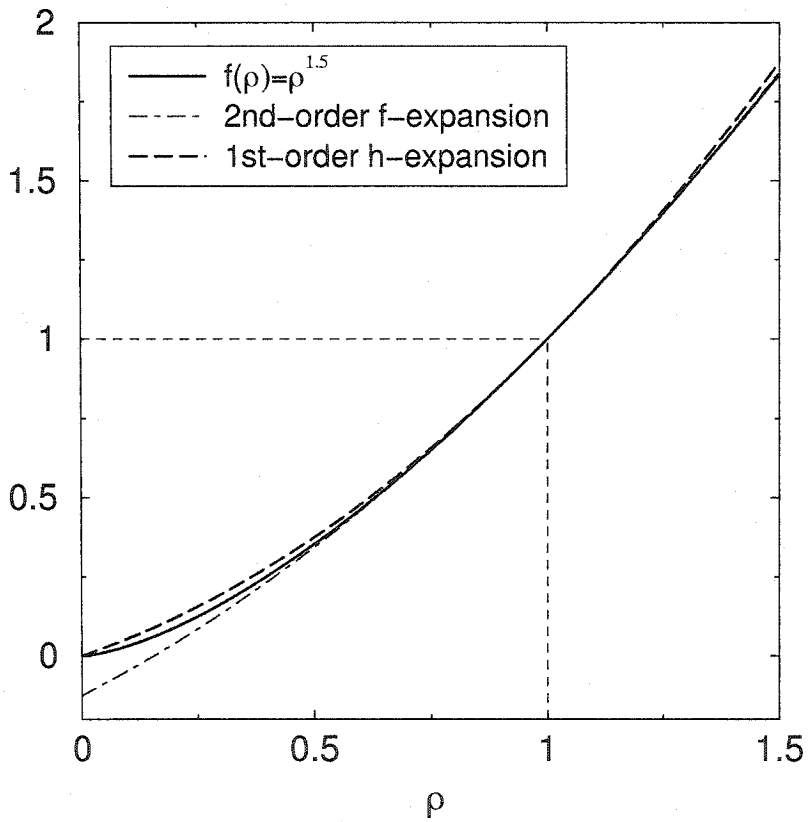

Figure B.1: A demonstration of the Taylor expansion of Eq. (B.12) compared to a direct Taylor expansion of $f(\rho) = \rho^{1.5}$ about $\rho_0 = 1$.

Upon substitution of the expansion of Eq. (B.12) in the convolution, Eq. (B.11), we find:

$$\begin{aligned}(K_i * f_2)(\mathbf{r}) &\simeq [h(\rho_0) - \rho_0 h'(\rho_0)] \sum_j L_i^{(1)}(\mathbf{r} - \mathbf{R}_j^{II}) \\ &\quad + h'(\rho_0) \sum_{j,k} L_i^{(2)}(\mathbf{r} - \mathbf{R}_j^{II}, \mathbf{r} - \mathbf{R}_k^{II}), \\ L_i^{(1)}(\mathbf{R}) &\equiv \int K_i(\mathbf{r}')\rho^{\mathrm{at}}(\mathbf{R} - \mathbf{r}')\mathrm{d}\mathbf{r}', \\ L_i^{(2)}(\mathbf{R}, \mathbf{R}') &\equiv \int K_i(\mathbf{r}')\rho^{\mathrm{at}}(\mathbf{R} - \mathbf{r}')\rho^{\mathrm{at}}(\mathbf{R}' - \mathbf{r}')\mathrm{d}\mathbf{r}'\end{aligned} \tag{B.13}$$

$L_i^{(1)}$ is the convolution of an atomic density with $K_i(r)$. $L_i^{(2)}(\mathbf{R}, \mathbf{R}')$ is the convolution of the product of two atomic densities with $K(r)$, and consequently vanishes when the two atomic densities do not overlap. The integrand is non-zero only where the overlap occurs. It thus makes sense to express $L_i^{(2)}$ in terms of new coordinates, illustrated in Fig. B.2:

$$M_i(\mathbf{R}_{\mathrm{c}}, \mathbf{R}_{\mathrm{rel}}) \equiv L_i^{(2)}(\mathbf{R}_{\mathrm{c}} - \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}}, \mathbf{R}_{\mathrm{c}} + \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}}) \tag{B.14}$$

In the case of a spherically symmetric $\rho^{\mathrm{at}}(\mathbf{r})$, $L_i^{(1)}(\mathbf{R})$ is a function only of $|\mathbf{R}|$, and $M_i(\mathbf{R}_{\mathrm{c}}, \mathbf{R}_{\mathrm{rel}})$ will depend only on $|\mathbf{R}_{\mathrm{c}}|$, $|\mathbf{R}_{\mathrm{rel}}|$, and $\mathbf{R}_{\mathrm{c}} \cdot \mathbf{R}_{\mathrm{rel}}$. We now argue that the dependence on $\mathbf{R}_{\mathrm{c}} \cdot \mathbf{R}_{\mathrm{rel}}$ is weak. We can write the expression for M_i as:

$$\begin{aligned}M_i(\mathbf{R}_{\mathrm{c}}, \mathbf{R}_{\mathrm{rel}}) = \int & K_i(\mathbf{R}_{\mathrm{c}} - \mathbf{r})\rho^{\mathrm{at}}(\mathbf{r} - \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}}) \\ & \times \rho^{\mathrm{at}}(\mathbf{r} + \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}})\mathrm{d}\mathbf{r}\end{aligned} \tag{B.15}$$

If we expand K about $\mathbf{R}_{\mathrm{c}}$, we find:

$$\begin{aligned}M_i(\mathbf{R}_{\mathrm{c}}, \mathbf{R}_{\mathrm{rel}}) = \int & [K_i(|\mathbf{R}_{\mathrm{c}}|) - \mathbf{r} \cdot \nabla K_i(\mathbf{R}_{\mathrm{c}}) + \cdots] \\ & \times \rho^{\mathrm{at}}(\mathbf{r} - \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}})\rho^{\mathrm{at}}(\mathbf{r} + \tfrac{1}{2}\mathbf{R}_{\mathrm{rel}})\mathrm{d}\mathbf{r}\end{aligned} \tag{B.16}$$

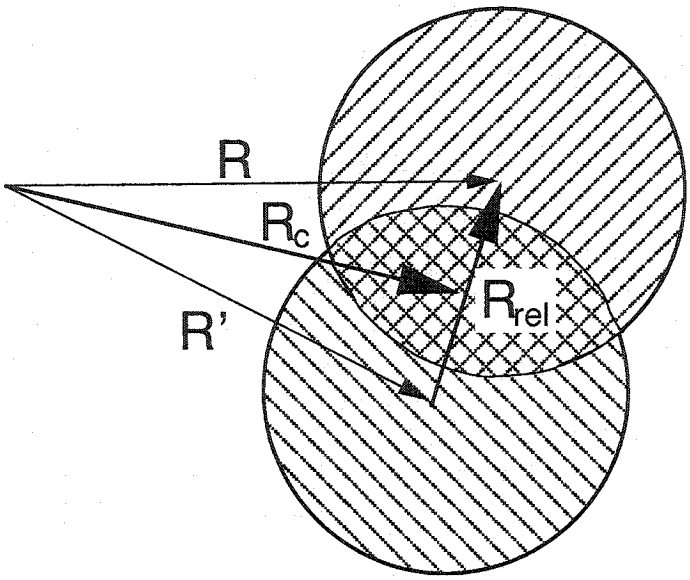

Figure B.2: New coordinates $\mathbf{R}_c$ and $\mathbf{R}_{rel}$ for evaluating the three-center integrals of $L_i^{(2)}$.

The integrals over the odd powers of $\mathbf{r}$ in the expansion of K_i vanish by symmetry. Thus if we truncate the series at first order, the first order term vanishes, leaving only the 0th order term:

$$M_i(\mathbf{R}_c, \mathbf{R}_{rel}) \simeq K_i(|\mathbf{R}_c|)P(|\mathbf{R}_{rel}|), \tag{B.17}$$

$$P(|\mathbf{R}_{rel}|) \equiv \int \rho^{at}(\mathbf{r})\rho^{at}(\mathbf{r} - \mathbf{R}_{rel})d\mathbf{r} \tag{B.18}$$

Truncating the expansion of K_i at the first order is reasonable, because $K_i(r)$ oscillates around the Fermi wavelength of the system, a length scale close to that of ρ^{at}. This approximate form, Eq. (B.17) will behave quite badly at small R_c, because K_i diverges at the origin, and the radial averaging of this divergence that occurs in Eq. (B.15) is not reflected in Eq. (B.17). Thus we replace $K_i(|\mathbf{R}_c|)$ there with the convolution of K_i with a Gaussian of unit weight and a variance r_0 given roughly by the length-scale

of the overlap regions of the atomic densities (i.e. some fraction of the range of $\rho^{\rm at}(\mathbf{r})$):

$$\begin{aligned} M_i(\mathbf{R}_{\rm c}, \mathbf{R}_{\rm rel}) &\simeq K_i'(|\mathbf{R}_{\rm c}|)P(|\mathbf{R}_{\rm rel}|), \\ K_i'(\mathbf{r}) &\equiv (K_i * w)(\mathbf{r}), \\ w(\mathbf{r}) &\equiv \pi^{-3/2} r_0^{-3} e^{-(r/r_0)^2} \end{aligned} \tag{B.19}$$

Summarizing these results, we find that we can approximately evaluate $(K_i * f_2)$ as:

$$\begin{aligned} (K_i * f_2)(\mathbf{r}) \simeq [h(\rho_0) - \rho_0 h'(\rho_0)] \sum_j L_i^{(1)}(\mathbf{r} - \mathbf{R}_j^{II}) \\ + h'(\rho_0) \sum_{<j,k>} K_i' \left(\left| \mathbf{r} - \tfrac{1}{2}(\mathbf{R}_j^{II} + \mathbf{R}_k^{II}) \right| \right) P\left(\left| \mathbf{R}_j^{II} - \mathbf{R}_k^{II} \right| \right) \end{aligned} \tag{B.20}$$

where the summation over $< j, k >$ indicates that we need only sum over pairs of region II atoms with overlapping densities. The derivation of Eq. (B.20) involved several approximations, and thus is not expected to be precise. We only propose that Eq. (B.20) be used to generate the boundary conditions for Eqs. (B.10) that determine the $(K_i * f_2)(\mathbf{r})$ within region Ω^I, and we have found that the resulting $(K_i * f_2)(\mathbf{r})$ is more dependent on the source term than the boundary conditions. Nevertheless, because of the inaccuracies of Eq. (B.20), we define a new region, $\Omega^{I'}$, that contains and extends a bit beyond Ω^I, and we use Eq. (B.20) to obtain the boundary conditions for points $\mathbf{r}$ that lie on the boundary of $\Omega^{I'}$, and we solve Eqs. (B.10) for all $\mathbf{r} \in \Omega^{I'}$, so that the resulting $(K_i * f_2)(\mathbf{r})$ are accurate for all $\mathbf{r} \in \Omega^I$.

Thus we have all the pieces necessary to compute $T_{\rm K}^{\rm int}$. In summary, we do this as follows:

- Using Eq. (B.20), we can evaluate $(K_i * f_2)(\mathbf{r})$ and $(K_i * g_2)(\mathbf{r})$ for points $\mathbf{r}$ on the boundary of a region $\Omega^{I'}$ that is slightly larger than Ω^I.

- Using those boundary conditions, the Helmholtz equations (B.10) are solved, yielding $(K_i * f_2)(\mathbf{r})$ and $(K_i * g_2)(\mathbf{r})$ for all points $\mathbf{r} \in \Omega^{I'}$.

- Then $(K * f_2)(\mathbf{r})$ and $(K * g_2)(\mathbf{r})$ are constructed with Eq. (B.7), and we can evaluate $T_{\mathrm{K}}^{\mathrm{int}}$ via Eq. (B.3).

Ther kernel interaction energy $T_{\mathrm{K}}^{\mathrm{int}}$ also gives a small contribution to the forces on region II atoms near the 1-2 boundary. By differentiating Eq. (B.1), we find:

$$\begin{aligned}
\frac{\partial T_{\mathrm{K}}^{\mathrm{int}}}{\partial \mathbf{R}_j^{II}} &= -\int \nabla\rho(\mathbf{r}-\mathbf{R}_j^{II})\,[f'_{12}(\mathbf{r})(K*G)(\mathbf{r}) \\
&\quad +g'_{12}(\mathbf{r})(K*F)(\mathbf{r}) + F'(\mathbf{r})(K*g_2)(\mathbf{r}) \\
&\quad +G'(\mathbf{r})(K*f_2)(\mathbf{r})]\,\mathrm{d}\mathbf{r}, \qquad \text{(B.21)} \\
(K*F)(\mathbf{r}) &\equiv \int K(\mathbf{r}-\mathbf{r}')F(\mathbf{r}')\mathrm{d}\mathbf{r}', \\
f'_{12}(\mathbf{r}) &\equiv f'(\rho^I(\mathbf{r})+\rho^{II}(\mathbf{r})), \\
F'(\mathbf{r}) &\equiv f'_{12}(\mathbf{r}) - f'(\rho^{II}(\mathbf{r})), \text{ etc.}
\end{aligned}$$