Chapter 1

Introduction

> If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the atomic *fact*, or whatever you wish to call it) that *all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another.* In that one sentence, you will see, there is an *enormous* amount of information about the world, if just a little imagination and thinking are applied.
>
> – Richard P. Feynman [1]

1.1 Interatomic Forces

The realization that matter is composed of tiny corpuscles called atoms is perhaps the greatest breakthrough in the history of science. The atomic hypothesis identifies the (usually) indivisible carriers of chemical identity and structure, which opens the possibility of predicting macroscopic materials phenomena from the microscopic level. Obviously, we could not understand chemical reactions like dissolution, catalysis and burning without talking about atoms because they are needed to identify the reacting substances, but the atomic hypothesis is also essential in cases not involving chemical changes. By thinking of matter as a collection of incompressible, indestructible atoms of finite size and mass that stick to each other, we can define physical concepts like heat (kinetic energy of atomic motion) and cohesion (potential energy of atomic arrangement). These ideas suffice for an intuitive picture of processes like sound propagation, evaporative cooling, melting, crystal growth, viscous fluid flow, solid deformation and fracture. Indeed, a central task of modern materials science is to understand macroscopic phenomena such as these in terms of the underlying atomic mechanisms.

The crucial property of atoms that determines such behavior is how they prefer to stick to each other; in other words, what are the interatomic forces? This question can be answered from first principles ("*ab initio*") by solving the quantum mechanical equations of motion for the atomic constituents, electrons and nuclei. While this is certainly the most reliable approach, there are two basic reasons to look for simpler descriptions that somehow capture the essential physics of the quantum mechanical treatment.

The first is that an *ab initio* solution of atomic motion is prohibitively expensive to calculate for more than a few hundred atoms, even on the fastest supercomputer. With simpler, classical models called *empirical interatomic potentials*, the same computers can perform simulations of millions of particles, making possible atomistic studies of incredibly complicated processes like melting, diffusion, sintering, amorphization, surface growth, radiation damage, cracking and plastic deformation in single crystals. The danger, of course, with such simulations is that in switching to the simple model, the essential physics has been lost, rendering the results meaningless. Certainly quantitative accuracy is reduced, but often even the qualitative behavior is incorrect.

The possible rewards of large-scale atomistic simulations, however, are so great that intense effort has been focussed recently on developing computationally efficient models of interatomic forces with increased realism. The goal is to make full use of high performance computers, which double in speed almost every year. If more realistic interatomic potentials can be designed, large-scale simulations may someday allow us to understand complex phenomena like the brittle to ductile transition, for example, from an atomistic level. That kind of knowledge would be extremely powerful in predicting materials properties and even engineering improved materials through computer experimentation.

While accurate, large-scale simulation is the usual motivation, there is another reason to develop simple models of interatomic forces that is rarely mentioned, namely to build our intuitive understanding of chemical bonding. Scientists like Cauchy, Poisson and Born were pondering the nature of interatomic forces long before the invention of the computer. Already then it was a nontrivial problem to explain experimental data, like cohesive energies and elastic constants, in terms of simple physical principles.

Today, the situation is much more challenging, because *ab initio* calculations, based on density functional theory [2] in the local density approximation (LDA) [3], have tremendously extended the body of accurate "experimental" data available that needs explanation [4]. Unfortunately, the output of every *ab initio* calculation is merely a number, the total energy of a particular atomic configuration[1]. The number is reliable, but we have little guidance in how to interpret it in terms of atomic interactions (or if such a thing is even possible within a simple framework).

In applied science, the primary goal is perhaps to predict physical properties with the maximum accuracy possible for the situation of interest. In more basic science, however, there is an inherent value in simple explanations, because a unified view of complex and seemingly disparate data can often be achieved.

[1]Actually, the meaningful output also includes the band structure and density distribution of the Kohn-Sham quasiparticles (which are believed to closely resemble the real, interacting electrons), but this information does not aid much in the quantitative understanding of interatomic forces.

> Although simplification is a sheer mathematical necessity, for many-body problems, there is also a more positive reason for it. What is it we really want from a theory? In the most interesting cases, what we are seeking is enlightenment, a general understanding of what is going on, a physical picture, something essentially qualitative that could be explained in relatively few words.... Simplification is an art rather like that of the cartoonist who captures the key features of a familiar face in a few deft strokes to make it instantly recognizable.
>
> – Sir Alan Cottrell [5]

One of our goals in this thesis is to represent a familiar covalent solid (Si) with a few "strokes" (or rather, potential energy functions) as deftly as we can, and see if it is still recognizable. Less colorfully, we aim to determine to what extent interatomic forces in covalent solids can be understood in simple terms and what degree of realism is possible with empirical potentials. We shall proceed by developing new methods of extracting classical interactions from *ab initio* data and by using that information, along with much blood, sweat and tears, to produce the best model we can. In the end, by seeing how well our model performs in a wide range of applications, we may learn something about the general strengths and limitations of interatomic potentials for covalent solids.

1.2 Why Atoms Rather Than Electrons and Nuclei?

Of course, we know that atoms are not really indivisible "little particles", so what do we mean when we talk about interatomic forces anyway? In the simple cases of noble elements or ionic solids, the physical picture given by Feynman above is quite accurate. Atoms (or ions) in these substances basically maintain their shape and chemical identity like spheres interacting via pairwise, radial attractions and repulsions. In the important cases of metals and semiconductors, however, the subatomic constituents, play crucial and vastly different roles in cohesion. One cannot begin to understand the subtleties of cohesion in these cases without considering electronic structure and its interplay with nuclear positions. Delocalization of electrons can also make the conceptual identity of an atom rather vague. For example, conduction electrons in a metal flow across macroscopic regions of space at enormous speeds averaging 10^{16} Å/s, typically undergoing collisions (strong interactions) with at least one out of every thousand atoms they pass. This means that in just one second, a typical metallic electron may be "shared" by over 10^{14} nuclei! Nevertheless, even though individual electrons and nuclei cannot be associated with each other, each nucleus is surrounded by a relatively static cloud of electron density, that may be conceptually divided among the nuclei to identify atoms.

In this regard, covalent solids, whose (semiconducting) electronic structure interpolates between the strongly localized (insulating) cases of noble or ionic solids and the delocalized (conducting) case of metals, are even more difficult to view from the atomic perspective. In these materials, valence electrons partially localize along "chemical bonds" with appreciable density concentrated in between bonded pairs of atoms. A valence electron in a bonding state is more or less equally shared between the two nuclei in the bond but may also resonate among a number of nearby bonding states. Thus, in a covalent solid the picture of well-defined atoms (resembling isolated atoms in a gas) sticking to each other at a preferred distance seems a bit strange. A covalent solid is more like a huge molecule, made up of around 10^{23} atoms.

In spite of these complications, however, the atomic picture of cohesion is justified in most condensed matter systems for one simple reason: nuclei are much heavier than electrons. The proton-electron mass ratio is 1836.15, and a Si^{28} nucleus is 5.157×10^4 times more massive than an electron. As a consequence, the fast and complex motions of the electrons are superimposed upon the (relatively) slow meanderings of the nuclei. Thus, if we identify each nucleus as the center of an atom, we can forget about explicitly keeping track of the electrons. Instead, we envision a "ball-and-spring model" of the material: the atoms (soft balls centered at nuclear positions) interact via an implicit force law (springs connecting the balls) determined by the electron density in the presence of the nuclei. Since core electrons stay tightly bound to the nuclei, it is more accurate to think of the atomic balls as representing ions consisting of the nuclei and their sheaths of core electrons, and the interatomic springs as the forces due to valence electron densities in the presence of the ionic pseudopotentials. In covalent solids, the springs can even assume a physical identity of their own, as chemical bonds, and we have then a simple language to describe atomic mechanisms in terms of the distorting, breaking and reforming of bonds. Since we do not need to solve explicitly for electronic structure in this approach, our task is to derive a classical interaction that somehow mimics the effect of the electrons on the nuclei. Although the problem is nontrivial, the reward for success is a tremendous conceptual and computational simplification.

1.3 The Born-Oppenheimer Energy Surface

An empirical interatomic potential is not just a toy model for qualitative understanding, akin to the Ising Hamiltonian for magnetic spin systems; instead, it can in principle provide a faithful *quantitative* reproduction of *ab initio* quantum mechanical predictions. This is a consequence of the adiabatic approximation, first applied to molecules by Born and Oppenheimer, which provides rigorous support for the arguments given above [6]. The adiabatic approximation justifies separation of the nuclear and electronic variables based on their disparate masses. The resulting errors in energy are smaller than the typical energy level spacings by a factor of order the mass ratio, less than 10^{-4} for most materials. Therefore, to a very good approximation, electrons move quantum-mechanically in a quasi-static external potential determined by the instantaneous positions of the nuclei, always in equilibrium due to their greater speeds. Conversely, the nuclei move in a force field determined by the time-averaged electron densities. In the context of molecules, the perturbative Born-Oppenheimer method or the variational method of Messiah may be used to derive precise *classical equations of motion for the nuclei* [6], and in the context of solids, the method of Car and Parinello [7] derives similar equations from self-consistent *ab initio* calculations of the instantaneous electronic ground state using density functional theory [2] in the local density approximation [3].

Whatever first principles method is used, the adiabatic approximation justifies the existence of the Born-Oppenheimer (BO) energy surface $E(\{\vec{R}_i\})$, which expresses the total energy of the system of electrons and nuclei as a function of the nuclear positions $\{\vec{R}_i\}$ alone. The force on atom i due to the presence of all the other atoms is simply $-\nabla_i E$, so *if* the BO surface were known, perhaps after being tabulated from many *ab initio* quantum-mechanical calculations, then atomistic dynamics could be performed using classical mechanics. Even though quantum mechanical equations of motion would not be integrated, the dynamics of the nuclei would be *exact* within the (very good) adiabatic approximation.

The difficulty is that *the BO surface is astronomically complicated*, except in special (very restrictive) cases. For example, in the simple case a 100 atom periodic lattice at finite temperature with up to 10% bond length distortions, in order to tabulate the total energy with a spatial resolution of 0.1% of the average bond length, we would need to perform a billion *ab initio* energy calculations. Now suppose we could somehow compile this data, it would still be a nontrivial task to design a data structure to store the massive table and an algorithm to access it efficiently. For more interesting situations involving larger systems with greater disorder, like the 1728 atom liquid simulations described in Chapter 6, it is clearly intractable to calculate, store and access the relevant regions of the BO surface, and no advantage over an *ab initio* quantum mechanical approach is achieved.

An obvious alternative is to design an empirical potential as follows: guess a simple functional form with adjustable parameters that allows efficient computation of forces, and fit it to a few carefully selected points on the *ab initio* BO surface. This approach is motivated by necessity, but there is no *a priori* guarantee that the potential is *transferable*, *i.e.* that it faithfully approximates regions of the BO surface to which it was not explicitly fit. Unfortunately, there is no small parameter, like $\hbar$ in semi-classical approximations, to bound errors, which may be unpredictably large or small in different cases. Very little theoretical guidance exists to select the correct form of an empirical potential. As a result, designing transferable potentials is a challenging and frustrating business, but, nevertheless, remarkable progress has been made for a wide range of materials [8].

1.4 Cohesion in Covalent Solids

The class of materials that has most resisted a transferable description by an empirical potential involves covalent bonding. In the prototypical case of silicon, over thirty potentials have been produced in recent years (reviewed in Chapter 2), but a satisfactory description has not been achieved, even for bulk material. To appreciate the subtleties involved in covalent solids, let us consider the simplest picture of interatomic forces, described above by Feynman, namely a pair potential in which atoms are attracted toward each other but resist being squeezed too close together. This kind of model, exemplified by the Lennard-Jones 6-12 potential for Van der Waals forces and the Coulomb electrostatic force law, captures the essential physics of noble, ionic, and, to a some extent, even metallic solids, but it is oversimplified in the case of covalent solids. For example, if we "apply a little imagination and thinking", we would predict qualitatively wrong crystal structures. The preference of atoms attracting via radial forces is to have as many neighbors as possible, since atoms simply want to be close to each other (up to a minimum distance). Thus, with pair potentials, crystal structure is mostly a matter of geometry, the close-packing of hard spheres in three-dimensions, which leads to lattices like face-centered cubic (FCC), hexagonal close-packed (HCP), body-centered cubic (BCC) and simple cubic (SC) with high density and coordination

(6–12). Covalent solids, however, crystallize in much more open structures like the diamond or graphitic lattices with low density and coordination (3–4), and usually *increase* their density upon melting.

Before the advent of quantum theory, the idea of pair potentials (radial forces) was advocated by influential scientists like Cauchy, and it was not until Born's seminal paper on diamond elastic constants in 1914 [9] that the need for non-radial forces to model covalent bonding was fully appreciated [10, 11]. He realized that additional forces are needed with explicit dependence on the *angles* subtended by the lines connecting atoms, not only to lower the energy of open lattices versus close-packed ones, but also to stabilize them against shear deformations. The model of Born was modified (for rotational invariance) and generalized by Harrison in 1956 [12]. Finally, in 1985 the conceptual framework of pair bonding and angular forces was extended to disordered structures with the potential of Stillinger and Weber (SW) [13], which has proven to be one of the most successful empirical models for covalent solids.

The original ideas of Born were motivated by elastic constant analysis, and thus are primarily relevant for small distortions of the diamond crystal structure. The SW potential illustrates, quite surprisingly, that the same concepts work fairly well for a broad range of configurations including crystal defects, liquid and amorphous states, but new ideas about the functional form of interatomic potentials are needed. The most obvious feature lacking in the SW model is *environment dependence*, or adaptation of the force law to changes in the local bonding environment. For example, liquid silicon is a metal with greater density and coordination than the solid, and metallic bonding is known in other materials to be described best by embedded-atom potentials [14], which usually have density-dependent bond strengths and no angular forces. It also seems unphysical that the SW model does not adapt to changes in covalent hybridization, for example, between the diamond and graphitic lattices. Experience with semi-empirical, quantum-mechanical (tight-binding) models has shown that transferability can be substantially increased by including environment-dependence [15], which suggests that the same may be true for classical, interatomic potentials.

Therefore, a crucial and ongoing theme in recent research is the environment dependence of interatomic potentials for covalent solids. Motivated by theoretical work [16, 17], environment dependence was first introduced by Tersoff in 1987 [18]. Since then numerous generalizations have appeared [19], and one version recently received theoretical justification from approximations of quantum theories [20, 21]. The next breakthrough in environment-dependence came with dangling bond vector of Chelikowsky *et. al.* [22], which is important for cases of broken lattice symmetry, like surfaces and clusters. In spite of these innovations, however, a significant improvement in transferability over the (much simpler) SW potential has not yet been achieved [19], which suggests that new ideas are needed to augment the Tersoff and Chelikowsky models. Unfortunately, the standard approach of fitting *ad hoc* functional forms has resulted in frustration, as increasingly complex and flexible functional forms have failed to yield substantial gains in transferability. Thus, new methods are needed to identify an improved functional form and to then guide the arduous fitting process.

1.5 Scope and Outline of the Thesis

This brings us to the questions we seek to answer in this work: Are there any indisputable facts about the functional form of an interatomic potential that can be deduced from *ab initio* calculations? Can potentials be derived directly from *ab initio* data, without fitting any adjustable parameters? How should environment dependence be included in the functional form? How well do theoretical results translate into accurate potentials, in practice? Is it really possible to attain a fully transferable description of a covalent material in all its important phases with a computationally efficient empirical potential, and if so, what methods might lead us to discover it?

We attempt to answer these important questions by developing new theoretical methods and applying them to the prototypical case of silicon. In recent years, Si has emerged as the representative covalent material due to its great technological importance and the vast amount of useful experimental and theoretical studies available to test new ideas. Our theoretical methods are equally applicable to related materials like Ge, C, and with minor extensions even alloys of these elements, but investigation of whether any of our specific results for Si carry over to these materials is beyond the scope of this thesis. Furthermore, a satisfactory description of bulk Si (crystal, defects, liquid and amorphous) has not yet been achieved, so here we shall focus on bulk interatomic forces, and postpone analysis of surfaces and molecules for subsequent work. Bulk Si already contains sufficiently complicated physics that we may use it to make progress toward answering our motivating questions.

We begin in Chapter 2 by comparing and contrasting existing empirical potentials for silicon, and mention some useful results from analytic approximations of quantum mechanical models. In Chapter 3, elastic constant relations are derived for various functional forms in the diamond and graphitic crystal structures to better understand interatomic forces mediated by hybrid covalent bonds. In Chapter 4, novel methods are developed to obtain many-body interatomic potentials directly from *ab initio* cohesive energy curves, which shed light on global changes in bonding across covalent and metallic structures. In Chapter 5, these theoretical results are incorporated into a new functional form called the "Environment-Dependent Interatomic Potential" (EDIP), which is fitted and tested for crystal phases and bulk defects. The computational efficiency and transferability of the fitted EDIP for disordered phases is studied in Chapter 6 using molecular dynamics techniques. Finally, Chapter 7 contains concluding remarks on our successes and failures and prospects for future research.