INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

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 bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI 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.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Diffusion in Semiconductors:
A Theoretical Study

Blas Pedro Uberuaga

A dissertation submitted in partial fulfillment of
the requirements for the degree of

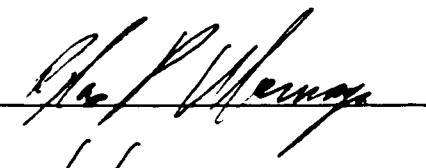
Doctor of Philosophy

University of Washington

2000

Program Authorized to Offer Degree: Department of Physics

In presenting this dissertation in partial fullfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Bell and Howell Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date 6/9/00

University of Washington

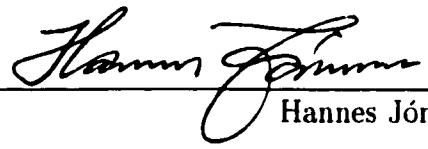
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Blas Pedro Uberuaga

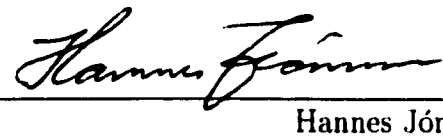
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

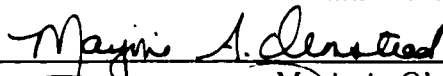


Hannes Jónsson

Reading Committee:

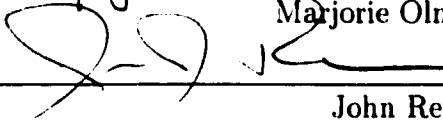


Hannes Jónsson



2 June 2000

Marjorie Olmstead



John Rehr

Date: 6/9/00

University of Washington

Abstract

Diffusion in Semiconductors: A Theoretical Study

by Blas Pedro Uberuaga

Chair of Supervisory Committee

Professor Hannes Jónsson
Chemistry

As the dimensions of semiconductor devices become smaller and smaller, greater control is needed in the fabrication process. It thus becomes more and more important to understand details of the processes that determine these critical dimensions. For example, accurate spatial profiles of dopants in Si are essential to achieve desired electronic characteristics. Experimental measurements give very good insight into the nature of diffusion, but they cannot discern the details of the atomic motion. This is where theory can play a role by calculating the pathways on an atomic level. This thesis is a contribution towards a detailed understanding of the mechanisms of diffusion.

We apply both empirical potential calculations as well as more sophisticated density functional theory methods to the problem of diffusion in semiconductors. We are interested both in the behavior in bulk material as well as the initial stages of growth on the surface. In particular, we have studied the self-diffusion of both Si and Ge, as well as foreign atom diffusion in Si. Ge diffusion in Si is studied in most detail, but some aspects of the diffusion of dopants, such as B, As, and Sb, are also described.

While density functional theory (DFT) can give valuable qualitative information about the shape of the energy landscape, and thereby a detailed microscopic description of the various diffusion mechanisms, we have found it underestimates the formation energy of defects in Si. For example, the formation of an interstitial is underestimated by by about 1 eV. Efforts to correct this shortcoming, thought to be due to the poor description of exchange and correlation used in the periodic DFT code, are discussed. Using cluster calculations, where an improved description of exchange and correlation can be implemented, has proven to be a viable way to estimate the corrections to the periodic DFT calculations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AED: Auger electron diffraction

BA: Bolding-Andersen empirical potential for Si

CE: concerted exchange

DFT: density functional theory

DV: dimer vacancy

ES: elastic sheet method

GGA: generalized gradient approximation

LDA: local density approximation

LEED: low energy electron diffraction

MBE: molecular beam epitaxy

MEP: minimum energy path

ML: monolayer

NEB: nudged elastic band method

OL: overlayer

PED: photoelectron diffraction

PW91: Perdew-Wong GGA exchange-correlation functional

TST: transition state theory

VASP: Vienna ab-initio simulation program

ACKNOWLEDGMENTS

I would like to acknowledge the guidance and support of my thesis advisor, Hannes Jónsson, who has both motivated me and given me the opportunity to explore other avenues while working with him.

I would like to thank Roland Stumpf of Motorola, who was my mentor during my internship with Motorola. I would also like to acknowledge his coworker, Wolfgang Windl, for many useful discussions and for both of their collaboration on part of the work presented here. In addition, I would like to thank their supervisor, Mike Masquelier, for giving me the opportunity to work with the Computational Materials Group at Motorola.

Professor Marjorie Olmstead and her student, Dr. Michael Leskovar were very instrumental in motivating much of the work I have done at the University of Washington. Working with Brett Schroeder on the CaF project has also been a pleasure.

The Jónsson Research Group has been a very stimulating and motivating group to work with. I would especially like to acknowledge Enrique Batista, Graeme Henkelman, Kiril Tsemekhman, and Arthur Smith for valuable discussions. Enrique and Graeme also deserve special thanks for their collaboration on several projects described in this thesis.

I would like to thank all of the friends and classmates I have at the University of Washington, both for the sanity breaks and occasional beer, as well as the great environment they helped create for me here.

I would like to thank J. Tersoff, E. Kaxiras, and M. Lagally for helpful discussions, and S. Meng for assistance with AED modeling.

Finally, I would like to thank Lisa Van De Graaff for all of her patience with me.

DEDICATION

To Pedro and Monica Uberuaga,

my parents,

and to Lisa Van De Graaff.

Chapter 1

INTRODUCTION

The semiconductor industry is continuously shrinking the critical dimensions of the devices used to build sophisticated computer circuits. The reason for this drive is to increase device speed as well as the density of devices that can be built on a given area of Si substrate. For example, the minimum feature size of dynamic memory (DRAM) components has gone from 250 nm to 180 nm from 1997 to 1999 and is expected to reach dimensions of 50 nm by the year 2012. This reduction in individual device size, accompanied by increased chip size, will result in a projected increase in the number of transistors on a chip from 11 million in 1997 to 1.40 billion in 2012. The control of the deposition processes necessary to achieve this small dimension and high volume of transistors will correspondingly have to improve. [3]

Devices are built by a series of deposition steps, in which either a dopant atom is implanted into the underlying Si substrate, Si is deposited to act as a buffer or fix damage, metallization layers are added, or the surface is oxidized to create insulating layers. The purpose of the dopant atoms is to change the electronic properties of Si. Some elements, such as B, act as acceptors. They only have three electrons in their valence shell, as opposed to four in the valence shell of Si, and thus grab another from the Si crystal, or "accept" an electron from the Si crystal. This leaves a positive charge carrier, a hole, in the crystal which is free to move in the crystal, creating current flow. Elements with a fifth, extra electron in their valence shell, such as P, As and Sb, act as donors and "donate" excess electrons to the Si substrate. In both cases,

the dopant atoms alter the electronic properties of the underlying Si substrate. The interplay between regions doped with acceptors (p-doped regions) and regions doped with donors (n-doped regions), along with properly placed insulating and conducting regions, give Si the necessary electronic characteristics to make a device operate.

To achieve predicted nanoscale device dimensions within the next decade, great control has to be exercised on the placement of the dopants in the Si substrate. The spatial distribution of dopants determines device properties and the final location of the dopants is, in turn, determined by the diffusion constant of each type of dopant. Many factors determine the overall diffusion constant for a given element; these will be discussed in more detail later. They include the concentration and mobility of defects, the temperature, and the charge state of both the dopant atom and the defects involved. Most common dopants used today are thought to diffuse via a defect mechanism, involving either vacancies or interstitials, or both. In some cases, an exchange mechanism is also possible, in which atoms just swap positions in the ideal lattice. The starting point, then, in determining the diffusion constant is to understand the motion of defects themselves, or self-diffusion.

Experiments can give great insight into the motion of defects in the lattice. They try to use measurements of the diffusion constant on the macroscopic scale to infer what is occurring on the atomic scale. Recent experiments measuring the overall diffusion constant and activation energy of self-diffusion of Si using isotope diffusion agree on the values for these quantities. However, estimating the contributions to the overall diffusion of exchange mechanisms and Si self-interstitial and vacancy mediated diffusion is very difficult and is usually done via indirect methods in which a dopant that is believed to diffuse via a defect is studied, and the diffusion of that defect is then inferred from the dopant diffusion. Thus, the description of self-diffusion depends critically on the particular model chosen to analyze dopant diffusion. This has lead to disagreement among different experiments. Knowledge of the individual contributions to diffusion of interstitials and vacancies is critical to modeling Si wafer

processing. For example, diffusion of interstitials determines the motion of boron in the system, and diffusion of vacancies determines the motion of arsenic. Theory approaches the problem from the opposite perspective, looking at the atomic motions and then trying to make the connection to the macroscopic diffusion constant. Together, theory and experiment can give a more complete description of diffusion by confirming or eliminating possible but different interpretations that one alone cannot distinguish. Part of the goal of this work is to determine which experiments have described self-diffusion more accurately.

Besides being critically dependent on the model chosen for analysis, experiments have the further shortcoming that the various aspects of atomic motion that comprise the diffusion constant cannot be extracted from experiments. For example, the primary temperature dependence of the diffusion constant is controlled by the activation energy, which in turn can be split into the formation energy and the migration energy. It is very difficult for experiments to separate the activation energy into the formation and migration energy, but it is relatively straight forward to find these two energies theoretically.

Si self-diffusion refers to the diffusion of Si atoms within a Si crystal. Self-diffusion is determined by the motion of vacancies and interstitials, as well as the direct exchange of atoms in the crystal. The motion of a given defect is described by four quantities: the entropies of migration and formation and the energies of migration and formation. Knowledge of these four quantities for all relevant migration processes is required to determine the diffusion constant. A process is relevant if, at a given temperature, it contributes significantly to the total diffusion. Once the diffusion constant is known for both self- and dopant diffusion, it can be used in larger scale Monte Carlo or coupled differential equation simulations that can describe diffusion on a macroscopic level and be used to predict dopant profiles.

Once self-diffusion is understood, dopant diffusion can be analyzed. Dopant diffusion is governed by the interaction of the dopant atom with the same point defects

that control self-diffusion. Vacancy mediated diffusion, for example, involves exchange of a substitutional dopant atom with a neighboring Si lattice vacancy. Which defect-mediated diffusion mechanism dominates depends on the nature of the dopant atom. Smaller dopants tend to diffuse via interstitials while larger atoms diffuse via vacancies.

The goal of the research described in this thesis is to calculate the four quantities that determine the diffusion constant: the energy of formation, the entropy of formation, the energy of migration, and the entropy of migration. To do this, various methods have been used, improved, and implemented. By analyzing all of the possibly important diffusion mechanisms for the diffusion of a given species, we can determine which ones need to be considered in the larger scale simulations. By expanding our study to include the Si(100) surface, we can also gain insight about the initial growth of films on the surface, which is important not only for the growth of epitaxial layers and superlattice structures, but will also govern the evolution of complex structures such as self-assembled quantum dots.

There have been previous theoretical studies of diffusion in Si, but they are incomplete in various ways. They either used empirical potentials which do not give good descriptions for the material under non-ideal conditions (conditions, such as transition states, that were not considered when the potentials were developed) or, because of computational limitations, have not been able to do a complete study of all relevant quantities. We have calculated all of the quantities that are necessary in describing the diffusion constant from an atomic perspective, within the harmonic approximation.

In chapter 2, a brief description of diffusion is given, including how macroscopic, experimentally measurable quantities are related to the atomic scale motions that are treatable with our theoretical techniques. Each of the physical properties of atomic motion are discussed in detail, along with an explanation of how each can be calculated within the framework of our methods.

Density functional theory is a sophisticated tool for studying solids at an atomic level. It allows a quantum mechanical treatment of the electrons important for bonding, but is also computationally efficient enough to allow the study of 100 or even 1000 atom systems. The basics of density functional theory are briefly introduced in chapter 3 with the intention of giving an idea of what parameters determine the accuracy of a calculation. The specific values of these parameters used in the calculations described in this thesis are also given in chapter 3.

The rest of the thesis describes the actual calculations conducted. Chapter 4 describes Si self-diffusion. The concerted exchange, interstitial and vacancy mechanisms are studied in detail. Entropies of formation and migration and energies of formation and migration are calculated for the important atomic mechanisms.

Chapter 5 describes mechanisms important for Ge diffusion in the bulk Si crystal. The concerted exchange, interstitial and vacancy mechanisms are discussed and the importance of each is evaluated.

In chapter 6, I discuss the interdiffusion of Ge into the Si(100) surface, important for understanding the epitaxial growth of Ge on Si. The theoretical calculations are examined in light of recent experiments, also described herein, in which sub-monolayer deposition of Ge leads to interdiffusion to the fourth layer of the Si substrate.

The next two chapters deal with specific mechanisms of diffusion for other dopants. Chapter 7 describes results of calculations of the concerted exchange for common dopants in Si. The following chapter describes vacancy assisted diffusion for As and Sb, two common donor dopants used in Si that have been determined experimentally to have large vacancy components of diffusion in Si.

Ge self-diffusion is discussed in chapter 9. As in the case of Ge diffusion in Si, we examine the vacancy, interstitial and concerted exchange contributions to Ge self-diffusion. This system exhibits to an even greater degree the problems of of commonly used methods in describing the formation energy of defects.

It has been found by several groups, us included, that density functional theory,

the theoretical frame work we use to study these systems, using the generalized gradient approximation developed by Perdew and Wang (the PW91 functional) [4] finds activation energies that are on the order of 1 eV lower than experiment. Chapter 10 discusses this problem and an approach to correct it. Using cluster calculations and a more exact exchange-correlation functional, we are able to get activation energies that are in the experimental range.

Mills, Jónsson and Schenter introduced a method for finding minimum energy paths and, thus, saddle points between any two states, if those states are known. This method is known as the nudged elastic band (NEB). We have made improvements on the original NEB method that allow the method to find saddle points more precisely with little or no additional computational costs. These improvements, done in collaboration with Graeme Henkelman, are described in chapter 11.

Finally, a new method for calculating zero-flux surfaces, useful in partitioning the electron charge density into "atoms", is described in chapter 12. Examples of the method applied to Si bonding structures and H_2O clusters are described.

The result of this work will be a better understanding of the relative contributions of interstitials and vacancies to the self-diffusion of Si, as well as to the diffusion of Ge and dopant atoms in Si, which in turn will help resolve the disagreement in experimental measurements of these contributions. Better understanding of individual defect motion will allow for better modeling of Si wafer processing.

Chapter 2

DIFFUSION IN SEMICONDUCTORS

2.1 Introduction

Ideally, to model device fabrication and processing, one would like to be able to describe the macroscopic motion of a large number of dopant atoms introduced into a Si substrate. Describing diffusion at this level ignores the individual motion of each atom, instead treating the average collective motion. However, this average, macroscopic diffusion is the direct result of atomic scale processes, and these processes are precisely the characteristics that we are able to study using density functional theory and the nudged elastic band method. Density functional theory will be introduced in the next chapter, while the nudged elastic band method will be described below.

This chapter describes how calculable quantities describing the individual atomic processes are linked to macroscopic diffusion. Specifically, we discuss how multiple atomic scale processes combine to determine the the macroscopic diffusion. The connection between the atomic scale and the macroscopic scale is done by assuming the processes are governed by transition state theory (TST) and that the atoms, on average, execute a random walk through the crystal. The details of how to calculate each of these microscopic quantities are also given.

2.2 The Diffusion Constant

The diffusion constant (sometimes also referred to as the diffusivity) is the quantity that describes how readily a gradient in the concentration of some species will lead to diffusion of that species at a given temperature. This is the quantity that we want

to know. It is composed of four key physical quantities that completely describe a given atomic level mechanism: the energy of formation, the entropy of formation, the energy of migration and the entropy of migration. As we will see, these four quantities can be calculated, with certain approximations, within the density functional theory framework.

Self-diffusion may be due to vacancies, interstitials, or direct exchange of atoms in the lattice (concerted exchange). The diffusion constant D is thus composed of three terms:

$$D = D_V + D_I + D_{CE} = \frac{C_V}{C_S} d_V + \frac{C_I}{C_S} d_I + d_{CE}. \tag{2.1}$$

d_X is the diffusivity due to mobile X, where X labels the species responsible for that component of diffusion (I for interstitial, V for vacancy and CE for the concerted exchange mechanism). That is, d_X is the diffusivity of X after it has been created in the lattice. In the case of the concerted exchange mechanism, every lattice site has the potential of undergoing this transition, so the diffusion constant due to CE is just the diffusivity of CE, d_{CE}. However, in the case of a defect, only lattice sites that neighbor the defect can diffuse via the defect mechanism. If there are no defects, then defect-mediated diffusion cannot occur. Thus, the diffusion constant due to defect X is proportional to the ratio of the concentration of lattice sites that contain the defect, C_X, to the total concentration of lattice sites, C_S. The ratio of C_X to C_S is the probability that a lattice site contains a defect of type X.

Macroscopic diffusion is related to D by Fick's first law:

$$F_X = -D_X \nabla C_X \tag{2.2}$$

where F_X is the flux of species X, D_X is the diffusion constant of species X, and C_X is the concentration of species X, where, again, X represents I, V, or CE. The flux of X is proportional, then, to both the diffusion constant and the concentration gradient

of species X. The diffusion constant is the constant of proportionality (dependent on temperature) linking the concentration gradient of a species and the flux of that species through the crystal. It can also be related to the average distance L species X travels in time t through $\langle L \rangle = \sqrt{D_X t}$.

In general, C_X and d_X can be further broken down into contributions due to different charge states of defect X. In some cases, the charge states dominate the diffusion constant [5]. However, the diffusion of neutral defects and dopant atoms will be important in certain situations and will also give a qualitative description of diffusion under other charge states. For that reason, in the work that follows, we will be concentrating on the neutral defect. Diffusion by charged defects is a topic for future work.

If we minimize the free energy of a defect X in the perfect crystal, and assume that the concentration of defects is much smaller than the concentration of lattice sites, we find that the concentration of an intrinsic defect has the form

$$C_X = C_S \exp\left(\frac{TS_f - E_f}{k_B T}\right) \tag{2.3}$$

where S_f is the entropy of formation and E_f is the energy of formation. S_f is composed of two parts, the entropy of configuration and the entropy of vibration. The entropy of configuration is the natural log of the number of ways that the defect may be placed on a given lattice site. In the case of interstitials, the interstitial atom can be placed 6 different ways on a lattice site, while a vacancy can be placed 3 different ways. These 3 placements of the vacancy correspond to the lattice with one atom removed and the four neighboring atoms undergoing Jahn-Teller distortions, which involve two bonds forming from the four dangling bonds left behind by creating the vacancy.

Fick's law is derived by examining the random hopping of an atom on a lattice. This leads to a relationship between the diffusivity d_X and the hop rate Γ:

$$d_X = \frac{1}{6} f z a^2 \Gamma = \frac{1}{6} f z a^2 \nu \exp\left(\frac{TS_m - E_m}{k_B T}\right). \tag{2.4}$$

d_X is proportional to Γ, the rate at which species X hops from site to site. The constants of proportionality are z, the number of equivalent pathways the diffusing mechanism can follow, a, the jump distance, which is approximately equal to the lattice spacing of the crystal, and the correlation factor f, which relates the diffusion of actual atoms to the diffusion of the mediating defect. Here, Γ is reexpressed in terms of the Gibbs free energy, $E_m - TS_m$, and the attempt frequency for the event, ν. S_m is the entropy of migration and E_m is the energy of migration. The factor of $1/6$ comes from the dimensionality of space in which the diffusion occurs (this factor goes as $1/2N$, where N is the dimensionality).

This equation assumes that transition state theory is valid, which is usually true if the thermal energy of the atoms is less than about one third the barrier height. There are three major assumptions in transition state theory. The electrons must always be in their ground state so that the nuclei are always on the same potential energy surface. Second, the motion of the nuclei is classical. Finally, there exists some dividing surface between two states in the potential energy landscape such that, if the system leaves one state and crosses the dividing surface, it will settle into the second state and equilibrate there before crossing the dividing surface again. This means that the system will cross the dividing surface only once between equilibrations. [6] The validity of the third assumption depends on the temperature of the system relative to the potential energy surface.

If all three assumptions are valid, then the rate for the system to move between two states is given by the ratio of the number of states at the dividing surface to the number in the initial state. Finding a good dividing surface that defines the transition state is a difficult problem. A further approximation can be made that greatly simplifies the problem and makes it computationally tractable. If the potential energy surface at both the initial state and the saddle point of the dividing surface is assumed

to be harmonic, then the rate constant Γ can be simply expressed as function of ν, S_m and E_m, as in equation 2.4, in which these quantities have mathematical definitions that depend only on the shape of the potential landscape near the critical points at zero temperature. This approximation is referred to as the harmonic approximation.

Often, the diffusion constant is written in an Arrhenius form as

$$D = D_0 \exp\left(\frac{-E_a}{k_B T}\right), \tag{2.5}$$

where D_0 is the diffusion prefactor containing all entropy terms, and E_a is the activation energy. E_a controls the primary temperature dependence of the diffusion constant, where E_a is simply $E_m + E_f$. Migration entropies might also depend on temperature, but here we will assume that they do not, that the harmonic approximation (or zero temperature limit) is valid. Experiments typically measure D as a function of T which enables an extraction of E_a and D_0

To describe diffusion completely, the four quantities S_f, S_m, E_f, and E_m need to be calculated. ν is also needed, but in the calculation for S_m in the harmonic approximation, the actual quantity calculated is $\nu \exp(S_m/k_B)$. Whenever entropy of migration is referred to in this thesis, it is this combination of quantities that is meant. The purpose of this work is to calculate each of these quantities for the processes that govern diffusion of intrinsic (interstitials and vacancies) and extrinsic (impurity atoms) defects in Si and Ge, as well as concerted exchange.

2.3 Impurity Atom Diffusion

It is well accepted that impurity atoms diffuse, at least in part, via mediating intrinsic defects. The physical picture to keep in mind here is of the defect X binding to the substitutional, but immobile, impurity atom A. This creates a AX complex, which is then able to diffuse as a unit through the crystal. If dopant A diffuses via defect X, then we can write the effective diffusion constant for A as

$$D_A^{eff} = d_{AX}\frac{C_{AX}}{C_A}. \tag{2.6}$$

In this equation, d_{AX} is the diffusivity of the combined AX complex, C_{AX}/C_A is the fraction of A that is bound to X, and is, thereby, mobile (assuming that isolated A is immobile).

We can calculate d_{AX} for the mobile complex of AX just as we can for an isolated intrinsic defect X as mentioned above and detailed below. What is the fraction C_{AX}/C_A? C_{AX} is the concentration of A that is bound, in some sense, to X. This will depend on the concentration of both X and A present in the system and will be related to those concentrations by a temperature dependent equilibrium constant, $K(T)$:

$$C_{AX} = K(T)C_AC_X \tag{2.7}$$

where the equilibrium constant is just the exponential of the binding energy E_{AX}^b of the AX complex:

$$K(T) = \frac{\Phi_{AX}}{C_S}\exp\left(\frac{E_{AX}^b}{k_BT}\right). \tag{2.8}$$

Φ_{AX} is the number of ways to put the AX complex on a lattice site. Since Si has four nearest neighbors, this is 4 for the Si lattice. C_S is the concentration of Si lattice sites, as before. Thus, we have

$$C_{AX} = \frac{\Phi_{AX}C_AC_X}{C_S}\exp\left(\frac{E_{AX}^b}{k_BT}\right) \tag{2.9}$$

and D_A^{eff} becomes

$$D_A^{eff} = d_{AX}\frac{\Phi_{AX}C_X}{C_S}\exp\left(\frac{E_{AX}^b}{k_BT}\right). \tag{2.10}$$

We know what C_X is (equation 2.3). Plugging in, we get

$$D_A^{eff} = d_{AX}\Phi_{AX} \exp\left(\frac{TS_f^X - E_f^X}{k_B T}\right) \exp\left(\frac{E_{AX}^b}{k_B T}\right). \tag{2.11}$$

So, the effective diffusion constant of A is dependent on the formation of the mediating defect X through its entropy and energy of formation adjusted by the binding energy of X to A. The two energies can be lumped together and viewed as an effective formation energy of X bound to A. There is no need to consider the formation and subsequent migration of X before it binds to A separately. The assumption here is that an equilibrium population of X is obtained on the time scale of the experiment.

It should be noted that this gives the effective diffusion constant of mobile A, where A is mobile if it is interacting with a defect X. Under non-equilibrium conditions, it may be true that X must migrate for some time before encountering A (for example, if A is buried in the substrate and X is generated at the surface). The AX complex can also dissociate, rendering A immobile again. These are dynamical questions that need to be solved via coupled differential equations. However, the mobility of A, via the formation of the AX complex, will be governed by equation 2.11.

Also, as given in equation 2.2, macroscopic diffusion of species A will be

$$F_A \equiv F_{AX} = -D_A^{eff} \nabla C_{AX} \tag{2.12}$$

and C_{AX} will be determined dynamically by the local values of C_X and C_A as well as diffusion of AX in the system. So, D_A^{eff} is the diffusion constant of AX as a complex. D_A^{eff} does not say anything about how much AX will exist in the system.

2.4 Energy of Formation

The energy of formation (E_f) is the easiest of the four quantities to calculate. Within the harmonic approximation, this calculation involves finding the minimum energy structure for the various defects, e.g. a vacancy, and comparing it to the energy of the perfect crystal.

In the case of defects in Si, the reference system is the ideal Si crystal. E_f for a defect cell containing $N \pm 1$ atoms (+1 for an interstitial and -1 for a vacancy) is

$$E_f = E_{defect} - \frac{N \pm 1}{N} E_{crystal} \tag{2.13}$$

where the ideal crystal cell contained N atoms. This assumes that the defect atom either came from (for interstitials) or went to (for vacancies) an ideal lattice site (the formation of interstitial-vacancy pairs has an activation energy that is much higher than the energy required to bring either defect from a kink site to the bulk crystal). When discussing dopant diffusion, we will assume that the dopant atom is already present, possibly due to an implantation process or a similar source, and that, for diffusion of the dopant to occur, a vacancy or interstitial must first be formed in the crystal and then bound to the dopant.

When reading the first part of this thesis, one should keep in mind that the energies of formation reported are the values obtained using plane-wave density functional theory using the generalized gradient approximation (see chapter 3). These values are typically about 1 eV lower than the experimental value. As will be shown in chapter 10, this is due to the generalized gradient approximation. A more accurate description of the exchange correlation functional improves the values for the energy of formation. These corrections and their implications on our results will be discussed more thoroughly in chapter 10.

2.5 Energy of Migration

To calculate the energy of migration, we require the transition state between the two stable structures of interest. The method employed here is the nudged elastic band (NEB) method [7]. To use this method, both the initial state and the final state of the process must be known. Finding the right final state is often a question of trial and error, unless some other method, such as the dimer method [8] is employed. In

the cases discussed here, trial and error was indeed used. If there are only two states to consider, then typically only one path needs to be calculated in these systems. However, for a system with a more complex potential energy landscape, the number of paths to be considered grows very quickly with the number of states. For example, in bulk Si, there are three important interstitial structures, and one must consider paths between each of them as well as between structures of the same type. This leads to a total of six different paths that must be calculated.

After the initial and final states have been located, a chain of states is created between the two. This chain represents the system as it moves from the initial state to the final state. The images of this chain, the actual replicas of the system displaced along the path between the two states, are connected by fictitious springs whose purpose is to keep the images equidistant. If the springs are not included, the images of the chain will slide down to one of the defining minima. To complete the algorithm, the tangent of the path defined by the chain is estimated and the component of the true force due to the atoms and their electrons parallel to the tangent is zeroed, as is the component of the spring force perpendicular to the tangent. This insures that the only forces acting perpendicular to the path – the forces that move the chain to the valley connecting the initial and final state – are the true forces of the potential, while the only forces acting to keep the images equidistant are the spring forces.

This process is illustrated in figure 2.1 for a two dimensional potential. The smaller dots (path i) represent the initial guess for the path connecting the two minima, A and B. After the NEB method has been applied, the chain of images lies along the minimum energy path (path f). The potential energy barrier between states A and B can be obtained from the curve showing energy versus position of the images. The points can be interpolated using the extra information supplied by the force parallel to the path to get an accurate value for the saddle point energy. In harmonic transition state theory, E_m is equal to the saddle point energy.

We have made some further modifications of the NEB method that allow for

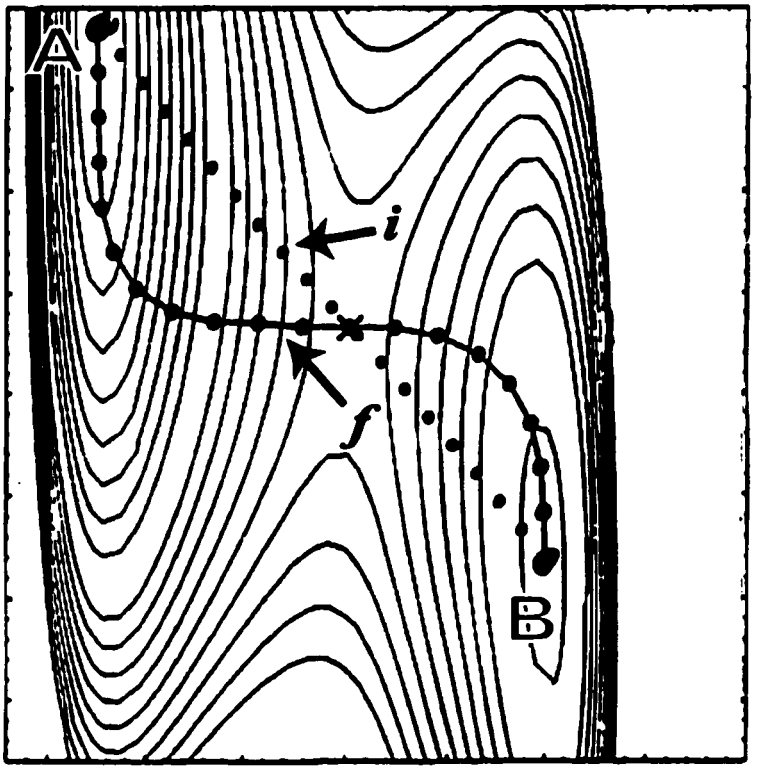


Figure 2.1 : Example of the NEB method. An initial guess of the path (i) is made between two states A and B. The NEB method is applied, and the resulting path f lies along the minimum energy path. The vertical and horizontal scales represent the two independent coordinates of this potential.

more accurate determination of the saddle point. These improvements, developed in collaboration with Graeme Henkelman, are discussed in chapter 11.

2.6 Entropy of Migration

Within harmonic transition state theory [9], the prefactor for diffusion from an initial state to a transition state is given by the ratio of the product of the normal mode frequencies of the initial state to the product for the transition state:

$$\nu^* = \frac{\prod_{j=1}^{N} \nu_j}{\prod_{j=1}^{N-1} \nu_j'} = \nu_N \frac{\prod_{j=1}^{N-1} \nu_j}{\prod_{j=1}^{N-1} \nu_j'} = \nu_N \exp\left(\Delta S / k_B\right) \tag{2.14}$$

Here, ν_j and ν_j' are the frequencies of the j^{th} normal mode of the initial state and the transition state, respectively. ν_N is the normal mode frequency in the minimum state whose eigen vector lies along the line of force connecting the minimum to the saddle point. It is the attempt frequency of the transition. Equation 2.14 also gives the relationship between the normal mode frequencies ν and the entropy of migration ΔS.

To get the full prefactor for diffusion, the exponential of the entropy needs to be multiplied by the attempt frequency, giving a factor of $\nu \exp\left(\Delta S / k_B T\right)$, which is what will be reported here as the entropy of migration.

This formulation of the entropy of migration measures the relative ease in which the system moves within the transition state hyperplane relative to the initial state. The frequencies measure the steepness of the potential well in which the system lies in both states. If the well is less steep, the frequencies will be smaller. Thus, if the system is freer to move in the transition state, the product of frequencies in that state will be smaller than for the initial state and the entropy of migration would be higher than if the system was more confined in the transition state.

Thus, to calculate the entropy of migration, one needs the normal mode frequencies in both the initial state and the transition state. The DFT code used, VASP, has been modified to make these calculations relatively straightforward. A set of degrees of freedom is selected, as is a displacement size, and each displacement is done one by one, building the force constant matrix of change in force versus displacement. This gives the matrix of second derivatives, the so-called Hessian matrix. By diagonalizing the matrix, the eigen frequencies ν_j and the eigen vectors (the normal modes) are obtained.

We have found that, to reduce the noise present in the DFT forces as well as get a better approximation to the harmonic well assumed by this procedure, several displacements must be done and the resulting force constants fitted to a harmonic function to give the final Hessian matrix. However, to get an order of magnitude estimate of the prefactor, one displacement is sufficient. Adding the second displacement changes the value typically by 10 to 20% which, while significant, does not change which mechanism has the larger prefactor. For the calculations quoted here, we have used a displacement size of 0.01 Å and, for important prefactors (ones that govern low energy pathways), we have also used a displacement of -0.01 Å. It will be noted in the text if one or two displacements were used.

The degrees of freedom included in the entropy calculation are chosen by finding the n atoms which move the most from the initial state to the transition state. These atoms govern the properties of the process more than others. More and more degrees of freedom are added until a convergence in the resulting entropy is seen. We have found that in most cases the entropy converges after 72 degrees of freedom, or 24 atoms, are included in the calculation. The entropy for some processes converges more quickly, but, in general, 72 degrees of freedom are required.

Finally, in order to reduce noise in the forces, we have found it necessary to use a high energy cutoff (188 eV for Si), as well as a tight convergence criterion for the electronic wavefunction. The actual criterion used is an energy difference between

successive electronic wavefunction minimizations of 10^{-8} eV.

One note about the forces given by VASP needs to be mentioned. For bulk Si, if an atom is displaced by 0.01 Å in one of x, y, or z, the restoring force on that atom is overwhelmingly in the opposite direction of the displacement. But there will be some component of the force in a direction perpendicular to the displacement. This component is typically less than 0.03%. If local harmonic analysis is done on all of the atoms of the system, the resulting spread of eigenvalues is less than 0.3%. Thus, because of the strong restoring force acting on the Si crystal atoms, the forces given by VASP should result in accurate prefactors.

2.7 Entropy of Formation

The entropy of a given state i for a harmonic solid is given by (e.g. [10])

$$S_\alpha = -k_B \sum_{i=1}^{N} \sum_{\beta=1}^{3} \ln \left[\sinh \left(\frac{\hbar \omega_{\beta i}}{2 k_B T} \right) \right] \tag{2.15}$$

where N is the number of atoms and the sum over β is over the degrees of freedom for each atom. The ω are related the the ν calculated before by $\omega = \nu/2\pi$. To get the entropy of formation of a defect involves finding this quantity for both the defect system and the system minus the defect:

$$S_f = S_{\text{defect}} - S_{\text{perfect-crystal}}. \tag{2.16}$$

In principle, one would have to calculate the complete normal mode spectrum of both states. One approximation is to use the harmonic frequencies, in which case the calculation of the ω is done just as for ν above. A further approximation can be made, called the local harmonic approximation, in which ω is found separately for each atom. That is, instead of displacing one atom and calculating the change in the force acting on it and all of its neighbors, only the forces on the displaced atom are included. This neglects the coupling of the vibrations of one atom from

the other atoms, which is neglected in the local harmonic approximation. Thus, N 3×3 matrices are diagonalized instead of one $3N \times 3N$ matrix. It has been shown that the local harmonic approximation can give an accurate entropy of formation [11]. However, as we will show, including the coupling to the other atoms in the system adds at least 1 k_B and we will use the full Hessian matrix to calculate the entropy. This will be referred to as the full harmonic approximation (FHA) in the text.

The basic computation necessary to find this quantity, the evaluation and diagonalization of the Hessian matrix, is the same as for the entropy of migration.

The numbers quoted in this thesis are calculated in the high temperature limit of equation 2.15. In that limit, the classical limit, the entropy of a state can be written as

$$S_\alpha = -k_B \sum_{i=1}^{3N} \ln\left(\frac{\hbar\omega_i}{k_B T}\right). \tag{2.17}$$

This equation gives values equivalent to equation 2.15 for temperatures above about 1000 K.

2.8 Conclusions

Thus, we have the framework for calculating the fundamental quantities that determine diffusion constants. The basic procedure is as follows. First, we find the stable states in the crystal. In some cases, this is simple, such as the vacancy where one lattice atom is removed. In others, such as the interstitial, various positions have to be tested to find the hierarchy of stable structures. Once stable structures have been identified and their energy calculated, comparing this energy to the chosen reference system gives the energy of formation. The entropy of formation can also be calculated now, since we have the stable structures. This involves finding the normal mode frequencies of both the perfect crystal and the defect structure and finding the difference in the entropy for each structure per equation 2.17. Once the low energy

structures have been identified, nudged elastic bands are run between them to find the minimum energy paths and the migration barrier between different states. Finally, once the transition state between two structures has been found, the entropy of migration can be calculated as detailed above.

If all of the important structures and the paths between them have been found, then we should have a complete description of diffusion on the microscopic scale. Diffusion on the macroscopic scale can then be simulated. Comparing to experiment offers some complications, mostly because of the approximations that have been made. We have made use of a harmonic description of the solid in calculating all four quantities that compose the diffusion constant. At low temperatures, this will be a good description of the Si lattice but, as the temperature is increased, the harmonic approximation will break down and a more accurate method of calculating these quantities must be used. Such methods exist, but are very costly computationally to use with DFT. However, future work will need to include a comparison between the harmonic approximation and a full free energy calculation in order to get better quantitative comparisons with experiment.

Chapter 3

DENSITY FUNCTIONAL THEORY

3.1 Background

Density functional theory (DFT) is a powerful tool for studying the properties of solids at an atomistic level. This approach is based on expressing the energy of a system as a functional of the electron density. In principle, this method can treat all of the electrons of a material quantum mechanically, while treating the atomic nuclei classically. In practice, because of computational demands, only the valence electrons are treated quantum mechanically. Also, the exact energy functional is unknown, and an approximate functional is used in actual DFT calculations.

The basic input to a DFT calculation is just the position and type of all atoms in the system. One also has to specify some information about how to treat the approximations mentioned above.

There are two theorems of DFT which make it such a useful tool. First, the energy of a system of electrons can be expressed as a functional of their density, n. Second, the density that minimizes the energy is the density of the ground state of the system.

The mathematical description of the theorems is outlined as follows. The Hamiltonian of a system of electrons in an external potential V_{ext} is given as

$$H = T + V_{ee} + \sum_{i=1}^{N} V_{ext}(\mathbf{r}_i) \tag{3.1}$$

where V_{ee} is the electron-electron interaction and T is the kinetic energy of the electrons. In a solid, the external potential is the ion coulomb potential, plus any applied fields. Defining the functional

$$F[n] = \min_{\Psi \to n} \langle \Psi | T + V_{ee} | \Psi \rangle \tag{3.2}$$

where we take the minimum over all Ψ, the antisymmetric wavefunctions of the electrons, that give the density n, we have the following two basic theorems of DFT:

$$\begin{aligned} E[n] &= \int d\mathbf{r} V_{ext}(\mathbf{r}) n(\mathbf{r}) + F[n] \geq E_{GS} \\ \int d\mathbf{r} V_{ext}(\mathbf{r}) n_{GS}(\mathbf{r}) &+ F[n_{GS}] = E_{GS} \end{aligned} \tag{3.3}$$

where GS represents the ground state of the system.

The computation is greatly simplified by expressing the wavefunction as a sum of plane waves. By using a periodic system (a repeating cell with the diffusing atom in all directions), we can exploit Bloch's theorem:

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_n(\mathbf{r}) \tag{3.4}$$

where u_n is a function with the same periodicity as the supercell. k is a wave vector representing a position in the Brillouin zone. The complete wavefunction for state n is then

$$\psi_n(\mathbf{r}) = \sum_{\mathbf{k}} \psi_{n,\mathbf{k}}(\mathbf{r}) \tag{3.5}$$

The functions $u_n(\mathbf{r})$ can be expanded in a plane wave basis set and we end up with the relationship

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{g}} c_{i,\mathbf{k}+\mathbf{g}} \exp[i(\mathbf{k} + \mathbf{g}) \cdot \mathbf{r}] \tag{3.6}$$

where $\mathbf{g}$ are the reciprocal lattice vectors of the supercell lattice.

This equation shows us the origin of the two parameters that most determine the accuracy of a DFT calculation for a given cell. In principle, the sum over $\mathbf{g}$ should

be infinite. In practice, the sum is truncated at some cutoff $\mathbf{g}_{cut}$, which is usually expressed in terms of the equivalent energy $E_{cut} = \frac{\hbar^2 \mathbf{g}_{cut}^2}{2m_e}$, where m_e is the mass of the electron. There is also the sum of $\mathbf{k}$ over the entire Brillouin zone. However, because the wavefunctions at k-points close together are nearly identical, the wavefunction can be evaluated for a special set of k-points that approximate the entire Brillouin zone. Thus, when doing a DFT calculation, both the energy cutoff and the k-point sampling must be specified.

For a more complete introduction to DFT, see references [12, 13].

3.2 Exchange-Correlation Functional

In principle, DFT is exact. However, in practice, one of the terms of the DFT energy, the exchange-correlation energy, is not known, so approximations must be made for this term. The simplest approximation is the local density approximation (LDA). It involves approximating the exchange-correlation energy for the charge density $n(\mathbf{r})$ at point $\mathbf{r}$ by the exchange-correlation energy of a uniform electron gas of density $n(\mathbf{r})$:

$$E_{XC}^{LDA}[n] = \int d\mathbf{r} n \epsilon_{XC}^{unif}[n(\mathbf{r})] \tag{3.7}$$

where $\epsilon_{XC}^{unif}[n\mathbf{r})]$ is the exchange-correlation energy per particle of a uniform electron gas of density $n\mathbf{r})$.

LDA can be improved by including gradients of the charge density. This approach, termed the generalized gradient approximation (GGA), formulates the exchange-correlation energy as

$$E_{XC}^{GGA}[n] = \int d\mathbf{r} f(n, \nabla n). \tag{3.8}$$

Here, f is some function of the density n and its derivative ∇n. There are several choices for this. The GGA used in this work is the version formulated by Perdew and Wang and is referred to as the PW91 functional [4].

As mentioned by Perdew [14], GGA functionals improve upon LDA results for such properties as total energies, atomization energies, energy barriers and structural energy differences. GGA tends to soften and expand bonds compared to LDA.

Another improvement can be made if exact exchange is included in the electron-electron interaction. These types of functionals, referred to as hybrid functionals, mix some exact exchange with the exchange of GGA. The form used in the cluster calculations described in chapter 10, referred to in the literature as B3LYP, has the form

$$E_{XC}^{hyb} = E_{XC}^{LDA} + a_0\left(E_X - E_X^{LDA}\right) + a_X\left(E_X^{GGA} - E_X^{LDA}\right) + a_C\left(E_C^{GGA} - E_C^{LDA}\right) \quad (3.9)$$

where E_X is the exact Hartree-Fock exchange energy and a_0, a_X, and a_C are parameters that are found by fitting to a data set of measured atomization energies. This functional was proposed by Becke in 1993 [15] who used values for these parameters of $a_0 = 0.20$, $a_X = 0.72$ and $a_C = 0.81$. Such hybrid functionals have been seen to give improved atomization energies, bond lengths, and vibration frequencies of most molecules, compared to GGA [16]. The form of this functional, however, makes it impossible to implement in the plane-wave periodic formulation of DFT that is used to study bulk and surface systems. It can only be used for cluster calculations.

3.3 Details of Calculations

The density functional theory calculations described in this work were conducted with the VASP (Vienna ab-initio simulation program) code [17], using ultrasoft pseudopotentials [18]. For systems where Si is the substrate material, the plane wave basis set had an energy cutoff of 188 eV, unless otherwise noted. For all bulk calculations, a cell containing 64 atoms (± 1 for defect structures) was used. The cell used for surface calculations also contained 64 atoms with vacuum spacing of 11 Å, unless pathways were being studied, in which case larger cells were used. The size of the

cell in those cases will be specified in the text. The generalized gradient approximation (GGA) exchange-correlation functional, in the form developed by Perdew and Wang [4], is used for the majority of the calculations, though, at times, local density approximation (LDA) results are quoted for comparison. (GGA and PW91 are used interchangeably in this work.) Finally, unless otherwise specified, a $2 \times 2 \times 2$ k-point sampling mesh of the Monkhorst-Pack type [19] in the irreducible Brillouin zone was used. In the case of surfaces, only 1 k-point was used in the direction perpendicular to the surface, because the simulation box is very long in that direction to incorporate vacuum spacing.

As mentioned previously in the discussion of the factors determining the diffusion constant, the NEB method [7] was used to calculate the minimum energy paths between states. It was further modified to include the climbing method (see chapter 11). The entropies of these processes were calculated within harmonic transition state theory, as described by Vineyard [9]. Part of the work involved in this research program was to implement the calculation of the entropy into VASP so that such a calculation was easy and convenient for the user. Details about how to conduct one of these calculations can be found at the following web address:

`http://www-theory.chem.washington.edu/vasp/`.

Chapter 4

SELF-DIFFUSION IN SILICON

4.1 *Introduction*

Self-diffusion of Si has been the subject of intense investigation for many years. However, it is still not known what the basic diffusion mechanism is. It is known that the activation energy for self-diffusion in bulk Si is between 4 and 5 eV (e.g. [20]). Whether diffusion is predominantly due to vacancies, self-interstitials, or the concerted exchange [21] mechanism is not certain. Early experiments concluded diffusion is dominated by interstitials above a cross-over temperature of about 1270°C and by vacancies below [22, 20, 23]. More recent experiments suggest a combination of all three mechanisms active at once, with a slightly larger contribution due to interstitials [24], or a predominance over all temperatures of the interstitial mechanism [25, 26]. Two recent papers ([26, 27]) report values for the total entropy and activation energy for both interstitial and vacancy self-diffusion. Bracht *et al.* find [26] a total diffusion constant of

$$D = 530 \exp^{(-4.75/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \tag{4.1}$$

that they separate into

$$D_I = 2980 \exp^{(-4.95/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \tag{4.2}$$

$$D_V = 0.92 \exp^{(-4.14/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1}. \tag{4.3}$$

Ural *et al.* find [27] an overall diffusion constant for Si self-diffusion of

$$D = 560 \exp^{(-4.76/k_B T)} \text{cm}^2\text{s}^{-1} \tag{4.4}$$

in very good agreement with Bracht *et al.* However, they estimate the various contributions to be

$$D_I = 149 \exp^{(-4.68/k_B T)} \text{cm}^2\text{s}^{-1} \tag{4.5}$$

$$D_V = 636 \exp^{(-4.86/k_B T)} \text{cm}^2\text{s}^{-1}. \tag{4.6}$$

These values differ greatly from the estimates of Bracht *et al.* Part of the goal of this work is to determine which of these experiments most accurately describes Si-self diffusion.

The experiments of Ural and coworkers were conducted in the temperature range of 800 to 1100°C. They injected interstitials into the Si substrate via thermal oxidation and then measured the deviations from equilibrium of the diffusion constants of P, Sb, and Si self-diffusion. Three different structures were made and anneals of varying length (1 hour to 100 hours) were performed. Profiles were measured with secondary ion mass spectroscopy (SIMS).

Bracht and his coworkers used SIMS as well to measure the profile of ^{29}Si and ^{30}Si in an epilayer of ^{28}Si grown on natural Si for anneals at temperature of 885 to 1388°C. They were able to split the total diffusion constant into interstitial and vacancy terms by relying on previous data on Zn diffusion in Si and the enhancement of Zn diffusion due to interstitials.

The analysis of both experiments is based upon some assumptions. Bracht *et al.* assume values for ν and a in equation 2.4. They also have to extract interstitial mediated diffusion of Si from difficult Zn diffusion experiments. Ural *et al.* assume that P diffuses solely by interstitials and Sb by vacancies. Both groups ignore the possibility of diffusion via concerted exchange.

There have been many theoretical studies of Si self-diffusion. In 1993, Blöchl *et al.* did density functional theory (DFT) within LDA calculations of self-diffusion constants in Si [28]. They found diffusion to be dominated by self-interstitials and were able to get the formation energy and entropy of both the interstitial and vacancy. They found a formation energy of 3.3 eV for the interstitial and 4.1 eV for the vacancy. Other calculations have concentrated on the interstitial mechanisms. Clark and Ackland also performed DFT/LDA calculations, finding three stable interstitial sites and likely paths between them [29]. Their interstitial formation energy is 2.3 eV. Kato's DFT/LDA studies of the Si self-interstitial conclude that diffusion due to self-interstitials has a barrier of 1.2 eV [30], though he used the drag method to find the barrier and this method is known to have serious flaws. Lee *et al.* did LDA calculations and found a relatively high formation energy for the interstitial of 4.1 eV and estimate a low energy barrier for interstitial migration of 0.15 eV [31]. Early (1984) Green's function calculations find a qualitatively different interstitial landscape than do the DFT/LDA calculations, predicting a different hierarchy between the various stable structures. [32]. Another early paper by Bar-Yam and Joannopoulos was the first to identify that the interstitial atom sharing a lattice site with a lattice atom is of the lowest energy [33]. Tight-binding molecular dynamics (TBMD) studies found that vacancies dominate interstitials, but also found a path for interstitial migration [34]. Finally, Nastar and coworkers used the Stillinger-Weber empirical potential to find saddle-point configurations for the self-interstitial [35].

None of these calculations have determined the four quantities determining the diffusion constant: E_f, E_m, S_f, and S_m for the three likely mechanisms. In addition, there is considerable disagreement between the various calculations. The goal of this work is to calculate all four of these quantities with a more accurate theoretical method for all of the important self-diffusion mechanisms.

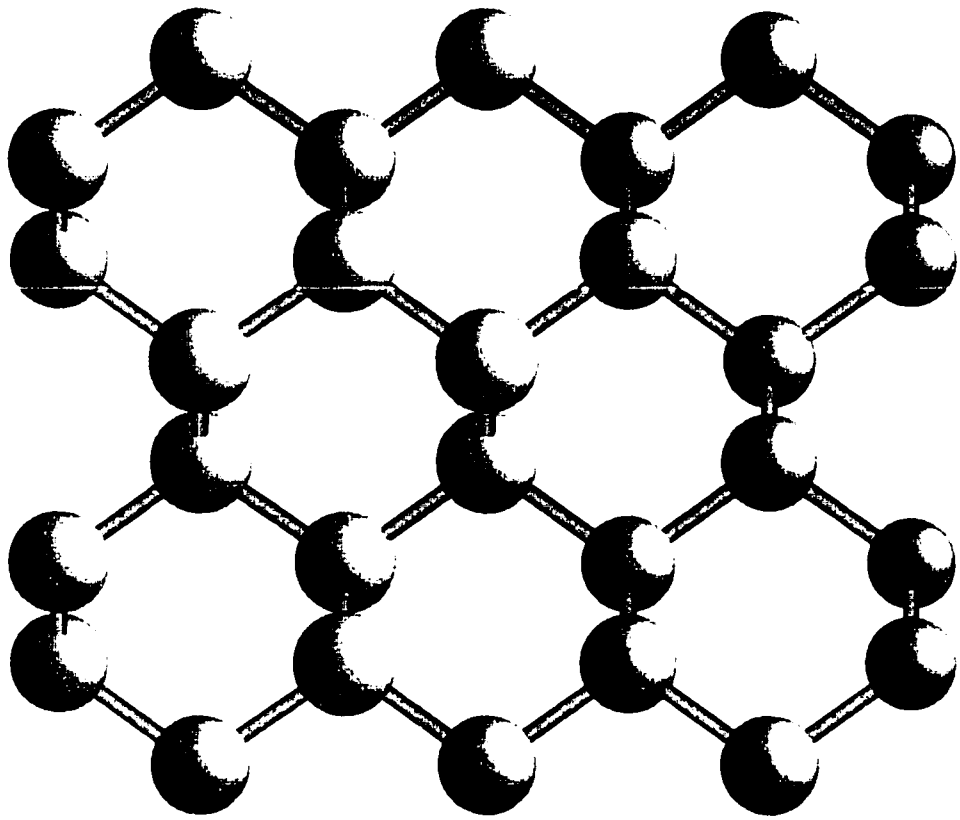


Figure 4.1 : The ⟨110⟩ projection of the perfect crystal, illustrating the 110 channels that run through the crystal.

4.2 Perfect Lattice

For a system of 64 atoms, we find that the lattice constant is 10.775 Å using LDA and 10.91 Å using GGA. For reference in interpreting the figures of defect structures and atomic pathways given later, the ⟨110⟩ projection of the perfect lattice is shown in figure 4.1.

4.3 Concerted Exchange

The nudged elastic band (NEB) [7] method was used to calculate minimum energy paths and extract activation energies for various processes. Pandey proposed the concerted exchange (CE) process in which two atoms exchange positions in the lattice in a concerted motion [21]. do the calculation completely with DFT, instead having to fit an empirical potential to DFT results. Pandey calculated a value of 4.5 eV for this process by mapping out the potential landscape under some assumptions and finding the pathway for the CE process (his final value is 4.3 eV, which is a value corrected via an estimation of long range relaxation). We use the NEB method to calculate the concerted exchange barrier for both LDA and GGA.

With LDA, we obtain a barrier height of 4.35 eV. This is slightly smaller than the value found by Pandey, which is due to our being able to relax the entire minimum energy path. Pandey was only able to relax the structure at the saddle point. One difference, however, is the presence of stable local minima on the shoulders of the minimum energy path. These shoulders have an energy of 2.74 eV and they consist of the two atoms involved in the exchange process rotated partially through the exchange. The CE barrier is shown in Figure 4.2.

With the GGA functional, the energy of the barrier is 4.57 eV. The stable side minima still exist, with an energy of 2.90 eV.

Concerted exchange involves the concerted motion of two atoms in the crystal. They first undergo a rotation about their bond center in a ⟨110⟩ plane. This is followed by another rotation in a ⟨100⟩ plane. Finally, another rotation in an equivalent but orthogonal ⟨110⟩ plane brings the two atoms back to the ideal lattice structure. This sequence of rotations is illustrated in figure 4.3. The shoulder minima occur at the points when the rotation begins in a new plane.

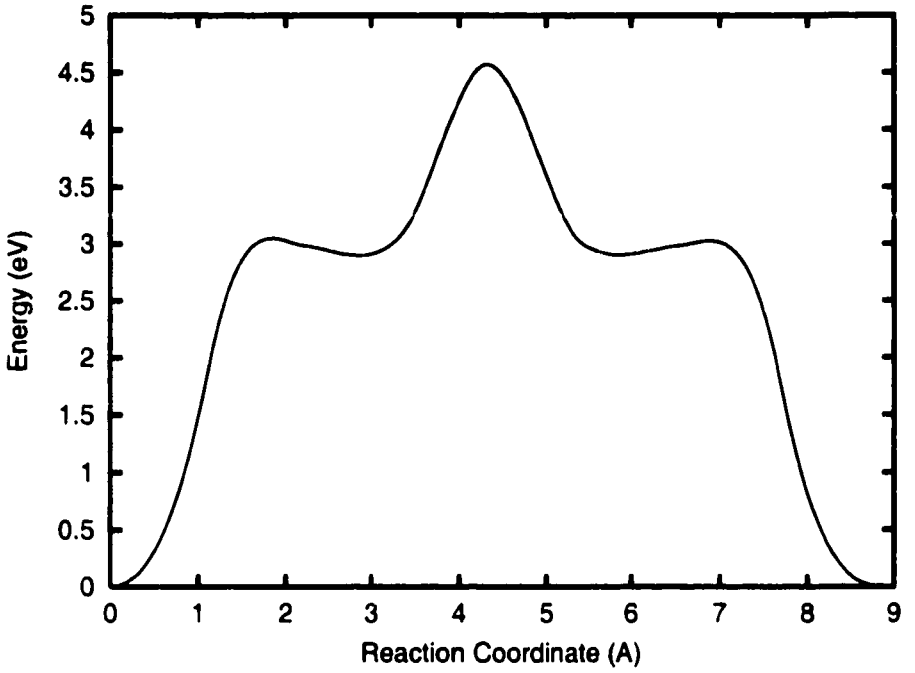


Figure 4.2 : Energy profile of the minimum energy path for the concerted exchange mechanism using the GGA energy functional.

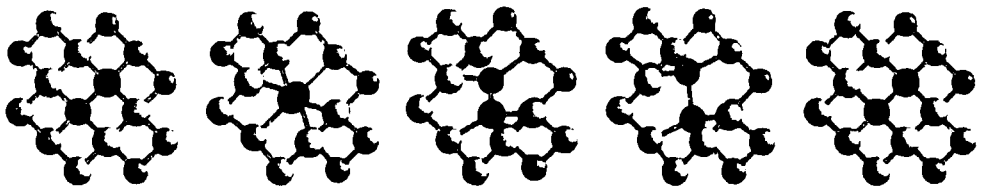


Figure 4.3 : The path in space executed by the concerted exchange mechanism. The first and last frame are ideal lattice positions, while the second and fourth frames correspond to a reaction coordinate of 2.3 Å and 6.54 Å, respectively. The third frame shows the transition state.

4.4 Self-Interstitial

4.4.1 Energy of Formation

We have found three stable interstitial structures in bulk Si: the split or dumbbell interstitial (X), the hexagonal interstitial (H), and the caged or tetrahedral interstitial (T). We calculated the relative energy of these three stable configurations and obtained the energies of 0.00, 0.07, and 0.17 eV, respectively, in the LDA and 0.00, 0.09 and 0.31 eV with GGA, where the zero of energy is taken as the split interstitial energy. This is in contrast to the study of Clark and Ackland [29], who found that X is the lowest energy interstitial with a formation energy of 2.3 eV, T is next with an energy relative to X of 0.13 eV, and H has an energy of 0.29 eV. X is also the most stable interstitial in our calculations, in agreement with Clark and Ackland. The caged interstitial is perfectly tetrahedral in our calculations, not the distorted tetrahedron reported by Clark and Ackland. Geometries of these three interstitial configurations are given in Figure 4.4 and the radial distribution function for each type of interstitial is given in Figure 4.5. A fourth interstitial configuration, found by substituting the interstitial atom into the bond between two atoms, is also stable, but has an energy of 1.37 eV in LDA.

The formation energy for the dumbbell interstitial is 3.35 eV in LDA and 3.72 eV in GGA.

It is useful to look at the electronic band structure for these different structures, as these are instrumental in determining the relative energy differences. Figure 4.6 shows the occupation of states near the band gap for the 3 stable structures mentioned above: X, H, and T. The band structure for the different types of interstitials is qualitatively different. In the case of the X interstitial, extra electronic states are created at the top of the valence band. The overall band structure is just a perturbation of the bulk structure. For the H interstitial, a state is created in the middle of the band gap of the perfect crystal, while for the T structure, states are created at the bottom

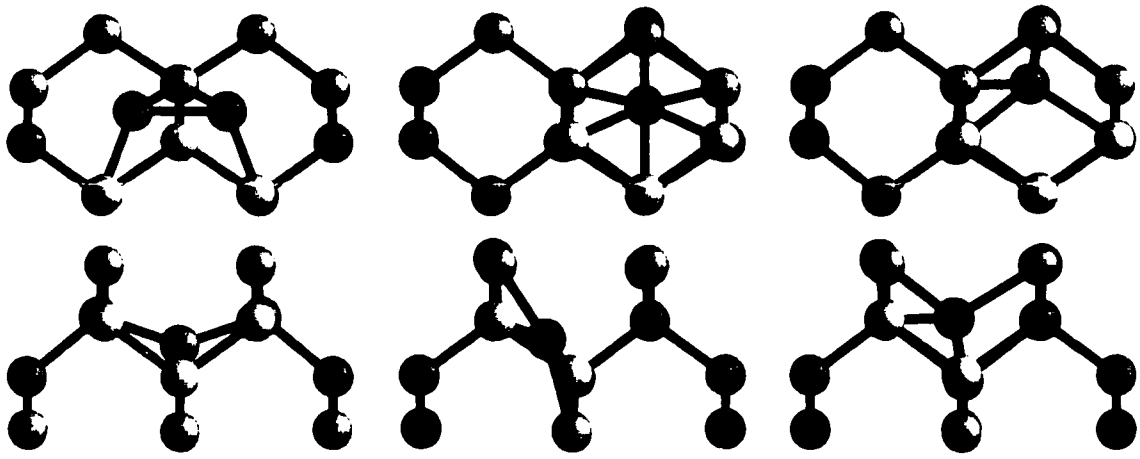


Figure 4.4 : Geometry of the Dumbbell, Hexagonal, and Caged interstitials, from opposite [110] planes.

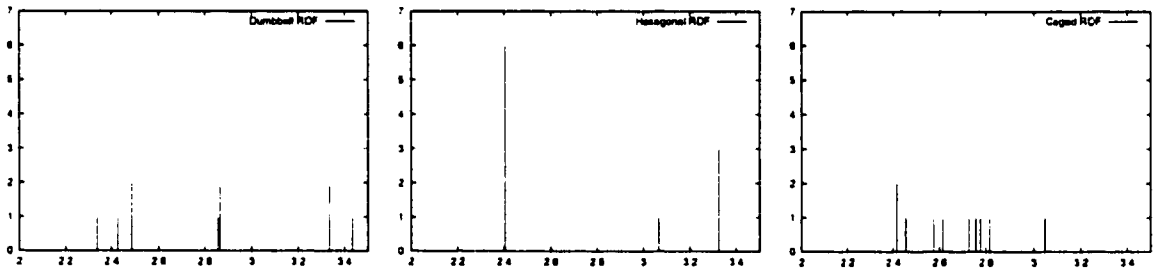


Figure 4.5 : Radial distribution functions for the Dumbbell, Hexagonal, and Caged interstitials, from the GGA calculations.

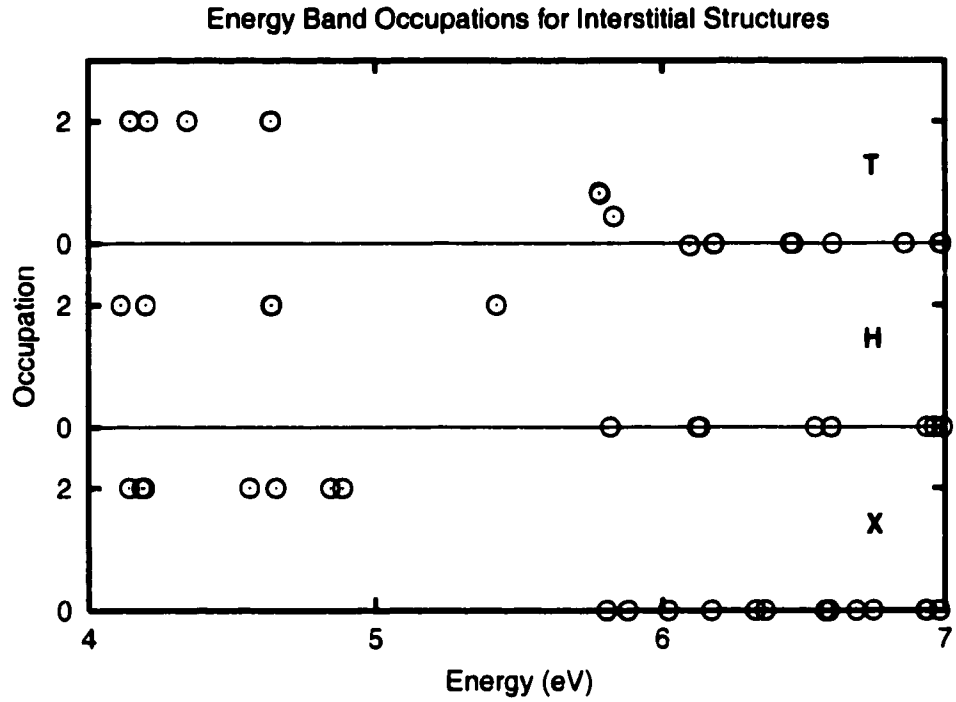


Figure 4.6 : Occupation of electronic energy states near the bandgap for the interstitial structures X, H, and T. States to the left of the gap have an occupation of 2, while states to the right of the gap have an occupation of 0.

of the conduction band and are filled, at least partially. The higher energy of the T structure compared to the X interstitial is at least partially due to higher energy electrons.

4.4.2 Energy of Migration

We used the NEB method to calculate the minimum energy paths between the three interstitial structures identified above. We used paths of 10 images, including the fixed end points.

Transitions from the dumbbell interstitial are important because this interstitial is the most stable configuration. It is the starting point for interstitial diffusion. The dumbbell to dumbbell interstitial transition (X→X) is additionally interesting

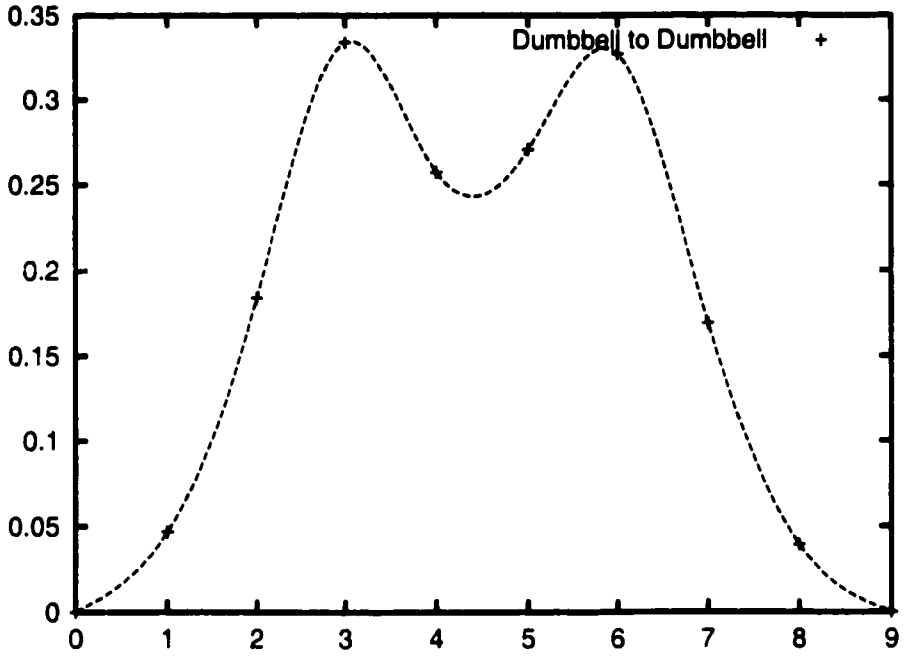


Figure 4.7 : Barrier height for an interstitial to move from one dumbbell configuration to another. The minima in energy between the two peaks is an intermediate hexagonal configuration which has not been completely relaxed in this curve. The height of the barrier is 0.33 eV in GGA, relative to the dumbbell interstitial.

because it is the one transition, as pointed out by Clark and Ackland, that changes the identity of the interstitial atom. Either atom can be viewed as the interstitial atom. One of them will move away from the lattice site during a transition with the second atom returning to the lattice site shared previously by both atoms.

We find a doubly peaked minimum energy path along the X→X transition with a hexagonal interstitial as an intermediate configuration. The barrier between the dumbbell and hexagonal interstitials is 0.33 eV both with LDA and GGA. The MEP is shown in Figure 4.7. This result agrees qualitatively with the TBMD study [34], with the intermediate caged interstitial found there being a hexagonal interstitial here. The path for the X→H motion is illustrated in figure 4.8.

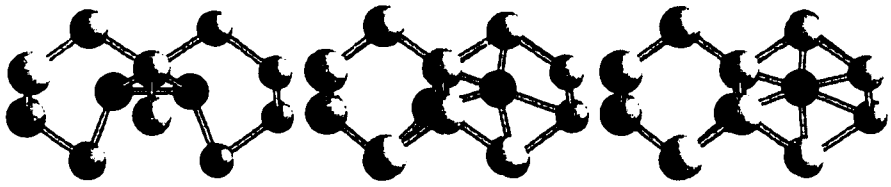


Figure 4.8 : The path taken by the interstitial atom as it moves from the dumbbell site to a hexagonal site. The first frame is the stable dumbbell site and last frame is the stable hexagonal interstitial site. The second frame is the transition state for this path.

The hexagonal to hexagonal interstitial transition proves to be an important process. An interstitial atom can transverse the entire crystal via H→H type transitions. The barrier between two hexagonal interstitials is 0.22 eV in GGA (0.09 eV in LDA). This transition state has an energy slightly lower than that for X→H transition by about 0.02 eV. The potential energy curve for the H→H process is shown in Figure 4.9. It consists of two parts. The hexagonal interstitial does not actually sit at the six-fold symmetry point in the channel. Rather, it sits off to the side (from the perspective of the ⟨110⟩ plane perpendicular to the direction of motion; see figure 4.4). There is a slight barrier for the interstitial to move from one side of the hexagonal ring to the other. This barrier, referred here as the pucker barrier (the barrier to move from one "pucker" configuration to another) is on the order of 0.01 eV. The traversal of this pucker barrier is not illustrated in figure 4.10.

The H→H transition takes the interstitial down the hex channels of the crystal (the hex channels are perpendicular to the (110) planes). In an almost straight path, the interstitial moves from one hexagonal site in a channel to the next one. In addition, each hexagonal site is at the center of two orthogonal channels, so, by traversing the hex channels, the interstitial can reach any location in the crystal. The H→H path is illustrated in Figure 4.10.

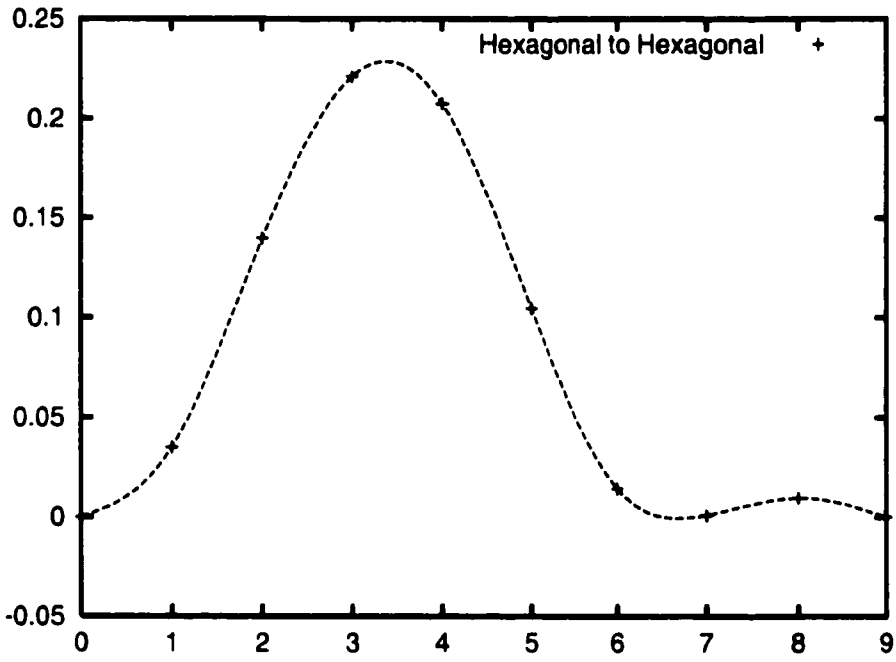


Figure 4.9 : Energy profile as the interstitial moves from a hexagonal site to another. At the end of the path, the barrier is relatively flat, as the interstitial is in a more or less hexagonal environment the entire time. The height of the barrier, relative to the hexagonal interstitial energy, is 0.221 eV in GGA (0.344 eV relative to the dumbbell interstitial, the lowest energy in this system.

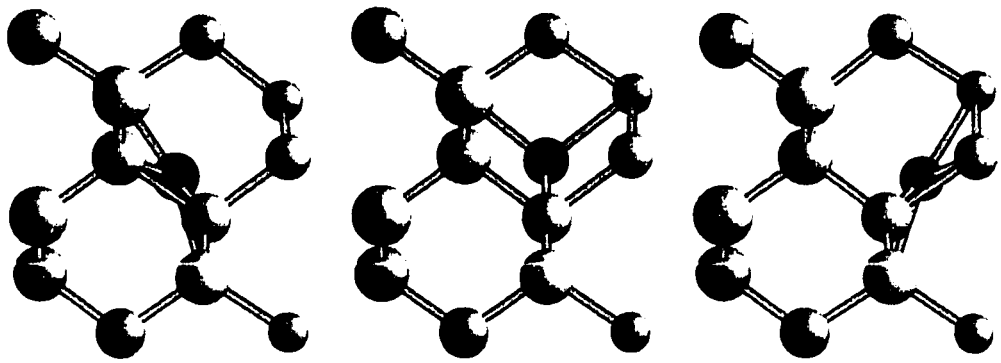


Figure 4.10 : The path taken by the interstitial atom as it moves from one hexagonal site to another. The first and last frames are stable hexagonal interstitial sites and the second frame is the transition state for this path. This process will take the interstitial down a hex channel.

The final picture for interstitial diffusion is the following: the atom starts as a dumbbell interstitial, this being the most stable site. It is excited to the hexagonal meta-stable configuration, from which it can diffuse with equal barriers to one of six neighboring hexagonal sites or one of six split sites (each of the six Si lattice atoms it is bound to is a potential site for a split interstitial). Thus, there is no preferred channel for the interstitial to diffuse by. Rather, it will jump from split to hexagonal sites with equal probability.

For both of these transition states (X→H and H→H), the structure is similar to the T interstitial structure. We have calculated more exactly the transition state of these paths (using the climbing method described in chapter 11) and table 4.1 gives the radial distribution of the atoms around the interstitial atom. While the structures are qualitatively similar and, of the stable interstitial structures found, they most resemble the T interstitial, they are not identical to the T structure. The transition state for H→H is a distorted T structure, but the X→H transition state is

Table 4.1 : Radial distribution of atoms around the interstitial atom at the transition state for both the X→H and H→H paths. For reference, the distribution of atoms around the T and T2 structures is also given.

X→H		H→H		T	
Distance (Å)	Number atoms	Distance (Å)	Number atoms	Distance (Å)	Number atoms
2.28	1	2.44	2	2.48	4
2.32	1	2.50	1	2.75	6
2.50	1	2.52	1		
2.56	1	2.62	1		
2.59	1	2.74	2		
2.83	1	2.78	2		
2.87	1	2.88	1		
3.00	1				
3.16	1				
3.32	1				

not overly tetrahedral. There is a wide range of bond lengths between the interstitial atom and its four nearest neighbors.

The overall barrier of 4.0 eV is at the lower end of the experimental range. It has recently been suggested that carbon could trap Si interstitials, effectively raising the barrier for diffusion [36]. To test this, we calculated the energy difference between a system with a Si dumbbell interstitial pair far from a C atom and a mixed Si-C dumbbell interstitial pair. The energy difference is 1.08 eV (GGA). Thus, a Si interstitial forming a dumbbell with a C impurity has an additional barrier of 1.08 eV for diffusion (assuming the transition state is the same). Thus, the presence of C could lead to an overall impediment to Si self-diffusion. However, as will be discussed,

the reason DFT/PW91 does not agree with experiment is not because of unknowns such as C trapping, but rather because of errors in DFT/PW91 itself. These will be discussed in more detail in chapter 10.

4.5 *Vacancy-Mediated Diffusion Energies of Formation and Migration*

In a collaborative effort with us, Wolfgang Windl of Motorola has calculated the formation energy and migration energy of the vacancy in Si. He has found that E_f = 3.45 eV relative to a bulk Si atom and E_m = 0.26 eV. He also did some testing of convergence against cell size, and found that, with a 216 atom cell, these values become E_f = 3.60 eV and E_m = 0.23 eV. The migration barrier is very well converged even with a 64 atom cell. The formation energy is not as well converged, but the difference is small relative to the magnitude of the formation energies. This gives an overall activation energy for vacancy diffusion of 3.83 eV.

For comparison, with LDA and a cell of 63 atoms, we find that the formation energy is 3.38 eV.

4.6 *Entropy of Formation*

The entropy of formation for a defect is composed of two parts: the vibrational entropy of formation and the configurational entropy of formation. The configurational entropy measures the number of ways that a defect can be placed on a lattice site and the vibrational entropy measures how free the defect is to move in its potential well compared to the perfect crystal.

The entropy of configuration is the natural log of the number of ways of putting a defect at a lattice site. The dumbbell interstitial sits in ⟨110⟩ and equivalent planes. Therefore, there are 6 equivalent placements for the dumbbell interstitial. For the vacancy, the surrounding atoms can distort in three different ways, creating 3 different vacancy structures (see chapter 2). This gives entropies of configuration of $1.8k_B$ and

$1.1k_B$, respectively.

The previous calculation for the vibrational entropy of formation of both intrinsic Si defects was done by Blöchl *et al.* [28]. They used the local harmonic approximation (LHA) to find the vibrational entropy of formation of both the interstitial and the vacancy using the LDA exchange-correlation functional. We have repeated his calculation using GGA and also using the full harmonic approximation (FHA). That is, we find the normal mode frequencies for the entire system at once. We do not decouple the vibrations of atoms from one another. These results are summarized in table 4.2. We can only include up to 30 atoms in our calculation as any more atoms included will begin to see the vacancy or interstitial in the neighboring cell. We find that the LHA calculations for both the vacancy and interstitial converge with 30 atoms included in the calculation. The FHA calculations converge more slowly, but we see the interstitial entropy changing by less than $.1k_B$ when the last shell of 12 atoms is added. For the vacancy, the entropy changes by $0.45k_B$ when this shell of atoms is added. However, if an extrapolation based on an exponential form is done, we would get an "infinite" limit of $5.43k_B$, which is not very different than the value we obtain for 28 atoms.

Smargiassi and Car [37] calculated the anharmonic contributions to the entropy of formation of the vacancy in Si. They find that, at low temperatures (500 K), there is no anharmonic contribution and the harmonic approximation gives accurate results. However, at higher temperatures, such as 1000 K, the anharmonic contribution is $2 \pm 1k_B$. Thus, anharmonic contributions to the entropies of formation of Si defects will be important. These will need to be calculated more accurately in the future.

Thus, we find vibrational entropies of formation of $3.56k_B$ for the interstitial and $5.43k_B$ for the vacancy, using the full harmonic values for the eigen-frequencies of each state.

Adding the entropy of configuration, we obtain $5.4k_B$ for the split interstitial and $6.5k_B$ for the vacancy.

Table 4.2 : Entropy of formation for the interstitial and the vacancy using both the local harmonic approximation (LHA) and the full harmonic approximation (FHA). Also given are the results of Blöchl's calculations.

Interstitial				Vacancy			
Number of	Entropy (k_B)			Number of	Entropy (k_B)		
atoms included	Blöchl	LHA	FHA	atoms included	Blöchl	LHA	FHA
2	1.76	1.55	1.87				
6	3.16	2.28	2.66	4	2.72	3.80	3.12
18	3.72	2.73	3.46	16	3.84	4.69	4.82
30	3.9	2.67	3.52	28	–	4.64	5.27
Extrapolation		2.71	3.56	Extrapolation		4.67	5.43

4.7 Entropy of Migration

The processes considered for the calculation of the entropy of migration were X→H, H→H, V→V, and CE. The entropy of migration calculated for these processes exhibit a range of over 2 orders of magnitude. The CE mechanism exhibits the largest S_m, with a value of 2.4×10^{14} s^{-1}. The entropy of migration for the other processes considered are listed in table 4.3. The total entropy for the H→H process should be the same as for the X→H process, so that adjusts our value for the entropy of formation for the H interstitial accordingly.

4.8 Diffusion Constant

Combining all of the various components to the diffusion constant gives the following values for the interstitial and vacancy (this includes a correlation factor of 0.5 for the vacancy):

Table 4.3 : Prefactors for various self-diffusion processes in Si, calculated using DFT/PW91.

Process	Prefactor (s^{-1})	S_m (k_B)	S_f (k_B)	S_c (k_B)	Total S (k_B)
X→H	6.5×10^{12}	0.5	3.6	1.8	5.9
H→H	4×10^{14}	2.3	1.8	1.8	5.9
V→V	1.8×10^{13}	1.8	5.4	1.1	8.3
CE	2.4×10^{14}	4.0	—	1.8	5.8

$$D_I = 0.13\exp^{(-4.0/k_BT)}\,\mathrm{cm^2s^{-1}} \quad (4.7)$$

$$D_V = 2.3\exp^{(-3.8/k_BT)}\,\mathrm{cm^2s^{-1}} \quad (4.8)$$

$$D_{CE} = 0.54\exp^{(-4.6/k_BT)}\,\mathrm{cm^2s^{-1}}. \quad (4.9)$$

These values are calculated assuming that $a = 1.2$ Å for the interstitial and $a = 2.36$ Å for the vacancy and concerted exchange (these hop lengths are taken from the actually minimum energy paths calculated). z is assumed to be equal to 4.

The prefactors are much smaller than the values obtained by Ural and coworkers. They found 149 and 636 $\mathrm{cm^2s^{-1}}$ for the interstitial and vacancy, respectively, while we find 0.13 and 2.3 $\mathrm{cm^2s^{-1}}$. Our ratio of interstitial prefactor to vacancy prefactor is about 0.06, while Ural gets 0.23. This is in contrast to Bracht's result of 3239 for this ratio. In addition, the total entropy for the migration of the two defects is closer to what Ural reports. Ural finds $S_I = 10.2k_B$ and $S_V = 12.8k_B$ while we find $S_I = 5.9k_B$ and $S_V = 8.3k_B$. The total entropy Bracht reports are $S_I = 13.2k_B$ and $S_V = 5.5k_B$. The entropy of the two defects, according to Bracht, are very different from one another, disagreeing with what Ural found and what we have calculated. Our numbers are smaller than those found by Ural by about $4k_B$, but we find, like

they do, that the vacancy has a higher total entropy than does the interstitial by about $2k_B$. Further testing needs to be done in order to determine the importance of anharmonic contributions to the entropy of migration and formation. This is the subject of future work.

Our PW91 activation energies are about 0.5 to 1 eV lower than the reported experimental values. One recent paper by Leung *et al.* suggests that this discrepancy is due to the error in how DFT treats the exchange-correlation energy of the electrons in both LDA and GGA [38]. They use diffusion quantum Monte Carlo (DMC) to solve the exact Schrödinger equation. This involves recasting the Schrödinger equation by changing real time into imaginary time, which makes the Schrödinger equation look like a diffusion equation. If the equation is then evolved in time for some initial guess, the long time limit will give the solution for the ground state of the Schrödinger equation. This is, in principle, exact, but to be computational efficient, some approximations are made. With the DMC technique, they find that the energy of formation of the split (X) interstitial is 4.96 eV. They also find that the H interstitial is lower in energy by about 0.14 eV, which disagrees with our DFT results here. It should be noted that they get the hexagonal interstitial as the lowest energy interstitial even with GGA. The discrepancy might be due to cell size or because they did not do a full relaxation of the LDA results with GGA. [1] Adding the migration barrier, their total activation energy is over 5.0 eV, significantly higher than the findings by Ural of 4.68 eV for interstitial activation energy. The result of Leung is an upper bound, as they used LDA geometries in the DMC calculation. We have estimated the error in the GGA formation energy in a different way and report on that in chapter 10.

Even considering this error in the total activation energy, we find that the acti-

[1] It is not completely clear how their DFT calculations differ from ours, as they did not give many details. It can be inferred that they only calculated the GGA energy of the LDA structures but did not fully relax them with GGA. This is, at least, what they did for the DMC calculation, which then gives upper bounds for the activation energy. They will report the details of the DFT results in a future publication.

vation energy for both vacancy and interstitial diffusion are similar with that of the vacancy being 0.2 eV lower than the interstitial. This, too, agrees more closely with Ural *et al.*, who find a difference between the two activation energies of 0.2 eV, though they find the vacancy to have a higher activation energy, as opposed to Bracht and coworkers, who find a difference of 0.8 eV.

If we look at the contributions of each mechanism at a given temperature, say 1050°C, we find that

$$f_I = 0.06, \tag{4.10}$$

$$f_V = 0.94, \tag{4.11}$$

$$f_{CE} = 2 \times 10^{-4}. \tag{4.12}$$

At this temperature, we find that self-diffusion in Si is dominated by the vacancy mechanism. Not only is the prefactor for vacancy diffusion higher than for interstitial diffusion, but the activation energy is smaller, favoring the vacancy at all temperatures. This is in contrast to the results of Ural *et al.*, who find that the activation energy of interstitial diffusion is smaller than for vacancy diffusion. However, as mentioned, DFT has trouble finding the energy of formation of these defects. This will be discussed further in chapter 10.

4.9 Discussion

In several respects, our results disagree with previous theoretical work. In early work, conducted in 1984, both Bar-Yam and Joannopoulus and Car *et al.* did DFT studies of the Si self-interstitial. Because of computational limitations, they had to use small cells or not allow all atoms to relax. As a result, both efforts found the bond-center interstitial to be the lowest energy, which we do not find at all. In addition, they find the X interstitial to be of high energy and thus unimportant, also contrary to what we have found.

Blöchl possibly did the most complete study previous to this one. He found formation energies of 3.3 and 4.1 eV for the interstitial and vacancy, respectively, using LDA. The interstitial energy is identical to what we find. For the vacancy, we found an LDA formation energy of only 3.38 eV, much lower than the 4.1 eV found by Blöchl. Our entropy of formation in the local harmonic approximation agrees well with the results of Blöchl. The disagreement there is likely due to his use of LDA and our use of GGA to find the LHA values of S_f. However, we see that the FHA gives higher values for S_f by about $0.8k_B$ for both the vacancy and the interstitial compared to LHA.

Clark and Ackland find very different values for the energy of formation of the interstitial. They find a lowest energy interstitial structure of 2.16 eV. They only use a 4 k-point set to sample the Brillouin zone and this is the biggest difference from our calculation. They also used a different type of pseudopotential. However, the reason for the disagreement between the value found by Clark and Ackland and that found by both Blöchl *et al.* and us is not known.

Lee *et al.* conducted DFT calculations using LDA of the self-interstitial and find a high formation energy, 4.1 eV, for the split interstitial. They do find that it is the most stable structure, in agreement with our calculations. Their migration barrier for X→H→X is 0.15 eV, which is about half of the 0.33 eV we find with LDA. They estimate the barrier by assuming that the T interstitial is the transition state. We agree with their LDA value for the energy of the T interstitial, but we find an LDA barrier that is higher than the T interstitial energy. In addition, we find that the T interstitial has an energy higher with GGA than with LDA. They used an energy cutoff of 163 eV, compared to our 188 eV. They did use a larger k-point mesh than we did (using 3×3×3), but did not use the special scheme of Monkhorst and Pack that picks a more optimal mesh for a given number of points. In any case, it is unclear where the disagreement in formation energy lies.

Finally, Kato also performed a DFT/LDA study of the self-interstitial and found

a qualitatively different landscape from us, with the hexagonal interstitial as the most stable structure. The dumbbell or split interstitial was found to be 1.0 eV higher in energy. Kato used smaller cells of only 17 and 33 atoms. Only 2 k-points were used for the Brillouin zone sampling and LDA was used for exchange-correlation. This leads to a qualitatively very different landscape than we have found, with migration barriers of over 1 eV, compared to our barriers of about 0.3 eV.

4.10 Conclusions

We have been able to calculate the diffusion constant of Si self-diffusion due to both interstitials and vacancies as well as the concerted exchange mechanism. Our values for the diffusion constant prefactor agree qualitatively well with the experimental findings of Ural and coworkers, which differ significantly from the results of Bracht *et al.*, in the sense that we get activation energies for interstitial and vacancy diffusion that are comparable and a prefactor for vacancy-mediated diffusion that is several times that of interstitial-mediated diffusion.

This illustrates the need for accurate calculations. In this case of Si self-diffusion, the experimental data is varied with no clear consensus even amongst the most recent work. DFT can be used to elucidate which processes are qualitatively more likely and distinguish between several experimental results. This can help clarify which parameters should be used for the diffusion constant in larger scale simulations.

Chapter 5

DIFFUSION OF GERMANIUM IN BULK SILICON

5.1 Introduction

Si is an indirect bandgap semiconductor, which makes bulk Si unsuitable to use as the basis of optoelectronic devices. However, it is the material that is most thoroughly understood and has the most support in the semi-conductor industry. Thus, it would be beneficial to find some way to create an optoelectronic device using Si as the substrate material. It has been shown that Si/Ge superlattices and alloys would exhibit a direct bandgap, allowing the creation of Si based optoelectronic devices [39]. For high performance, the thickness of the superlattice structure needs to be controlled to within a monolayer. This kind of precision requires a good understanding of how Ge diffuses both in bulk Si and into the Si surface. This chapter will focus on the diffusion mechanisms of Ge in bulk Si, looking at the three components that could possibly contribute to diffusion: vacancy, interstitial, and concerted exchange. Surface diffusion is considered in chapter 6.

5.2 Previous Work

The prospect of direct band gap devices from a Si-Ge superlattice structure has focused attention on the diffusion of Ge in Si and vice versa. Most of this attention has concentrated on interdiffusion at the interface, which may or may not be indicative of bulk diffusion, as the presence of strain and defects may be very different at an evolving surface/interface than in the bulk.

Frank *et al.* summarize the data available in 1984 and report that the activation

enthalpies found by various experimental techniques range from 3.92 to 5.33 eV, though the 3.92 eV value seems to be anamolous, as the next highest value found is 4.7 eV [20]. This data puts the activation enthalpy at right around 5.0 eV, on average. The low value of 3.92 eV is speculated to be due to either a crossover to a new mechanism at lower temperature, or the result of oxygen defects in the crystal [20]. Measured diffusion prefactors range from 1500 to 6.3×10^{5} cm^2s^{-1}, with the experiment corresponding to 3.92 eV activation energy reporting a value of 0.35 cm^2s^{-1}.

The experiment giving 3.92 eV at lower temperatures and 4.97 eV at higher temperatures suggests a crossover temperature of 1050°C. This temperature was examined more closely by Fahey *et al.* [40], who used SIMS to determine that, at this temperature, between 30% and 40% of Ge diffusion in Si occurs via an interstitial mechanism, and the remaining 60% to 70% occurs by a vacancy mechanism. As they only study one temperature, they were unable to extract activation energies or diffusion prefactors. However, they do find the diffusivity for Ge in Si at 1050°C to be 5.3×10^{-16} cm^2s^{-1}, which agrees well with the result of 6×10^{-16} cm^2s^{-1} found by Södervall [41].

5.3 Concerted Exchange

5.3.1 Energy

Pandey first proposed the concerted exchange process as a possible pathway for Si self-diffusion [21]. There is no reason, a priori, to assume that such a process will not be active for Ge diffusion in Si (labeled here as Ge_{CE}). We have calculated the energy barrier for CE of Ge in Si. The profile of the potential energy vs reaction coordinate curve for Ge exchanging with Si is very similar to that of Si exchanging with Si. The curve exhibits side minima that represent a stable state of partially rotated atoms. The height of the barrier for exchange is found to be 4.54 eV, with the side minima having an energy of 2.93 eV and a barrier between them and substitutional Ge (Ge_S)

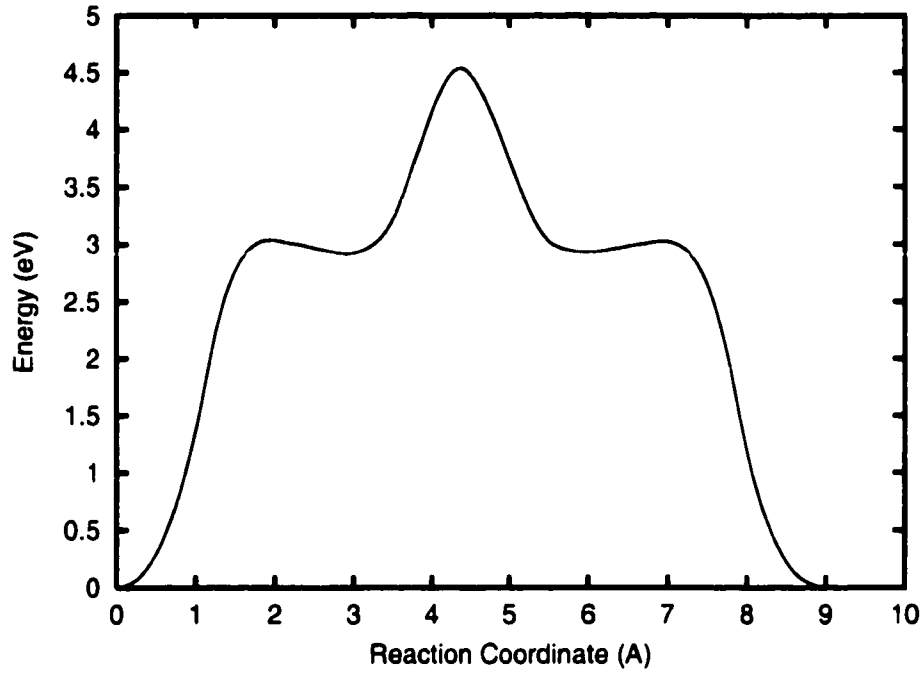


Figure 5.1 : Energy profile of the minimum energy path for the concerted exchange mechanism between Ge and Si in the Si lattice.

of 3.04 eV (these numbers are all relative Ge_S). The potential energy curve is given in figure 5.1 and the minimum energy path is given in figure 5.2.

5.3.2 Entropy

We calculated the entropy of migration for the Ge_{CE} process via the techniques outlined in chapter 2. The prefactor of migration for the concerted exchange of Ge_S and Si_S is quite large, 3.1×10^{14} s^{-1}. Only one displacement of 0.01 Å was used for this calculation.

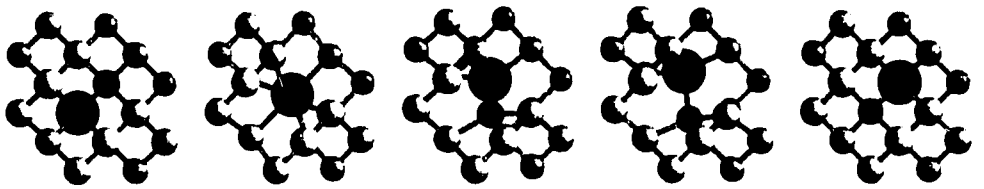


Figure 5.2 : The path in space executed by the concerted exchange mechanism. The first and last frame are ideal lattice positions, while the second and fourth are when the system is between the shoulder barrier and minimum on each side. The third frame is the system at the transition state. The larger atom is the Ge atom and the other dark atom is the Si atom involved in the transition.

5.4 Interstitial Mediated Diffusion of Ge in Si

5.4.1 Formation Energy

The energy of formation of the Ge_S+Si_I pair is just the energy of formation of Si_I in pure Si minus any binding energy of Si_I to Ge_S. This binding energy is about .11 eV. With GGA, as discussed above, the formation energy of Si_I is 3.72 eV, giving a formation energy of the Ge_S+Si_I pair of 3.61 eV.

As with the Si interstitial (Si_I), the Ge interstitial (Ge_I) exhibits a relatively large set of stable configurations. We have identified three stable geometries for the neutral interstitial (only neutral species have been considered at this time): the split (Ge_I^X), the hexagonal (Ge_I^H), and the tetrahedral (Ge_I^T) configurations. The relative energies of these structures are 0.00 eV, 0.42 eV and 0.44 eV where the zero of energy is the energy of Ge_I^X. In addition, Ge_I^H comes in two varieties: one, at 0.42 eV, is distorted, with the Ge lying to one side of the hexagonal channels existing in Si, and another in which the Ge interstitial is at the center of the channel, with an energy of 0.46 eV. The structure of Ge_I^X, Ge_I^T, and Ge_I^H (the distorted version) are illustrated in figure 5.3.

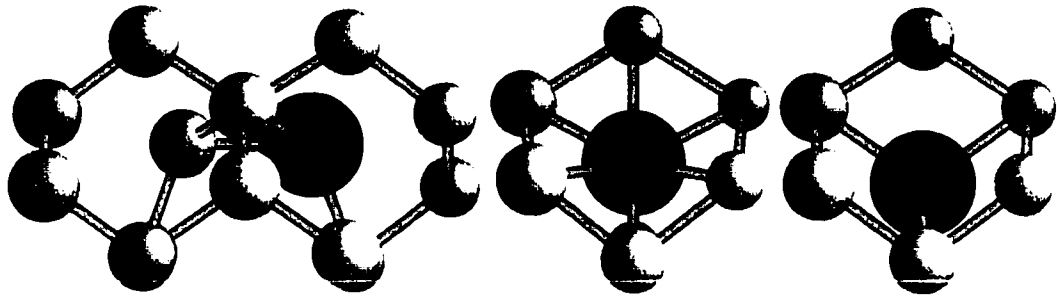


Figure 5.3 : Pure Ge interstitial structures (in which the interstitial species is Ge). The structures illustrated are the split interstitial (Ge_I^X), the hexagonal interstitial (Ge_I^H), and the tetrahedral interstitial (Ge_I^T).

The situation is made more complex if one considers Si_I structures with a nearby Ge_S. A Si split (or dumbbell) interstitial (Si_I^X) in which both of the Si_I atoms are bonded to Ge_S has an energy of 0.16 eV relative to Ge_I^X. If the Ge_S is in a lattice site such that only one of Si_I is bonded to it, the energy is 0.10 eV. This is about the same as Si_I^X far away from Ge_S, as is described below. A Si H-interstitial with a Ge_S neighbor has an energy of 0.24 eV (labeled here as Si_{H1}) while the same structure with Ge_S one ring away has an energy of 0.20 eV (Si_{H2}). Thus, as the Si_I^H moves away from Ge_S, it becomes more stable. Far from Ge_S, the most stable Si_I structure is Si_I^X, which has an energy of 0.11 eV (the previously mentioned binding energy of Si_I to Ge_S to form Ge_I^X).

5.4.2 Migration Energy

The energy barrier to move from a split Ge-Si interstitial (Ge_I^X) to a Ge hexagonal interstitial (Ge_I^H) is 0.53 eV. The potential energy curve for this process is given in figure 5.4. The minimum energy path is very similar to that seen in figure 4.8 for Si, except now the atom that moves into the hexagonal site is a Ge atom. To then move from one Ge_I^H to a neighboring Ge_I^H site, a barrier of 0.05 eV has to be overcome,

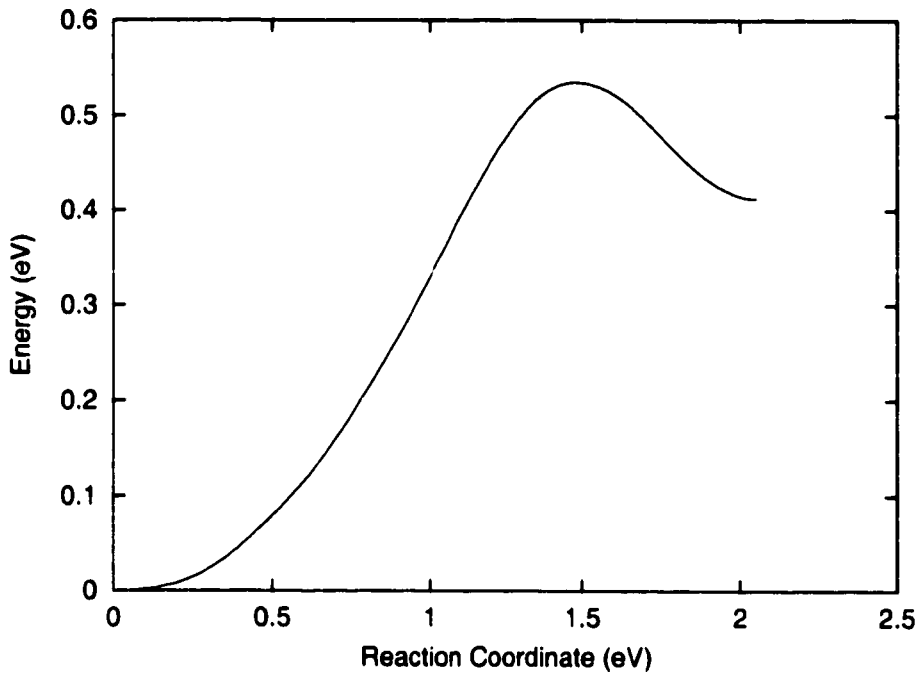


Figure 5.4 : Potential energy versus reaction coordinate curve for Ge_I to move from an X site to an H site.

which is 0.47 eV relative to the Ge_I^X structure. The highest energy point for Ge_I^H motion corresponds to the pucker barrier we saw in chapter 4 for Si_I^H motion, where it was actually a negligible barrier for Si_I^H motion. The path for this motion is shown in figure 5.5 and the barrier for complete Ge_I^H motion down a channel is shown in figure 5.6.

Thus, once a Ge atom is able to make it into a hexagonal channel, the barrier to just diffuse along one of those channels is very small. Unlike in the pure Si case, where the barrier for Si_I^H diffusion is biggest as the interstitial atom moves through the open space of the channel, here the larger barrier is when Ge_I^H passes through the hexagonal ring formed by the six Si atoms in the lattice.

The barrier for Ge_I^X to Si_I^H is 0.47 eV (both the energy curve and the path through space are similar to what we saw for $Ge_I^X \rightarrow Ge_I^H$ above). Not only is the formation

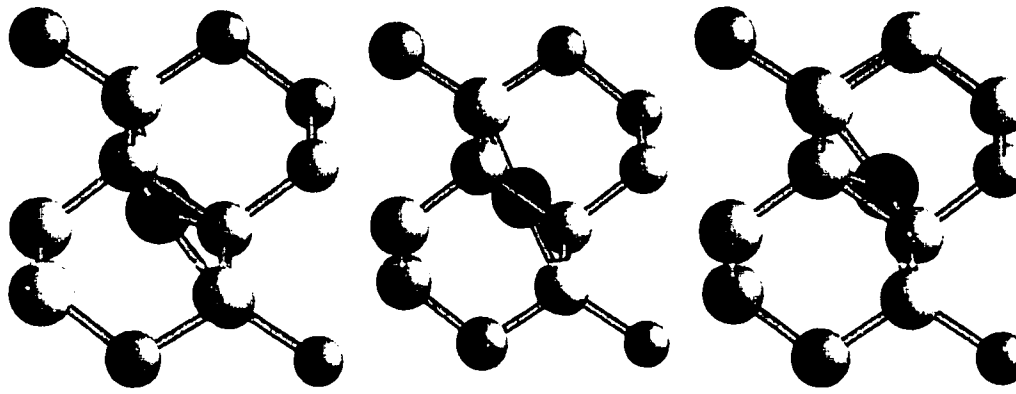


Figure 5.5 : Path in space for Ge_I^H to traverse the "pucker" barrier. The Ge atom is the dark atom. It is smaller in these images, relative to other images, so that the structure can be more easily seen.

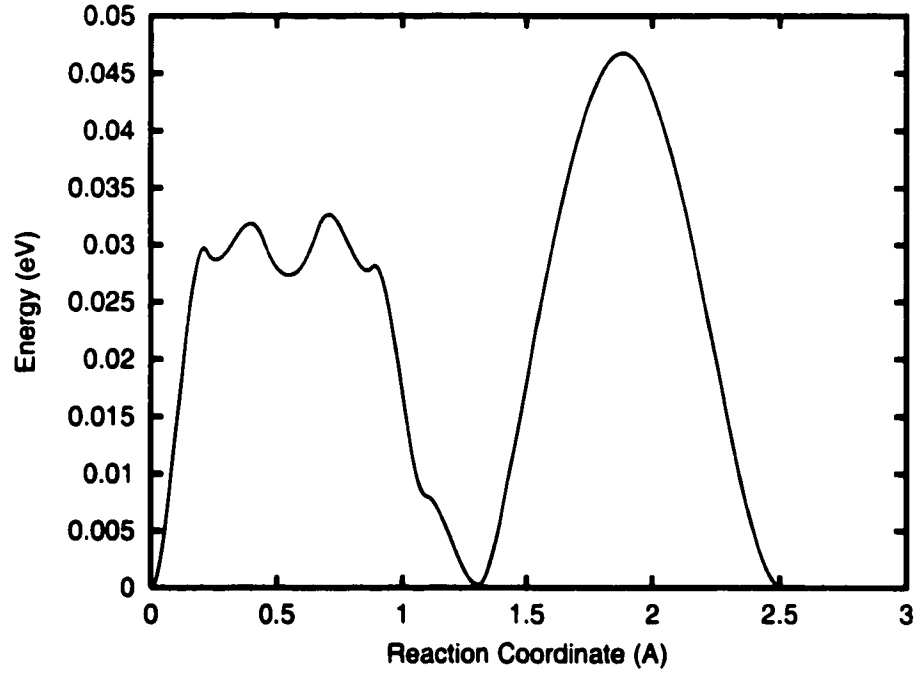


Figure 5.6 : Potential energy versus reaction coordinate curve for Ge_I^H motion in a channel in the Si crystal. This is the most complicated barrier found, with the myriad of extrema. The highest point on the curve is about 0.05 eV.

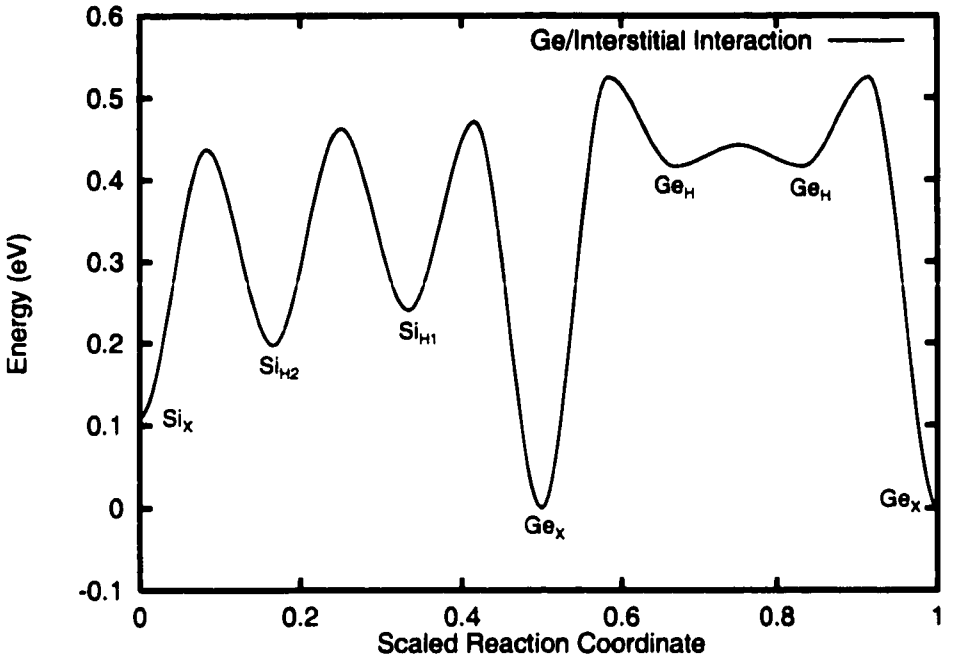


Figure 5.7 : Interaction of Ge_S and Si_I. The deepest minimum corresponds to Ge_I^X. To the right, the interstitial species is Ge and, to the left, Si.

energy for Si_I^H smaller than for Ge_I^H, the energy barrier for migration is also smaller (though, this difference is small). As Si_I moves away from Ge_S, the energy barrier decreases slightly just as did the formation energy. The barrier for Ge_I^X moving to Si_I^H is 0.47 eV, and for the Si_I^H to move one further site away it is 0.46 eV. The complete interstitial landscape is shown in figure 5.7.

One way to view this is that Ge_S traps Si_I. Ge doping is thus predicted to slow down the component of Si self-diffusion due to interstitials. Ge_S also slightly repels Si_I, with both the energy barrier and the energy of the stable Si_I going up as Si_I approaches Ge_S. However, the effect is very small. The energy of the stable interstitial near Ge_S increases by 0.04 eV and the barrier height for Si_I motion near Ge_S is increased by less than 0.04 eV compared to Si_I motion in the pure Si crystal.

The overall activation energy in the DFT/PW91 calculations is 4.07 eV.

5.4.3 Entropy of Migration

The important path for Ge_I diffusion through Si is via the Ge_I^H structure. The prefactor for migration for $Ge_I^H \rightarrow Ge_I^H$ has been calculated, and the value we have found is 1.2×10^{12} s^{-1}. (Only one displacement of 0.01 Å was used to calculate this prefactor.) This was found by the same methods as described in chapter 2.

5.5 Vacancy Mediated Diffusion of Ge in Si

5.5.1 Binding to substitutional Ge

As opposed to the case of pure Si, in which the position of the vacancy is equivalent for all lattice sites, when a substitutional Ge is added to the lattice, it breaks the equivalence of the surrounding lattice sites. The vacancy will be bound more strongly to sites closer to Ge_S. In addition, if we are trying to calculate the contribution to net diffusion of Ge caused by a particular vacancy, Ge_S has to be displaced by Si_V from opposite sides. The reason for this is that, if Si_V does not approach from a different direction than the initial displacement, but still remains bound to Ge_S, the two species will simply oscillate back and forth with no net diffusion of Ge_S. Thus, Si_V must approach Ge_S from another direction to cause net diffusion. Because of the structure of the Si lattice, one can see that Si_V actually has to move to the third nearest neighbor site in order to be able to approach Ge_S from a different direction. This means that the relevant barrier for Ge diffusion via vacancies is the barrier for Si_V to move to the third nearest neighbor site [42]. This is what we have calculated below.

This is illustrated in figure 5.8 for a square lattice. In this case, if the vacancy is bound so strongly to the dopant atom (the darker atom in the figure) that it cannot beyond a nearest neighbor position of the atom (sites labeled by "1"), then it cannot cause net diffusion of the dopant atom. It must move to the second nearest site ("2") to cause net diffusion. Similarly, in the Si lattice, the vacancy must move to the third

nearest neighbor site to cause net diffusion of a dopant atom. Otherwise, the vacancy and dopant atom will just wander within two lattice sites of the original position.

5.5.2 Formation Energy

The energy landscape for Si_V near Ge_S has been calculated to the third nearest neighbor. In the following discussion, the position of Si_V relative to Ge_S has been labeled as nn for the nearest neighbor site, 2nn for the second nearest neighbor, and 3nn for the third nearest neighbor site.

The energies, then, of Si_V near Ge_S are: nearest neighbor (nn): 0.00 eV, 2nn: 0.26 eV, and 3nn: 0.29 eV. There is a trend of higher binding energy of Si_V to Ge_S as the distance between the two decreases. The energy for Si_V far from Ge_S is 0.30 eV, which is very close to the energy of Si_V in the 3nn site, meaning that Si_V is then practically free of Ge_S.

In pure Si, Si_V has a formation energy of 3.60 eV. Si_V binds to Ge_S with an energy of 0.30 eV, trapping it slightly. Thus, this complex has a formation energy of 3.30 eV.

5.5.3 Migration Energy

The barrier for Si_V to move through the crystal near Ge_S was calculated using the NEB method as described before (chapter 2). The barrier for Si_V to exchange with Ge_S is 0.17 eV. To move from the nn site to 2nn, the barrier is 0.43 eV and to move from 2nn to 3nn, it is 0.35 eV (or 0.61 eV relative to nn). The effective barrier height, relative to nn, increases with distance. Si_V is effectively unbound to Ge_S after it reaches the 3nn site and the subsequent barrier for migration is 0.23 eV (see section 4.5). The complete potential energy curve is shown in figure 5.9.

The path for Si_V-Ge_S exchange is shown in figure 5.5.3. The Ge atom moves directly from one lattice site to the next. The mechanism responsible for the overall energy barrier for vacancy-mediated diffusion of Ge – the 2nn to 3nn barrier – is

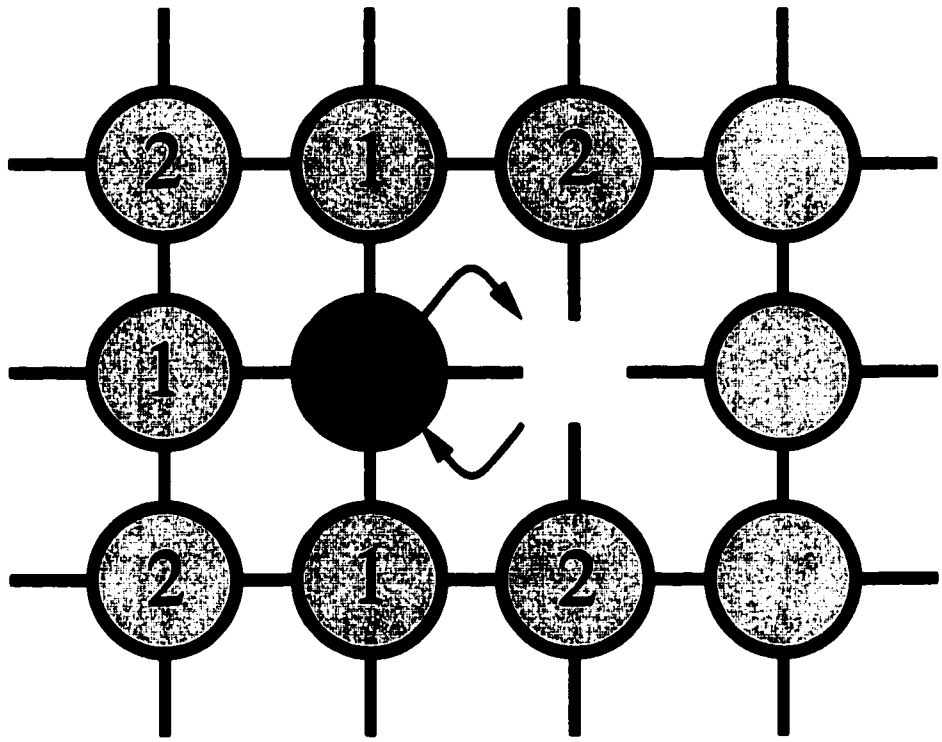


Figure 5.8 : Schematic figure illustrating the relationship between vacancy motion and net diffusion of an impurity atom for a square lattice. If the vacancy only moves among nearest neighbor sites (labeled "1") of the impurity (darker) atom, the impurity atom will never move more than one site from its initial location. The vacancy must reach a second nearest neighbor site (labeled "2") in order to approach the impurity atom from a different nearest neighbor site and cause net diffusion. The situation is similar in the Si lattice, except that the vacancy must reach a third-nearest neighbor site to cause net diffusion.

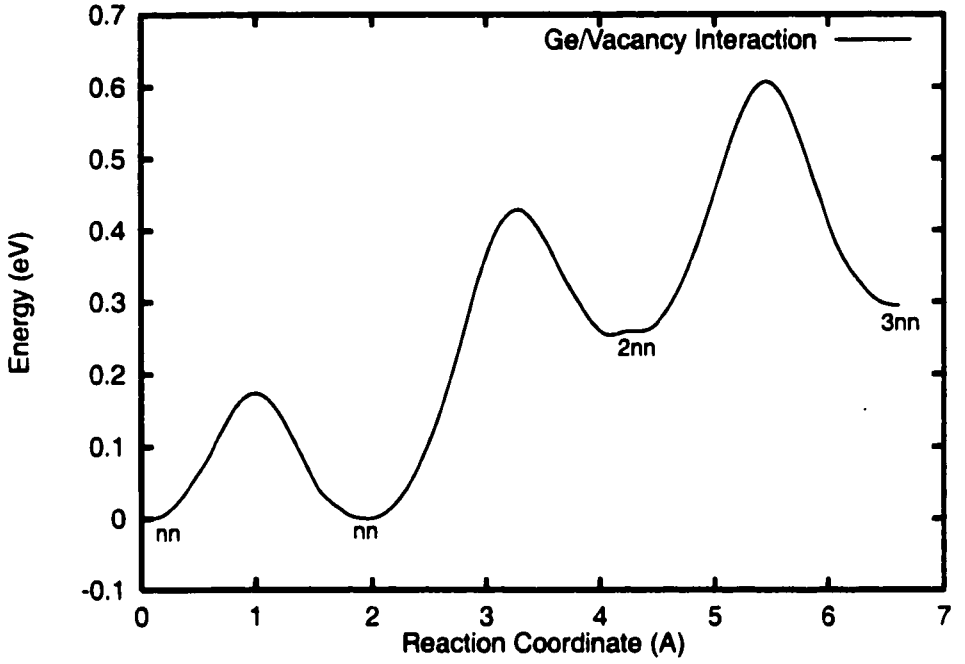


Figure 5.9 : Interaction of Ge_S and Si_V. The first two minima correspond to Si_V as the nearest neighbor of Ge_S and the next two correspond to Si_V in the 2nd and 3rd nearest neighbor sites, respectively.

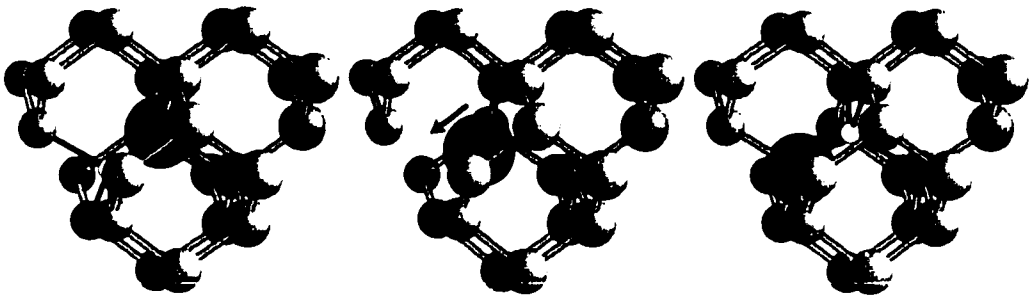


Figure 5.10 : Path for Si_V-Ge_S exchange. The view is a $\langle 110 \rangle$ plane slightly rotate by another 10° to more clearly show the position of the vacancy.

shown in figure 5.5.3. This path involves a Si atom moving from one lattice site to the site occupied by the vacancy, leaving its original site vacant.

From the perspective of Si transport through the Si lattice, Ge_S has the opposite effect on Si_V as it does on Si_I. The energy of Si_V, as well as the barrier for motion, decreases as Si_V approaches Ge_S. This is in contrast to the situation for Si_I, which sees slightly larger barriers and higher energies as it approaches Ge_S. Both intrinsic defects, however, are more stable when they incorporate Ge_S than they are far from Ge_S. Ge_S acts as a trap of Si_V and Si_I. In non-equilibrium conditions, when there is an excess concentration of interstitials or vacancies, Ge_S will act as a trap and will slow down the migration of both Si_V and Si_I. The overall activation energy of vacancy and interstitial diffusion will be unchanged, however.

The overall activation energy for vacancy mediated diffusion of Ge in Si, then, is 3.91 eV.

5.5.4 Migration Entropy

As mentioned before, the important barrier height is that between Si_V in the nearest neighbor site and in the 3nn site, as this is how far away Si_V must move from Ge_S in

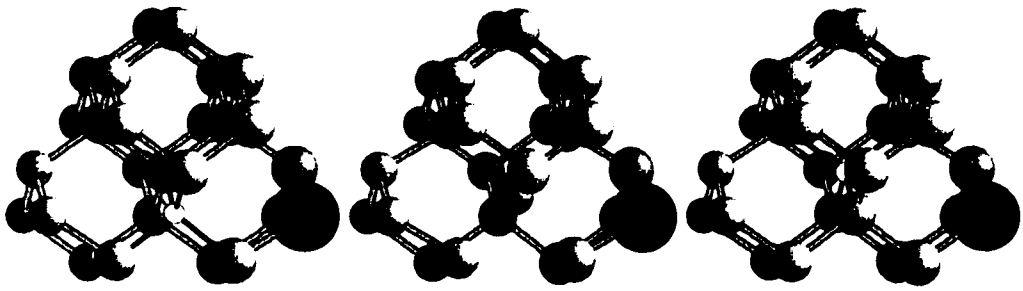


Figure 5.11 : Path for Si_V-Si_S exchange from the 2nn to the 3nn site. The view is a $\langle 110 \rangle$ plane slightly rotate by another 10° to more clearly show the position of the vacancy.

order to cause net diffusion. The prefactor for migration for this transition is 2.1×10^{13} s^{-1}. This is a factor of nearly 16 more than the entropy of migration found for Ge_I and a factor of 15 times smaller than that found for concerted exchange. As for the other migration entropies quoted in this chapter, only one displacement (0.01 Å) was used to obtain this number.

5.5.5 Entropy of Formation

Assuming that defects will be created relatively far away from Ge_S, we can use the values for the entropy of formation for the defects in pure Si. Those values, as reported previously in chapter 4, are $3.6k_B$ for the interstitial and $6.5k_B$ for the vacancy, within the harmonic approximation. The concerted exchange has an entropy due to the multiple ways in which it can occur on a given lattice site of $1.8k_B$ in addition to the entropy of migration mentioned previously.

5.6 Results

Summarizing, the activation energy for CE, interstitials, and vacancies is 4.5 eV, 4.1 eV, and 3.9 eV, respectively. We have estimated the entropy of migration for these

processes to be 3.1×10^{14} s^{-1}, 1.2×10^{12} s^{-1}, and 2.1×10^{13} s^{-1}. The formation entropy is the same as for Si for interstitials and vacancies: $3.6k_B$ and $6.5k_B$, respectively. Finally, CE has an entropy of configuration of $1.8k_B$. These values lead to diffusion constants for each mechanism of

$$D_I = 0.016\exp^{(-4.1/k_BT)}\,\mathrm{cm^2s^{-1}} \quad (5.1)$$

$$D_V = 11\exp^{(-3.9/k_BT)}\,\mathrm{cm^2s^{-1}} \quad (5.2)$$

$$D_{CE} = 2.8\exp^{(-4.5/k_BT)}\,\mathrm{cm^2s^{-1}}. \quad (5.3)$$

As in the case of Si self-diffusion, $a_I = 1.2$Å, $a_{V,CE} = 2.36$Å, and $z = 4$. For the vacancy, $f = 0.5$. Also, since there is four different ways of putting a defect and Ge together on a lattice site, there is an additional factor of $\Phi_{AX} = 4$. These diffusion constants lead to relative contributions at 1050°C (the temperature at which Fahey *et al.* conducted their experiments) of

$$f_I = 3.3\times10^{-4} \quad (5.4)$$

$$f_V = 0.999 \quad (5.5)$$

$$f_{CE} = 1.0\times10^{-3}. \quad (5.6)$$

At this temperature, the vacancy mechanism completely dominates. The other two mechanisms are negligible. The qualitative prediction that vacancy mediated diffusion of Ge should be stronger than of Si in the Si lattice relative to interstitial mediated diffusion because of its larger size is born out by these calculations. The ratio of $\frac{f_V}{f_I}$ is bigger for Ge diffusion in Si than for Si self-diffusion.

Thus, we get that the lower formation energy of vacancies in Si and their strong binding energy to Ge_S results in a domination of the vacancy mediated mechanism. Not only does the vacancy mechanism have a lower activation energy, but it has a much higher prefactor for diffusion. This is in contrast to the experimental results of Fahey *et al.*, which state that the interstitial is responsible for about 35% of the diffusion. Our activation energy is also low compared to those reported by Frank *et*

al. by nearly 1 eV in the case of the vacancy mechanism. This is the same problem we saw for the DFT/PW91 determination of the formation energy of the self-interstitial and vacancy in Si. Both are underestimated by about 1 eV. If 1 eV is added to the formation energy we use here, our overall activation energy for both the interstitial and the vacancy are right in the middle of the range reported by Frank *et al.*

Comparing to the case of pure Si (chapter 4), the vacancy mechanism dominates even more strongly for Ge diffusion in Si. This is to be expected, as the Ge atom is larger in size than the Si atom and would always be in a strained environment in the Si lattice, unless a vacancy was nearby.

Fahey *et al.* and of Södervall reported measured diffusivities at 1050°C. Plugging this temperature into equations 5.2,5.3,5.3, we find

$$d_I = 1.1 \times 10^{-18}\text{cm}^2/\text{s} \tag{5.7}$$

$$d_V = 3.4 \times 10^{-15}\text{cm}^2/\text{s} \tag{5.8}$$

$$d_{CE} = 3.5 \times 10^{-18}\text{cm}^2/\text{s}. \tag{5.9}$$

The diffusivities of interstitials and the concerted exchange are negligible compared to the vacancy. The diffusivity due to vacancies, d_V, is higher than the experimental results (see section 5.2) by about a factor of 10. The overestimation of d_V is most likely due to an underestimation of the activation energy. An increase in the activation energy of 0.1 eV would make the calculated diffusivity at 1050°C agree well with experiment. But, the calculated prefactor of 2.7 cm^2/s is two to five orders of magnitude too small. Most likely, both the activation energy and the prefactor are coming out to be too low in the DFT/PW91 calculations. We saw the same problem for Si self-diffusion.

5.7 Conclusions

While we are unable to quantitatively explain experimental results for Ge diffusion in Si with the DFT/PW91 calculations, there is qualitative agreement with experiment that Ge diffusion is dominated by the vacancy mediated mechanism.

Chapter 6

DIFFUSION OF GERMANIUM INTO THE SILICON SURFACE

The results presented here are published in *Physical Review Letters*, **84**, 2441 (2000).

6.1 Introduction

Interface abruptness in Si-Ge heterostructures has been the subject of intense research for many years [43]. Such heterostructures are of increasing interest for optoelectronic devices and quantum wells. Accurate modeling of these devices requires knowledge of the interface elemental profile. The lower surface energy and larger size of Ge relative to Si means that interdiffusion is thermodynamically favored when Si is deposited on Ge(100) [44]. In the reverse case, where Ge is deposited on Si(100), because of the lower surface energy and larger size of Ge, it has frequently been assumed that the interface is abrupt within 1–2 layers. Recently, evidence for Ge diffusion to the third [45, 46, 47, 48] and fourth layer [49] has been reported, although other work found no significant interdiffusion at similar temperatures [50].

Figure 6.1 illustrates the Si(100) surface structure. Dimerization of the surface layer results in several distinct sites. The open structure and inhomogeneous stress distribution may affect Ge incorporation in subsurface sites. The tendency of the larger Ge atom to occupy the tensile sites (3b, 4b) has been suggested as a mechanism for growth-induced ordering in $Si_{.5}Ge_{.5}$ alloys [51]. An alternative mechanism involves burial of Ge at single-height steps [52].

As we saw in chapter 5, diffusion of Ge in bulk Si has a large activation energy, on the order of 4–5 eV, and is negligible at the temperature of typical deposition experiments, about 500°C. This high activation energy for bulk diffusion has been taken as an indication that Ge diffusion below the second layer of the Si(100) surface would be kinetically hindered, even if thermodynamically favored. Yet, several experimental measurements have indicated the presence of subsurface Ge [45, 46, 47, 48, 49]. The experimental results described below, with significant interdiffusion at 500°C, suggest an activation barrier smaller than 2.5 eV for diffusion down to the fourth layer [1]. It has been suggested that the mechanism for subsurface diffusion must involve surface defects to facilitate transport [53].

In this chapter, we summarize experimental evidence for extensive interdiffusion during growth of Ge on Si(100), with Ge present throughout the top four layers after deposition of only 0.8 ML at 500°C. The results are not strongly dependent on step-structure (widely spaced, single-height steps vs. closely-spaced, double-height steps). We present our results of theoretical calculations explaining these measurements. Density Functional Theory (DFT) calculations of relative site energies were carried out and used to predict the thermodynamic distribution of Ge atoms. Extensive calculations of diffusion paths and activation energies revealed a mechanism that does not involve surface defects which could lead to near-equilibrium Ge distribution down to fourth layer without the presence of surface defects. Dopants are predicted to significantly affect the rate of Ge diffusion.

6.2 Calculations

For the most part, the DFT calculations are as described in chapter 3. For the surface calculations, there are some differences. Ge site energy calculations used a 64 atom

[1] Assuming the onset of the diffusion mechanism corresponds to the experimental conditions of the present study, ca. 100 sec at 500°C, and a standard prefactor for surface diffusion, 10^{-13} sec^{-1}, an activation energy of 2.3 eV is obtained within harmonic transition state theory.

cell with a vacuum spacing (the distance in the z direction between the periodically repeated slabs) of 11 Å and a 2 × 2 × 2 k-point sampling mesh of the Monkhorst-Pack type [19] in the irreducible Brillouin zone. Migration pathway calculations used a 96 plus 1 atom cell, a vacuum spacing of 6 Å, and a 2 × 2 × 1 k-point mesh. The Nudged Elastic Band (NEB) method [7], implemented to run in parallel on a cluster of computers, was used to find the minimum energy paths (MEPs) for the transitions. About 20 metastable structures were found using a 1 × 1 × 1 k-point set, not all of which remained unique with the 2 × 2 × 1 k-point set, and 20 MEPs between these structures were calculated.

6.3 *MBE Experiments of Ge epitaxy on Si(100)*

The work described in this section was done by Michael Leskovar and Marjorie Olmstead [54] and was used as support and justification for the theoretical model developed in the following sections.

MgK_α-excited Auger and photoelectron diffraction (AED and PED) were used to determine the near-surface Ge distribution in Ge/Si(100) films. High kinetic energy PED/AED is a powerful tool for obtaining element-specific structure of the near-surface region [55]. Emission from Ge atoms occupying different sites may be distinguished by different angles of forward-scattering by overlying atoms.

Figure 6.1 illustrates the dimerized Si(100) surface structure. The open structure and inhomogeneous stress distribution may affect Ge incorporation in subsurface sites. The tendency of the larger Ge atom to occupy tensile sites (labeled 3b, 4b) has been suggested as a mechanism for growth-induced ordering in $Si_{.5}Ge_{.5}$ alloys [51].

AED, PED, and LEED data were obtained for a matrix of Ge thickness (0.8 ML and 1.6 ML), substrate growth temperature (500 and 700°C), and substrate step structure (widely-separated, single-height steps (on-axis) and closely-spaced, double-height steps (4° miscut))[54]. Element-specific diffraction patterns for Si and Ge were

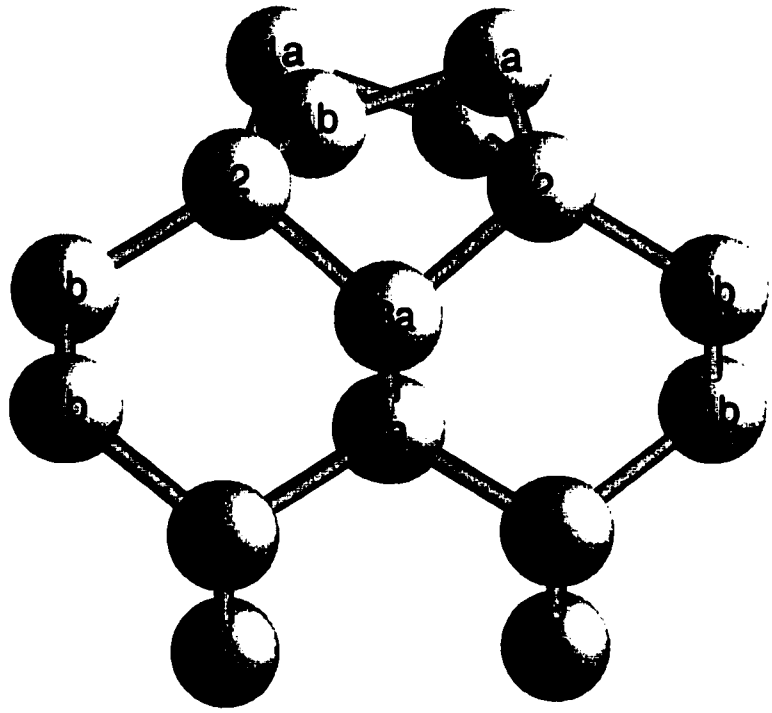


Figure 6.1 : Si(100)-c4x2 surface structure. The reconstruction results in two inequivalent sites in layers 1, 3, and 4, but one site in layers 2, 5, and 6.

obtained using intensities of Si 2p (PED) and Ge LMM (AED) emission, respectively.

In the absence of a kinetic pathway to subsurface sites, deposition of sub-ML Ge on Si(100) should result in Ge entirely in the top layer, with perhaps a small amount in layer 2 due to random burial or Si-Ge exchange. Such a distribution has been deduced from comparison of theoretical calculations with the coverage dependence of the $2 \times n$ reconstruction [50]. The predicted Ge photoelectron diffraction pattern for 0.8 ML deposition for such a distribution would be featureless except for a weak peak due to layer 2 Ge scattering from dimer atoms ($\theta \sim 60°$, $\phi \sim$13,77°). Experimental results for on-axis wafers for 0.8 and 1.6 ML deposition at 500 and 700°C are shown in Fig. 6.2a-c. The experimental results clearly show diffraction structure, indicating Ge buried to at least the fourth layer. Fig. 6.2d shows the predicted AED pattern for our model described in detail below.

The bright spots in Fig. 6.2 indicate diffraction peaks. At these high kinetic energies (1140 eV), AED is dominated by forward scattering peaks surrounded by dark rings. The peak positions marked by circles in Fig. 6.2 are those predicted for Ge emission from the various sites indicated in Fig. 6.1 (labelled in Fig. 6.2-d). The single-scattering AED calculations used coordinates we calculated by DFT for H-terminated Si(100). The weak but observable layer 2 emission is in contrast to the absence of layer 2 Ge predicted for Si-Ge alloys using the Tersoff potential [1], but in agreement with our own *ab initio* calculations (see below). Ge occupation of layers 3 and 4 is clearly visible in the data, with larger occupation in the sites between dimer rows (3b, 4b) than beneath them (3a, 4a). For example, note the layer 4 peaks along $\phi = 0°$ and 90°, especially at 700°C. Significant layer 4 (or deeper) occupation is indicated by the strong peak, present in all data sets, in the [110] direction ($\theta = \phi = 45°$). The [110] peak intensity increases by ~30% as the deposition temperature is increased from 500°C (6.2-b) to 700°C (6.2-a), and by another ~50% from 0.8 to 1.6 ML at 700°C.

Experiments on off-axis samples (single-domain terraces separated by double-

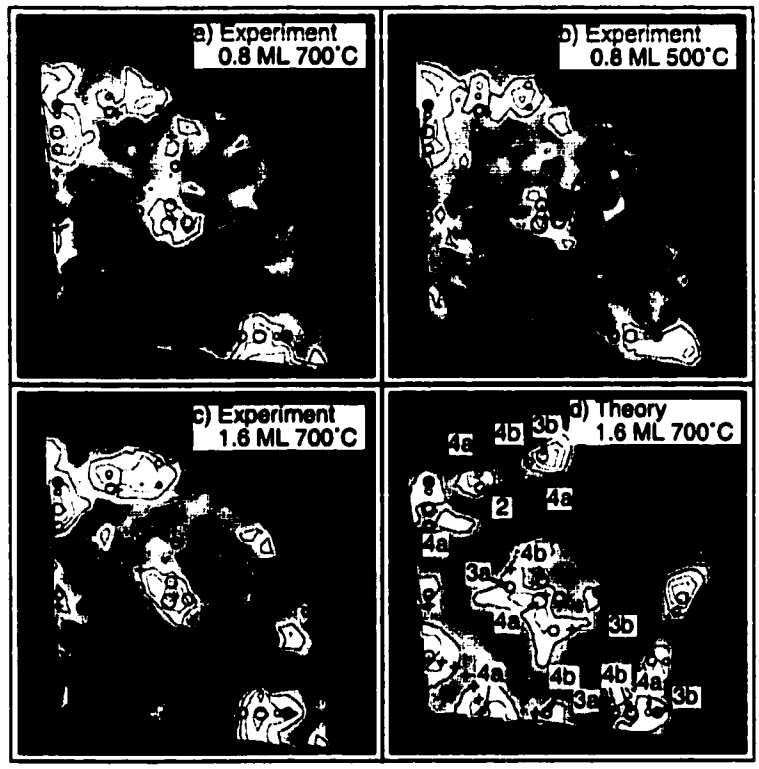


Figure 6.2 : Experimental (a-c) and Theoretical (d) AED stereographic projections for Ge deposition on on-axis Si(100): (a) 0.8 ML Ge, T_s = 700°C, (b) 0.8 ML, 500°C, (c) 1.6 ML, 700°C (d) simulated AED for 1.6 ML, 700°C based on occupations in Table 6.6. Circles denote expected positions of peaks due to various emitters, labelled in (d), with the size indicating intensity. Plus signs are on a 15° grid. Contour lines are 4% of average emission intensity. The data are normalized by the instrumental response and angles were aligned using the Si 20 emission.

height steps) show similar AED results to the on-axis data shown in Fig. 6.2. The data are difficult to quantify due to uncertainties in the ratio of 2×1 to 1×2 domains after deposition. However, all peaks characteristic of Ge buried to the fourth layer are clearly present for off-axis wafers, in roughly the same ratios as for on-axis wafers. This indicates an incorporation mechanism based on terraces rather than steps.

The experimental measurements indicate significant subsurface occupation at both 500 and 700°C, regardless of step structure. The dimer vacancy spacing for the on-axis samples is about the same as the step spacing for the off-axis samples, suggesting that neither greatly influence the results. The AED patterns in figure 6.2 are also similar to those of Yeom *et al.* [49], who deposited Ge at room temperature and annealed to 500°C. They deduced a Ge distribution in the ratio 4:1:1:1 for the top four layers, but did not distinguish between a and b sites. The questions remain: (1) why do Ge atoms get buried so deeply and (2) how are the kinetic barriers overcome? DFT calculations were carried out to answer these questions.

6.4 Effect of Surface Conditions on Ge Interdiffusion

The energetics of a substitutional Ge atom in the Si(100) surface were studied under various conditions. First, the perfect Si(100) surface with a c4x2 reconstruction was studied (table 6.1). This is the environment one would expect the Ge atom to see at the beginning of deposition. Next, one half of monolayer coverage was approximated by putting a Ge atom in the upper atom of the surface dimer, which we calculate to be the preferred spot for a Ge atom. The energy of an additional Ge atom relative to this one was tested (table 6.2). Third, a full monolayer coverage was examined, with the complete first layer being Ge (table 6.3). Finally, the influence of the dimer vacancy (DV) surface defect was tested, as this is a very common surface defect on Si(100) [56] and has been suggested as the "gateway" for Ge interdiffusion into the Si(100) surface [53] (table 6.4).

For the first three environments, the structure is depicted in figure 6.1, with the corresponding labels explained. For the DV structure, there are more inequivalent sites and the labels used in table 6.4 are explained in figure 6.3. These labels correspond to: 2.1 is a second layer atom in the DV, 2.2 is a second layer atom between two dimers in the same row as the DV, 3a.1 is a third layer atom in the DV, 3a.2 is a third layer atom in between the dimers in the same row as the DV, 3b is third layer atom in the trough between dimer rows, 4a.1 is a fourth layer atom in the DV, 4a.2 is a fourth layer atom under a dimer in the same row as the DV, 4b.1 is a fourth layer atom in the trough between rows next to the DV, and 4b.2 is a fourth layer atom in the trough away from the DV.

On the clean surface (table 6.1), the most preferred site is the upper atom of the surface dimer. This is consistent with experiment, which shows Ge incorporation first into this site [48]. The lower atom of the dimer is also favorable, compared to subsurface sites. In the subsurface (layers 2 through 4), there is a preference for the Ge atom to sit in the "open" sites (b), or those sites that sit underneath the trough between dimer rows, as opposed to under the surface dimers ("closed" sites, a). The preference for these sites, compared to the "closed" sites, is on the order of 0.06 eV. Both "open" sites, in the third and fourth layers, are of comparable energy. The second layer is slightly less favorable than these "open" sites.

After a half-monolayer has been deposited (table 6.2), in which the upper atom of the surface dimer has been replaced with Ge, the lower dimer is still the most favorable location for the next Ge. However, the second layer becomes more favorable than the "open" sites, by 0.03 eV in GGA. In addition, relative to the lower atom of the dimer, these "open" sites are less favorable than they were in the clean surface case. Compared to the 1b site, the energy cost of the 3b and 4b sites have gone up by about 0.04 eV, while the energy of the layer 2 site has only gone up by 0.01 eV.

When a full monolayer of Ge has been deposited, and the 0°K limit is assumed, then the first layer of the surface has been replaced by Ge. The energetics of this

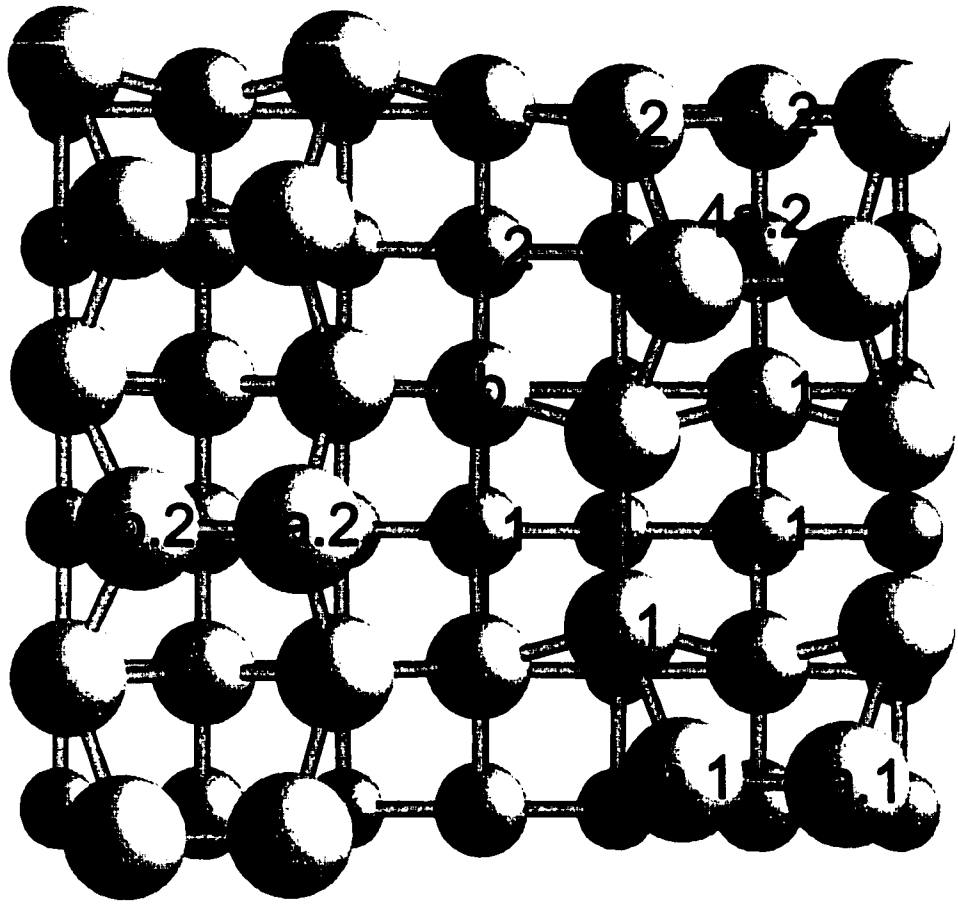


Figure 6.3 : Structure of the Dimer Vacancy on the Si(100) surface. The inequivalent sites in the second, third and fourth layer of the crystal are labeled. The dimer that was removed to form the dimer vacancy was positioned directly above the 4a.1 site.

Table 6.1 : Energy cost of substituting 1 Ge in an otherwise perfect Si surface. Energies are relative to the cost of substituting 1 Ge in the upper atom of the surface dimer.

	Energy (eV)	
Site	LDA	PW91
1a	0.00	0.00
1b	0.13	0.15
2	0.31	0.36
3a	0.36	0.42
3b	0.29	0.34
4a	0.34	0.42
4b	0.29	0.35

Table 6.2 : Energy cost of substituting 1 Ge in Si surface with Ge in upper atom of dimer. For comparison, the energy of the 1b site is set to the same value it had when the 1a site contained a Si atom (table 6.1).

	Energy (eV)	
Site	LDA	PW91
1a	—	—
1b	0.13	0.15
2	0.30	0.35
3a	0.38	0.44
3b	0.32	0.38
4a	0.37	0.44
4b	0.32	0.39

Table 6.3 : Energy cost of substituting 1 Ge in Si surface with Ge overlayer. For comparison, the energy of the layer 2 site is set to the same value it had when both the 1a and 1b site contained Si atoms (table 6.1).

	Energy (eV)	
Site	LDA	PW91
1a	—	—
1b	—	—
2	0.31	0.36
3a	0.40	0.48
3b	0.36	0.38
4a	0.40	0.48
4b	0.35	0.42

case are listed in table 6.3. The sites under the dimer row become even less favorable, as one would expect. The same is true of the 4b site. The layer 3b site, however, is nearly as stable as the layer 2 site when a full monolayer of Ge is deposited. It is not quite as stable as it was in the clean Si surface case, but is more stable than the surface with 0.5 ML coverage of Ge. The 3b site and the layer 2 site will be the preferred sites once 1 ML of Ge has been deposited.

The overall trend is that, as Ge is deposited and the top most layer contains more and more Ge, the subsurface sites become less and less favorable for Ge incorporation. That is, it is more favorable for Ge to interdiffuse into the surface at the beginning of deposition than after an appreciable amount has already been deposited (appreciable being on the order of half a monolayer). Interdiffusion is more likely to occur at the beginning of deposition and, as deposition proceeds, there will be less interdiffusion into the surface.

An important question is how surface defects might facilitate the interdiffusion

Table 6.4 : Energy cost of substituting 1 Ge in Si surface underneath a dimer vacancy. For comparison, the energy of the 2.2 site, being the layer 2 site furthest from the DV, is set to the value the layer 2 site had in the perfect Si surface (table 6.1).

	Energy (eV)	
Site	LDA	PW91
1a.1	—	0.04
1b.1	—	0.21
1a.2	—	0.08
1b.2	—	0.18
2.1	0.23	0.24
2.2	0.31	0.36
3a.1	0.44	0.48
3a.2	0.40	0.49
3b	0.35	0.42
4a.1	0.36	0.41
4a.2	0.40	0.48
4b.1	0.35	0.45
4b.2	0.33	0.40

of Ge into the surface. While actual migration barriers have not been studied, we have looked at the energetics of Ge atoms in the surface surrounding a DV, and those values are given in table 6.4. It is harder to compare these to the previous situations to get an idea of relative stability, but if we assume that a Ge atom in the layer 2.2 site, which is, in terms of local environment, equivalent to the layer 2 site of the clean surface (that is, it is bonded to two surface dimers), then we can get some idea of how the DV stabilizes Ge incorporation into the surface.

The energy cost of for Ge substitution in the surface sites in between DVs in our

DV configuration do not agree with the energy cost calculated for a perfect surface. This indicates that the cell is not big enough for a Ge atom to be far enough away from the DV as to avoid interaction with it. Nonetheless, some useful information can be extracted. First, the most stable site is 1a.1, the upper atom of the dimer adjacent to the DV. It is more stable than the equivalent site in the surface dimer far from the DV. The opposite is true for the lower atom of the surface dimer, with the 1b.1 site being less stable, though not by a large amount (0.02 eV), than the 1b.2 site. The second layer atom that is "bonded" to the DV is much more stable the the second layer atom that is bonded to two surface dimers, by over 0.1 eV. It is nearly as stable as the lower atom of the surface dimers (sites 1b.1 and 1b.2). Also, the 4a.1 site, which is underneath the dimer row and, normally, is not very favorable, is the most favorable layer 4 site, along with 4b.2 which is in the trough far from the DV. The DV does not help the incorporation of Ge into layer 3. The site underneath the DV is still 0.06 eV above the layer 3 site in the trough, while for the perfect surface it was 0.07 eV. It is harder to estimate the effect the DV has on the layer 4 trough sites. The energy difference is nearly 0.04 eV between them, but the layer 4 site further from the DV is more stable. This is counterintuitive, as one would expect that the more space provided by the DV would allow a larger atom such as Ge greater freedom for incorporation. However, we do not see that in the case of the layer 4 trough sites (as noted above, it is seen for the layer 4 site under the DV).

6.5 Occupation Model

Now having the energetics of a Ge atom in the Si(100) surface, we want to predict the thermodynamic behavior of the system. That is, we would like to use this information to predict which sites will contain Ge, and how much, versus temperature. There has only been one previous theoretical study of the temperature dependence of the surface Ge occupation [1]. This study was done using the Tersoff potential

and was a Monte Carlo simulation that calculated the equilibrium distribution of Si and Ge in a 50%/50% alloy of the two elements. We wanted similar information, but using the more accurate description of the interactions amongst Si and Ge given by DFT. Thus, a similar Monte Carlo simulation was out of the question. We developed another model, based only on the energy cost of placing Ge in the various inequivalent sites of the system, that allows us to predict the Ge distribution in the surface versus temperature for the epitaxial conditions present in the experiment previously described. This model is based on the energetics of the sites in the surface.

The free energy of such a site can be written, using the Gibbs entropy expression, as [10]

$$F = n(E - \mu) - kT\left[n\ln(n) + (1-n)\ln(1-n)\right].$$

The first term describes the energy of the site E relative to the chemical potential μ. The second term accounts for the entropy of the site, or the number of states it can have. In this case, the site can have two states: filled or empty, where filled refers to the site containing a Ge atom and empty to the site containing a Si atom. n then refers to the percentage of these sites that contains a Ge atom or, equivalently, the probability that a site contains a Ge atom.

Minimizing this free energy with respect to n and solving for n leads to

$$n(E,T,\mu) = \frac{e^{\frac{\mu-E}{k_BT}}}{1+e^{\frac{\mu-E}{k_BT}}},$$

the Fermi-Dirac distribution. E is the energy of a given site, T is the growth temperature, and μ, the chemical potential, is a parameter we use to fit to the experimental growth conditions. We tune μ so that the total amount of Ge in the model is equal to the amount deposited during the experiment.

To verify that this model works, we first compared with the Monte Carlo simulation of Kelires and Tersoff [1] of a 50/50 Si/Ge alloy. We used the Tersoff potential [57] to calculate the energies of the various inequivalent sites in the Si surface and

used Fermi-Dirac statistics to determine occupation. Our results are compared to the full Monte Carlo simulation in table 6.5, which shows that this model works well in reproducing the more expensive calculation within 10%. As can be seen, agreement is increased if an overlayer of Ge is included in the energy calculations.

Also shown are the results of using Boltzmann statistics with the site energies both for the clean Si surface and the surface with a Ge overlayer. The relative occupation for site i given by Boltzmann statistics is

$$n_i(E,T,\mu) = \frac{e^{\frac{\mu - E_i}{k_B T}}}{\sum_{j=1}^{N} e^{\frac{\mu - E_j}{k_B T}}} \tag{6.1}$$

where the denominator represents the total occupation of all sites in the surface and thus n_i is the fraction of the total occupation that is in site i. μ cancels out of this expression. As can be seen in table 6.5, Boltzmann statistics do not reproduce the results of the Monte Carlo simulation nearly as well as Fermi-Dirac statistics. The occupation of site i depends only on its energy relative to the other sites in the surface. The chemical potential drops out, with the same factor in the numerator and denominator. Because the first layer site dominates so strongly in energy, very little occupation is seen in the lower layers. The Boltzmann distribution assumes that the number of states in the system is much greater than the number of particles, but this assumption is not valid here. With the Fermi-Dirac model, however, the chemical potential allows the occupation of a specific site to be set, as in this example where the chemical potential is set so that the bulk site has an occupation of 50%.

The difference between Fermi-Dirac statistics and Boltzmann statistics is illustrated again in figure 6.4. There are two things to notice. First, in the zero temperature limit, Fermi-Dirac statistics give an occupation of 1 for sites whose energy is lower than the chemical potential (or the energy of the bulk site in this case) and 0 for those sites whose energy is higher. Boltzmann statistics give 0 occupation for all sites except the one site with the lowest energy. Second, the Boltzmann statistics do

Table 6.5 : Ge atom occupation at the surface of a for 50/50 Si/Ge alloy at a temperature of 300 K. All calculations are based on the empirical Tersoff potential. Shown are the values calculated via a Monte Carlo simulation [1] and the occupations predicted by the present study using the energy of substituting one Ge atom for a Si atom and Fermi or Boltzmann statistics. Site energies (E) and free energies (F) were calculated for a pure Si surface as well as one with a Ge overlayer (Ge OL).

		Fermi Statistics				Boltzmann Statistics	
Layer	Monte Carlo	E	E/Ge OL	F	F/Ge OL	E	E/Ge OL
1	1.00	0.995	1.00	0.997	1.000	0.961	1.00
2	0.17	0.028	0.086	0.030	0.11	0.0001	0.013
3a	0.40	0.52	0.48	0.50	0.45	0.005	0.114
3b	0.60	0.64	0.61	0.65	0.61	0.008	0.188
4a	0.22	0.20	0.23	0.19	0.22	0.001	0.037
4b	0.69	0.77	0.76	0.76	0.75	0.015	0.392
5	0.45	0.51	0.52	0.50	0.49	0.005	0.133
bulk	-	0.50	0.50	-	-	0.005	0.123

not converge to the Monte Carlo simulation values for higher temperature, while the Fermi-Dirac statistics do. The reason for these differences and the reason Fermi-Dirac statistics reproduce the Monte Carlo results is because each lattice site can be viewed as a fermion site: it either contains 1 or 0 Ge atoms, just in the same way that an allowed electron energy level contains 1 or 0 electrons. Boltzmann statistics allow a site to have any number of Ge atoms and the zero temperature limit will have all of the Ge atoms in the lowest energy site. With Fermi-Dirac statistics, those sites will fill up and the Ge atoms will be forced to occupy higher energy sites until there are no more Ge atoms left, with this limit set by the chemical potential. If the occupation number of Ge is small, much smaller than 1, then Boltzmann statistics should agree well with Fermi-Dirac and that is just what we would see. The chemical potential would equivalently lower until only the lowest energy site had any occupation. However, when the occupation of a given site is significant, then Boltzmann statistics no longer apply and Fermi-Dirac statistics must be used.

The free energy of the various sites was also calculated by the method of thermodynamic integration [58]. Here, the identity of the atom is continuously changed from Si to Ge as a function of some variable λ at a given temperature and force sampling is done. The average force is integrated over λ, giving the free energy difference between the crystal with Si in the site and with Ge in the site. This has the effect of including entropy of vibration into the calculation of the occupation. Thus, the only component included in the full Monte Carlo simulation missing in this model is the interaction between neighboring Ge atoms. This is partially included by adding the Ge overlayer, though interaction between subsurface sites is still neglected. The energy and free energy obtained are given in table 6.5.

As can be seen, the agreement between this model and the Monte Carlo simulation increases as the entropy of vibration and the interaction amongst Ge via the Ge overlayer are added. In the end, the agreement when both are included is about 5% in the occupation number or better.

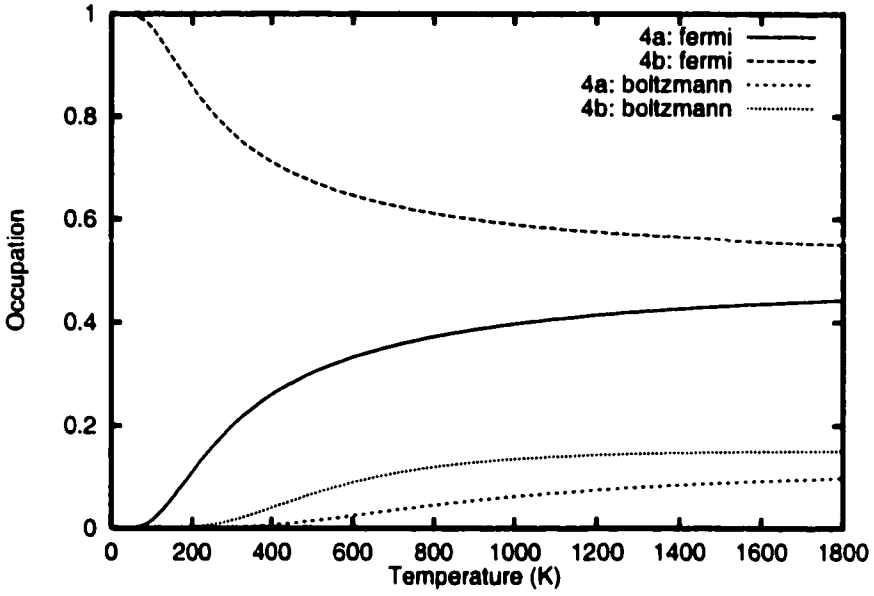


Figure 6.4 : Predicted occupation versus temperature for the two layer 4 sites using both Fermi-Dirac and Boltzmann statistics.

Table 6.6 : Empirical potential study, site energies in eV to substitute a Ge atom into various sites in the Si crystal. Both the energy and the free energy are given for the case of a Si(100) surface as well as for the surface with a Ge overlayer (OL).

Layer	E	E/Ge OL	F	F/Ge OL
1	0.00	-	0.00	-
2	0.23	0.23	0.23	0.23
3a	0.14	0.17	0.14	0.18
3b	0.12	0.16	0.13	0.17
4a	0.17	0.20	0.18	0.21
4b	0.11	0.14	0.11	0.15
5	0.14	0.17	0.14	0.18
bulk	0.14	0.17	0.15	0.19

Table 6.7 : Energy cost of substituting a Ge atom for Si in various sites near the Si surface relative to 1a and calculated fractional occupations. The site labels are explained in Fig. 6.1. Results of LDA and PW91 DFT calculations as well as an empirical interaction potential [2] are given. The fractional occupations are based on the PW91 values and equation 6.5, and correspond to experimental conditions of 0.8 and 1.6 ML Ge coverage at substrate temperatures of 500 and 700°C.

	Energy (eV)			Fractional Occupation 0.8 ML		1.6 ML	
Site	LDA	PW91	Tersoff	500°C	700°C	500°C	700°C
1a	0.000	0.000	0.000	0.904	0.831	0.987	0.965
1b	0.134	0.149	—	0.495	0.449	0.892	0.820
2	0.314	0.363	0.230	0.037	0.058	0.243	0.257
3a	0.361	0.415	0.136	0.017	0.032	0.127	0.156
3b	0.292	0.344	0.123	0.049	0.072	0.300	0.303
4a	0.344	0.419	0.174	0.016	0.031	0.121	0.150
4b	0.291	0.348	0.107	0.046	0.069	0.287	0.293

6.6 Theoretical Results for Occupation

We applied the Fermi-Dirac occupation model using the substitutional energies, E, found with DFT. Since the amount of deposition here is on the order of a monolayer or less, we are using the pure Si surface energies, not the energies with the Ge overlayer. Table 6.6 gives the calculated energies and resultant occupation numbers for the various sites after deposition of 0.8 and 1.6 ML at 500 and 700°C.

These calculations show that, even for 0.8 ML deposition at 500°C, an appreciable amount of Ge is predicted thermodynamically to be found in the 3rd and 4th layers. The amount increases with temperature and deposition amount. For 1.6 ML of deposition at 700°C, almost one third of both the 3b and 4b sites are predicted to contain

Ge. The subsurface sites underneath the trough are predicted to have the largest occupations, in agreement with earlier calculations using an empirical potential [1]. However, unlike that study, we find the occupation of layer 2 to be nearly the same as these tensile sites, which agrees with the experimental results. The tensile strain on these sites due to the surface reconstruction is more accommodating for the larger Ge atom. We have simulated the AED patterns based on these occupation numbers and find that the results are in good qualitative agreement with the experiment, as shown in figure 6.2. The stronger layer 4 signals in the experiment than in the calculated pattern could arise either from additional layer 4 occupation or from approximations in the AED simulations (single-scattering; all Si atoms).

6.7 Mechanism of Interdiffusion

The experiment thus reveals a near-equilibrium distribution of Ge in the first four layers of the surface. The activation energy for Ge diffusion in bulk Si is measured to be in the range of 4 to 5 eV [59, 60], much higher than the ~2.5 eV upper bound obtained from the present experiments [2]. The question is: What mechanism could be responsible for such rapid diffusion in the top four layers? We present here results of DFT calculations of diffusion paths, illustrating how an adatom can reach the fourth layer.

Experiments have shown that Ge diffuses in bulk Si by both substitutional-interstitial exchange and a vacancy mechanism, with comparable contribution from each [40]. Our work on Ge diffusion in bulk Si (chapter 5) shows that the vacancy mechanism dominates. We focus here on the interstitial mechanism. The exchange process where a Si interstitial displaces a Ge atom from a lattice site in the bulk,

---

[2] Assuming the onset of the diffusion mechanism corresponds to the experimental conditions of the present study, ca. 100 sec at 500°C, and a standard prefactor for surface diffusion, 10^{-13} sec^{-1}, an activation energy of 2.3 eV is obtained within harmonic transition state theory.

$$Si_I + Ge_S \rightarrow Si_S + Ge_I, \tag{6.2}$$

has a small activation barrier. Our DFT calculations give a barrier of 0.3 eV and the exchange is downhill by ~0.1 eV. This process, in which Ge diffusion is assisted by a Si interstitial, is of lower energy than Ge directly diffusing through the crystal. *The dominant part of the activation energy for Ge diffusion is the formation energy of the Si interstitial.* This has been estimated to be 3.3 eV using LDA [28] and our PW91 calculations give 3.7 eV. This gives a total activation energy for bulk Ge diffusion of 4.0 eV within DFT. The central question is: *What is the formation energy of Si interstitials near the surface?* Since Ge displaces surface Si upon deposition[48], both Si and Ge adatoms are present during growth; we calculate the interstitial energy with respect to the energy of the Si adatom. We find the formation energy of Si self-interstitials in the surface region to be significantly lower than in the bulk. Fig. 6.5 shows the formation energy of interstitials as well as a possible pathway for a Si adatom to reach the fifth layer. Even in the fifth layer, the nature of the interstitial is different than in bulk. The optimal configuration is a tetrahedral interstitial (fig. 6.5(iv)) rather than a dumbbell (fig. 6.5(iii)) [38]. The low energy of interstitials near the Si(100) surface is consistent with calculations of Tersoff for C diffusion in Si using his empirical interaction potential [2].

The migration path (fig. 6.5) involves the hop of an adatom (i) on top of a dimer row and then displacement of the adatom down between two surface dimers to a dumbbell interstitial geometry perpendicular to the surface dimers (ii) where it shares a third layer site with a lattice atom. One of these atoms pushes an atom out of the fourth layer, forming a dumbbell parallel to the surface dimers with an energy of 1.8 eV relative to the adatom (iii). One of these atoms can jump further down into the 5th layer, forming a tetrahedrally coordinated interstitial (iv). This configuration is about 2.2 eV above the adatom in energy. The barrier for formation of the interstitial in the fourth layer (iii) is 2.2 eV with respect to the Si adatom. The

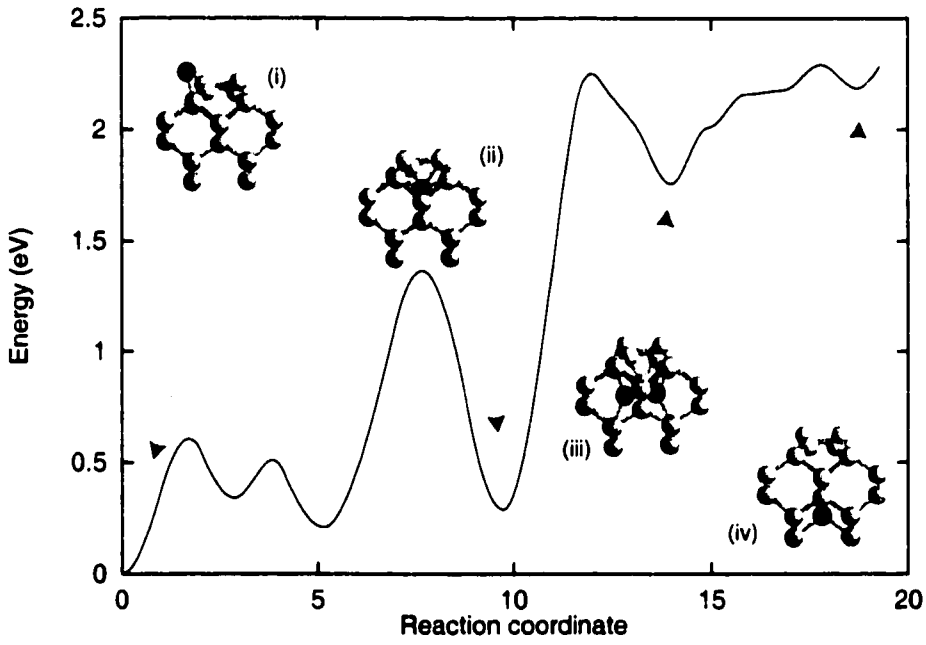


Figure 6.5 : Minimum energy path calculated by DFT for the diffusion of a Si adatom to subsurface interstitial sites down to the 5th layer. The activation energy to bring an adatom to layer 4 is 2.2 eV. Stable structures include: (i) the adatom, (ii) a dumbbell in layer 3, (iii) a dumbbell in layer 4, and (iv) a tetrahedral interstitial in layer 5.

overall barrier for Ge diffusion down to the 4th layer is estimated to be 1.8 eV+0.3 eV= 2.1 eV (after adding 0.3 eV for the Si/Ge exchange) [3], on the order of the upper bound set by the experimental conditions.

6.8 Effect of Charge

We have seen a significant effect in these calculations due to the charge state of the system. The stability of the 5th layer interstitial compared to the adatom goes from 2.6 eV at a charge of −2 to 1.6 eV at +2. The insertion of boron far from the defect structure in the cell [4], with the accompanying loss of an electron, has a similar effect, stabilizing the interstitial by 0.5 eV compared to pure Si, thereby lowering the activation energy for Ge diffusion to the 5th layer by 0.5 eV. This suggests that the Ge/Si interface may be sharper for n-type than for p-type substrates.

The difference in energy between the interstitial in the 5th layer and the adatom is due to a difference in the allowed energy states for the electrons. In figure 6.6, we show the band structure for both systems (the 5th layer interstitial and the adatom) with various charges.

The difference in energy between the interstitial structure and the adatom is entirely due to the differences in occupation of the electronic states. The calculation used a k-point mesh of 2×2×1. The two k-points remaining after symmetry is applied are shown in figure 6.6. K-point 1 shows a normal band gap, while k-point 2 shows a band gap with two states in the middle. For the neutral system, the energy levels at k-point 1 are full up to the valence band maximum. At k-point 2, the adatom structure has filled energy levels up to the valence band maximum, but the gap states

[3]The barrier is 2.7 eV with respect to a kink atom, which is 0.6 eV lower in energy than an adatom. Also, the reason we use 1.8 eV here is because the exchange between Si and Ge would occur from a stable site, not from the top of the barrier. So, once the Si atom has settled into its interstitial site with an energy of 1.8 eV, it can then exchange with Ge, which has a barrier of 0.3 eV.

[4]The B atom was inserted 3 layers below and 2 layers to the side of the interstitial in the cell. This corresponds to a distance between the B and the interstitial of about 6 Å.

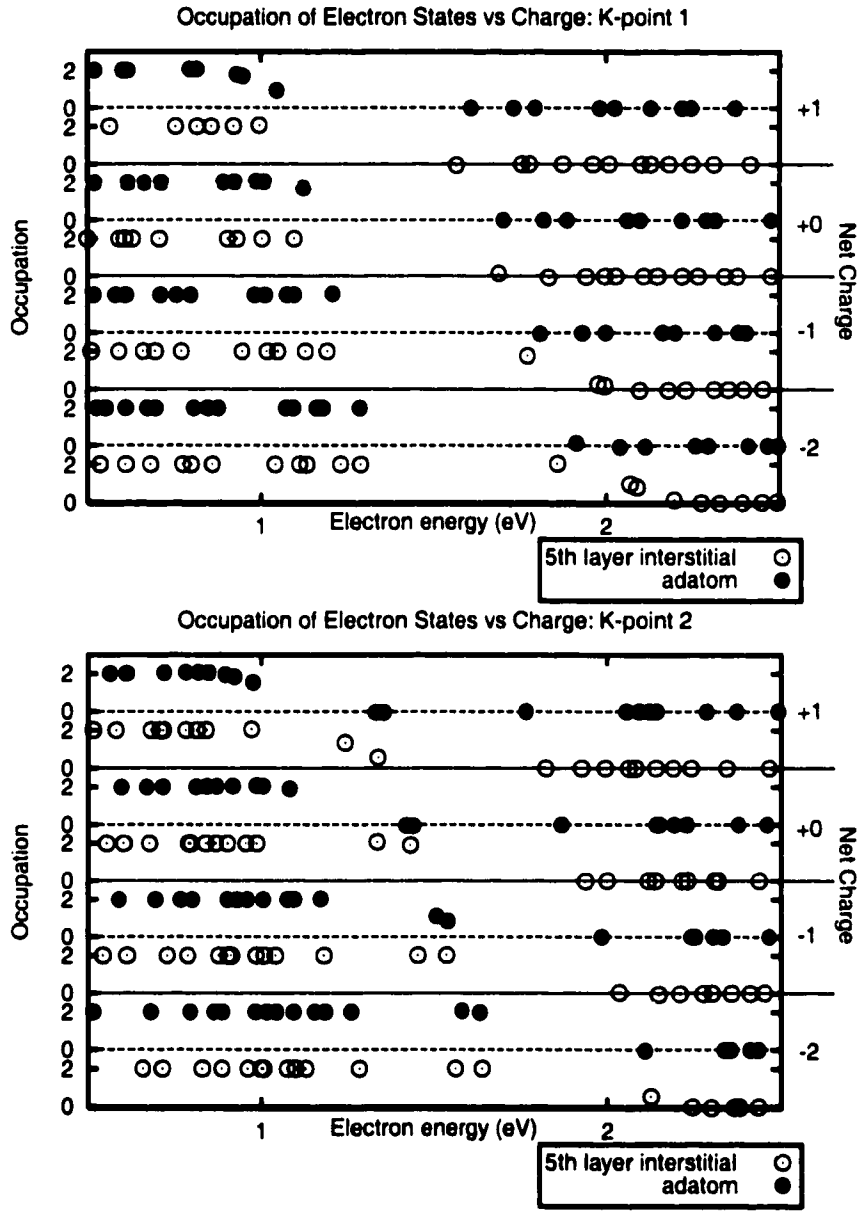


Figure 6.6 : Electron energy levels for 5th layer interstitial structure and adatom on the Si(100) surface for the four charge states +1, 0, −1 and −2 for the two irreducible k-points in the Brillouin zone resulting from a 2×2×1 k-point mesh.

are empty. This is not true for the interstitial, for which the two gap states are occupied. If an electron is removed and the net charge on the system is +1, the top of the valence band is depleted in the case of the adatom structure, while for the interstitial structure, the gap states lose occupation. They are not completely empty, but they are nearly so. So, the highest energy electrons in the +1 charged system have similar energy in both the adatom and interstitial configurations. In the neutral system, the interstitial structure had electrons at a higher energy than the adatom structure. As electrons are added to the system, the adatom structure has empty gap states that can be filled while the only empty states these new electrons can go in the interstitial structure are at the bottom of the conduction band, states that are much higher in energy. This explains why the interstitial structure, relative to the adatom, has a higher energy as the system becomes negatively charged than when it is positively charged.

6.9 Empirical Potential Studies

Preliminary empirical potential studies were done of the Si and Ge (100) surface in order to understand interdiffusion. However, we found that the empirical potentials available give a poor description of migration barriers near the surface. We discuss the calculations here for completeness.

Both the Tersoff potential [61, 62, 63, 57] and the Bolding-Andersen potential [64] were used. The original motivation was to understand the origin of the Si/Ge ordering seen in various experiments in which codeposited Si and Ge on a Si substrate segregate, forming ⟨111⟩ planes that are alternatively Si and Ge rich. Two mechanisms were proposed and we conducted a study of the barriers for exchanges on the surface in order to distinguish between the two mechanisms. Those results are summarized here. It should be pointed out that the two potentials did a poor job of describing bulk self-diffusion and would be expected to perform even worse on the surface.

The phenomenon of Si/Ge ordering and segregation in the ⟨111⟩ planes of crystals grown by the codeposition of Si and Ge on a Si substrate was observed by several groups in the early 90s [51, 65, 66, 67, 68]. Two competing mechanisms were proposed. One involved the local strain in the surface layers of the crystals due to the surface reconstruction [51, 67]. The other explanation invoked a preference for exchange at one type of step as opposed to another. The Si(100) surface, if miscut, will form single height steps. These steps come in two varieties: rebonded and non-rebonded. The "atomic pump mechanism" proposed that exchange is favored at the rebonded step because a high energy dangling bond is replaced by a lower energy dangling bond [65, 66]. To estimate the validity of these models, we calculated barriers with empirical potentials for exchanges on the surface.

Calculations of the barriers for exchanging atoms in the surface of Ge(100) were conducted with the Tersoff potential. Both the perfect Ge(100) surface and a surface with both types of single height steps (rebonded and non-rebonded) were examined. In addition, the barriers for the single height step processes were calculated with the BA potential for the Si(100) surface. Figure 6.7 shows which paths were calculated for the perfect surface and figure 6.8 shows those for the stepped surface. The results are summarized in table 6.8. In the BA calculations, all atoms are Si. In the Tersoff calculations, all atoms are Ge except the upper atom in the exchange, which is Si.

All of the barriers are high, much higher than would be active at typical experiment temperatures. The barriers for exchange in the Ge surface are especially high, being higher than the experimental values for Ge self-diffusion. At the stepped surface, the Tersoff potential predicts that decent of a Si atom in the Ge(100) surface is easiest at the flat surface (process F)[5], and most difficult by a significant amount (0.8 eV) at the rebonded step (process G). In addition, the driving force for the exchange to

[5]While, in figure 6.8 process F appears to take place between two steps and would feel the strain created by those steps, the figure is only a cartoon illustrating the types of processes considered. In reality, process F was calculated on a clean surface, free of steps

Table 6.8 : Barriers for exchanging atoms in the Ge(100) and Si(100) surface. The Tersoff potential was used to calculate the barrier for a Si atom to exchange sites in a Ge(100) surface. The labels describing the processes are explained in figure 6.7. A positive energy preference means that the final state is more preferable than the initial state. For the BA potential calculations, the entire surface is Si, so there is no energy preference for the exchange.

	Energy (eV)		
Process	Tersoff	Energy Preference	BA
A	5.114	0.005	—
B	5.086	0.038	—
C	4.564	-0.106	—
D	4.842	-0.044	—
E	6.008	—	5.374
F	3.391	0.204	2.966
G	4.702	0.222	3.648
H	3.910	0.119	3.417

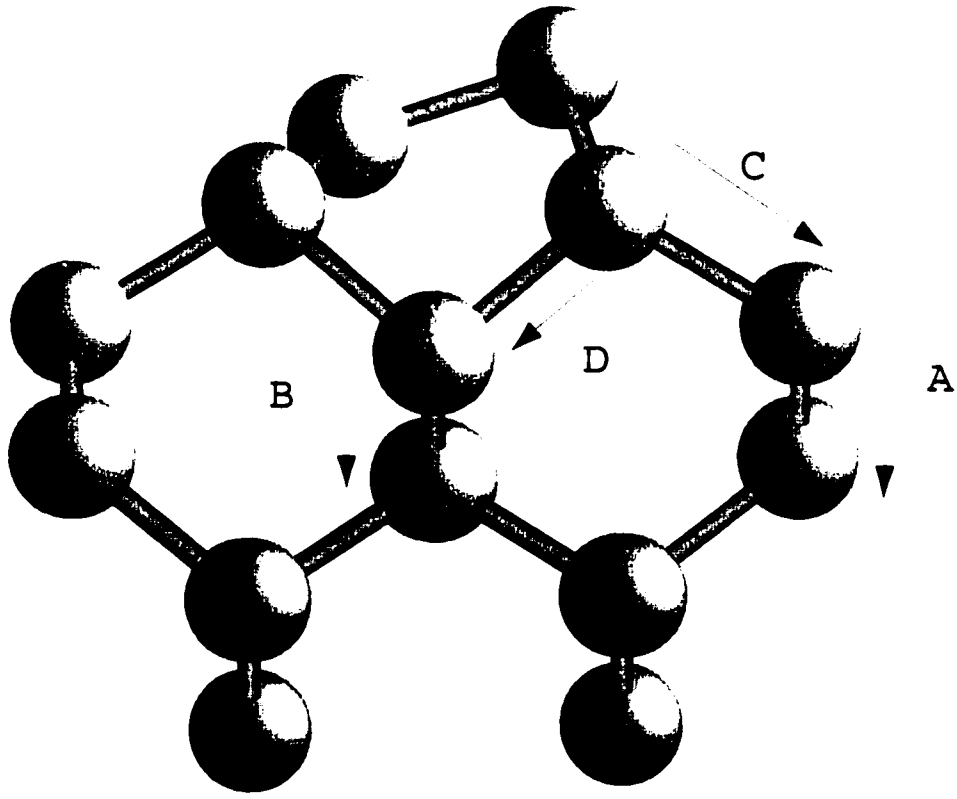


Figure 6.7 : Exchange processes calculated on the perfect (100) surface. The barriers and driving forces for these processes are given in table 6.8.

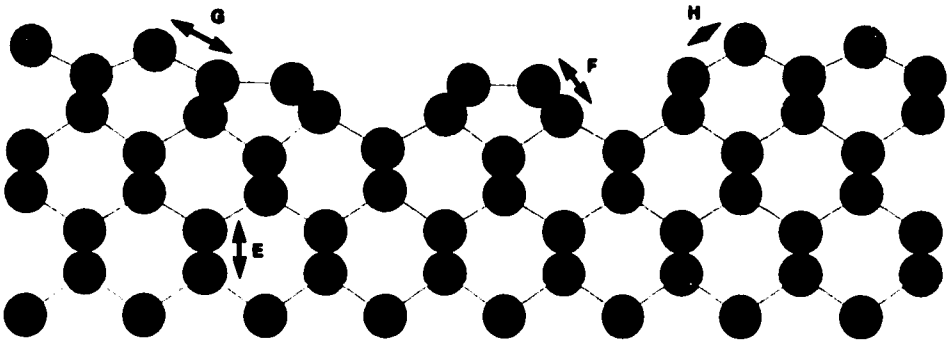


Figure 6.8 : Exchange processes calculated on the stepped Si(100) surface. Process G occurs at the rebonded step and process H at the non-rebonded step. Process F occurs on the flat surface and E represents the exchange in the bulk crystal. This figure is only a schematic illustrating the different pathways considered. It does not represent the actual structure of the cells used in the calculations.

occur is greater for process G than process H. Process E represents the exchange in the bulk, so there is no energy preference for this process to occur (the environment of the Si atom is identical in both the initial and final state).

These calculations agree with the atomic pump mechanism in that exchange is favored at the rebonded step (process G), but we find it is only favored by 0.1 eV, as opposed to the 0.45 eV estimated previously [65, 66]. In addition, we find that the barrier for exchange at the rebonded step is much higher than at the non-rebonded step (process H) with the Tersoff potential. Using the BA potential, it is still true that process G has a higher barrier than process H, but not by nearly such a high margin (only 0.23 eV). Both potentials agree that decent on the flat surface (process F) is the most favorable exchange. The high barrier for process G suggests that the atomic pump mechanism would not be active at the temperatures at which growth normally occurs.

Again, the usefulness of these calculations is doubtful. Neither potential describes well the fully coordinated motion of the exchange in the bulk and it is to be expected

that, on the surface where dangling bonds exist and a more complete description of electrons would be necessary, they would give an even worse description.

Chapter 7

CONCERTED EXCHANGE OF DOPANTS IN SILICON

7.1 Introduction

Pandey first proposed the concerted exchange (CE) mechanism as a possible component of Si self-diffusion in 1986 [21]. Since then, the size of the role CE plays in Si self-diffusion has been analyzed by various experimental groups. While, in general, it has been concluded that CE plays a small role, it has not been ruled out as a contributor to self-diffusion. Recently, it has been estimated that the contribution to self-diffusion by CE is in the range of 0 to 62% at 1100°C [24]. This same group finds that CE potentially plays a relatively large role in As diffusion as well, though its role in B and P diffusion was shown to be negligible. Our own results for Si self-diffusion and Ge diffusion in Si, given in chapters 4 and 5, show that concerted exchange is a small contributor to the diffusion of those systems, but that its contribution is nearly the same as interstitial mechanisms for Ge diffusion in Si. Thus, CE cannot automatically be discarded compared to vacancies or interstitials without more detailed analysis.

The exact nature of self-diffusion and, by extension, foreign atom diffusion in Si is still not known. While the roles of interstitials and vacancies have been analyzed from first principles calculations, their relative contributions have been difficult to extrapolate. This is because one needs to know not only the energy barrier to diffusion, but also the entropy barrier, as well as the energy and entropy of formation. Recent work has been done to complete the picture for interstitial and vacancy diffusion for Si self-diffusion for both neutral and charged defects (see chapter 4 and reference [5]).

The role of these defects in dopant diffusion is just beginning to be calculated with first principles, with the role of interstitials in boron (B) diffusion being an example [69]. However, the calculation of the CE mechanism for B and other common dopants has not been done. The goal of this work is to fill in the gaps of the first principles work done on diffusion in Si by providing the energy barrier for the concerted exchange mechanisms of column III elements B, Al, Ga, and In, column IV elements C, Si, Ge, and Sn, and column V elements N, P, As, and Sb in Si.

7.2 Previous Work

As mentioned previously, the concerted exchange was first mentioned as a possible self-diffusion mechanism in Si by Pandey [21]. Further studies of the CE mechanism have resulted in values for the entropy of migration [70]. The values reported place the CE mechanism within the range of experimental values for both the activation energy and prefactor for Si self-diffusion. Even so, it has been difficult to experimentally determine how much self-diffusion might be due to CE. Estimates of the fractional diffusion due to CE range from 0 to 0.6 [24]. The same paper puts similar bounds on the CE component of As diffusion in Si. Our results from chapters 4 and 5 show that DFT predicts concerted exchange to be a very small part of Si self-diffusion and Ge diffusion in Si, though comparable to the interstitial component. However, it is possible that, for some other atoms, it could play a larger role.

Previous to this, Nichols *et al.* studied the diffusion of B, Sb, P and As in Si [71]. Using DFT with the local density approximation, they calculated an upper bound for the CE mechanism for each of these elements. Their cells contained 32 atoms. They did not relax all atoms in their cells to find the minimum energy structure, but rather relax the first neighbor shell of atoms to three different positions and then interpolate to find the minimum energy distance. They obtained 4.9 eV for B, 4.6 eV for P, 3.9 eV for As, and 4.5 eV for Sb.

The experimental work mentioned previously consisted of oxidation and nitridation experiments at two temperatures for four different dopants. Oxidation injects interstitials into Si while nitridation injects vacancies. The profiles of these dopants were measured by SIMS for anneals of 1 hour at 1100°C and 5 hours at 1000°C. This gives enough information to numerically solve the resulting system of equations for the concentration profiles for the fractional contribution of interstitials, vacancies and concerted exchange to the diffusion constant. It was estimated that, at 1000°C, the fraction of As diffusion in Si due to concerted exchange (f_{AsCE}) is 0-0.36 and, at 1100°C, 0-0.61 while, f_{SiCE} is 0-0.27 and 0-0.62 for the same temperatures, respectively. Thus, CE might play a large role, but it might also be completely negligible. [24]

7.3 Details of Calculation

The energy cutoff used in the following calculations was 150.6 eV for Al, As, Ga, Ge, In, Sb, Si and Sn, 257.2 eV for B, 286.7 eV for C, 348.1 eV for N, and 173.5 eV for P. Otherwise, the calculations are as described in section 3.3.

7.4 Results

The potential energy barriers for CE in neutral systems for column III elements are shown in figure 7.1 and for column IV and V elements in figure 7.2. We find that the shape of the barrier for a given element is very much a function of the column to which that element belongs. The energy curves for column III elements are featureless, except for a slight shoulder on each side of the saddle point. In contrast, the barriers for column IV and V elements all show a minimum on each side of the saddle point. This is true even for Si, and this is the first time this has been reported (Pandey found a barrier that had only the saddle point as an extremum along the energy curve) (this was described in chapter 4). All the column IV elements have higher energy saddle

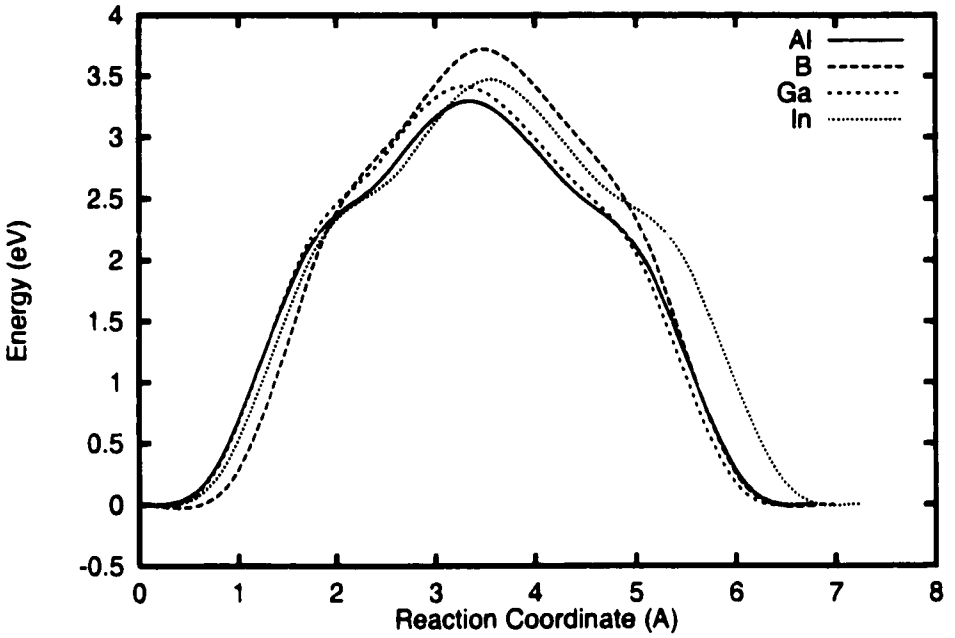


Figure 7.1 : Concerted exchange barriers for column III elements in Si.

points than the column V elements, with small differences in the height of the saddle point between Si, Ge and Sn. The path for B, a column III element, is shown in figure 7.3. Column IV and V elements are typified by the paths for Si and Ge concerted exchange, as shown in figures 4.2 and 5.2.

The resulting energy barriers are given in table 7.1, along with experimental values. In all cases, the CE barrier is close to the experimentally measured value for the diffusion activation energy. In some cases, the CE barrier we find is slightly higher, in other cases lower. In the cases of Ga, Ge, and Sn, the CE barrier is noticeably lower than the experimental value, suggesting that CE might play an important role in the diffusion of these elements in Si.

Our values for the activation energy of the concerted exchange mechanism agree well with the results obtained in the theoretical study by Nichols *et al.* [71] for As

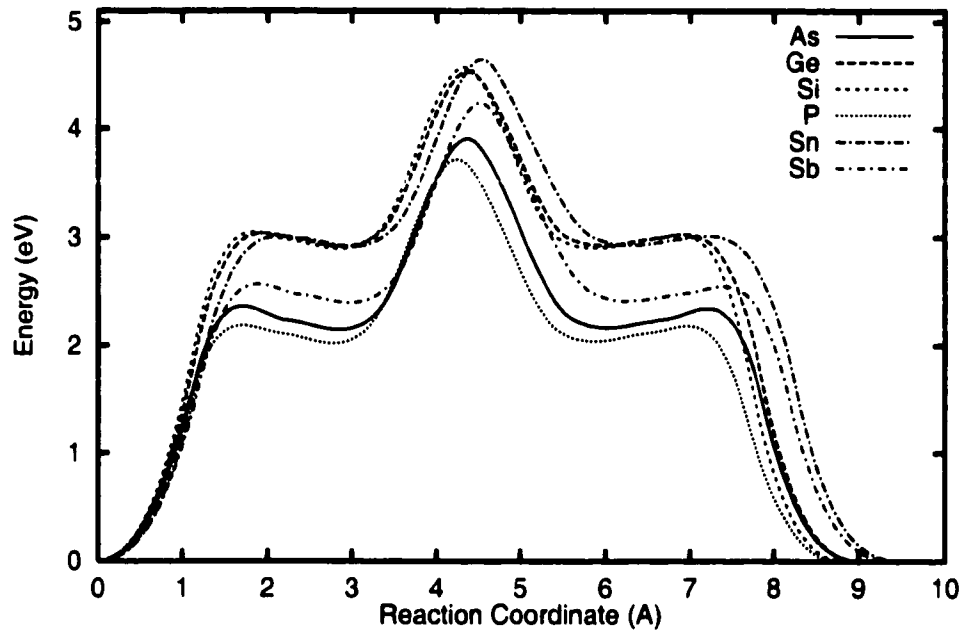


Figure 7.2 : Concerted exchange barriers for column IV and V elements in Si.

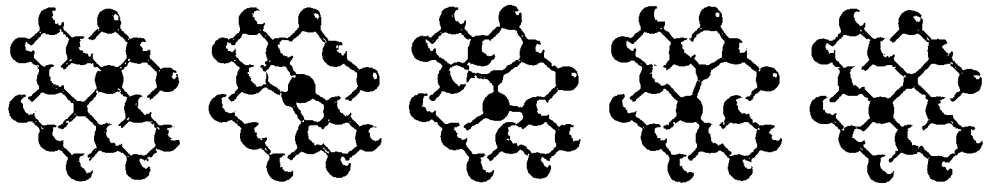


Figure 7.3 : The path in space executed by the concerted exchange mechanism for B in Si. The first and last frame are ideal lattice positions. The third frame is the system at the transition state. The smaller atom is the B atom and the other dark atom is the Si atom involved in the transition.

Table 7.1 : Energy barriers and prefactors for CE in Si for various dopants and impurities. In the case of column IV and V elements, the energy of the shoulder minimum is also give. Literature values of the diffusion activation energy and prefactor as determined by experiment are given for comparison.

Element	Shoulder Minimum (eV)	Calculated Barrier (eV)	Experimental Barrier (eV)
B	—	3.72	3.5[3]
Al	—	3.29	3.36[42]
Ga	—	3.42	3.75[42]
In	—	3.47	3.5[3]
Si	2.90	4.57	4.76[3]
Ge	2.92	4.54	4.7-5.3[20]
Sn	2.92	4.65	4.8[72]
P	2.02	3.72	3.68[3]
As	2.15	3.91	3.99[3]
Sb	2.41	4.24	3.88[3]

and Sb. We find values of 3.9 and 4.2 eV, respectively, and they found 3.9 and 4.5 eV, respectively. We do disagree significantly for B and P, by 1.0 and 0.9 eV, respectively. Nichols *et al.* did not relax the full path, but rather placed the impurity atom in the transition state found by Pandey [21]. Thus, their results are in all cases an upper bound only, and this is reflected in the comparison with our values.

7.5 Charged Systems

Dopants such as B and As are ubiquitous in Si technology because they readily accept (B) or donate (As) electrons from or to the Si crystal. This gives an excess of either positive or negative free charge carriers. Because of the importance of the charge states, and the likelihood that these dopants (and the others as well) will be charged in the real system, we have calculated the CE barrier for B- and As+. The barrier for B- is shown in figure 7.4 and that for As+ in figure 7.5. The barrier for CE of neutral B is 3.72 eV. For B-, it is 4.20 eV. The barrier increases after B has accepted an electron and become electronically active. The same is true for As. For neutral As, we find a barrier of 3.91 eV. Once the As has donated its electron, the barrier becomes 4.62 eV. If the charge is opposite, that is, if B is positively charged and As is negatively charged, the barrier is reduced. B+ has a barrier of 3.57 eV and As- has a barrier of 3.31 eV.

The qualitative shape of the potential energy curve changes for the B- system. It becomes similar to the shape seen for column IV and V elements, with a side minimum on the shoulder. This suggests that the shape of the curve and the bonding that occurs during the exchange depends on whether the Si atoms are fully coordinated. If they are, then there is a stable intermediate state for the exchanging atoms. Otherwise, the only stable configuration is to sit on the crystalline sites.

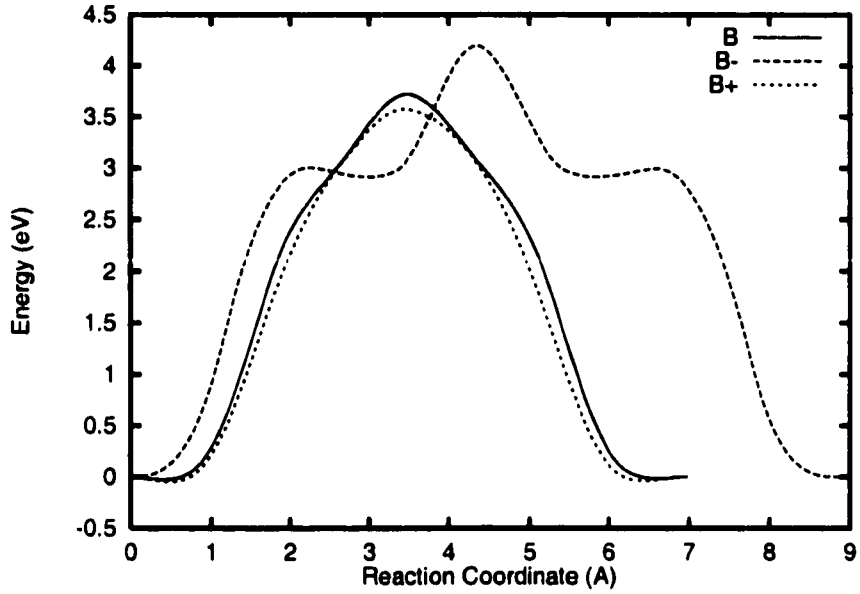


Figure 7.4 : Concerted exchange barrier for neutral, negatively charged, and positively charged B in Si.

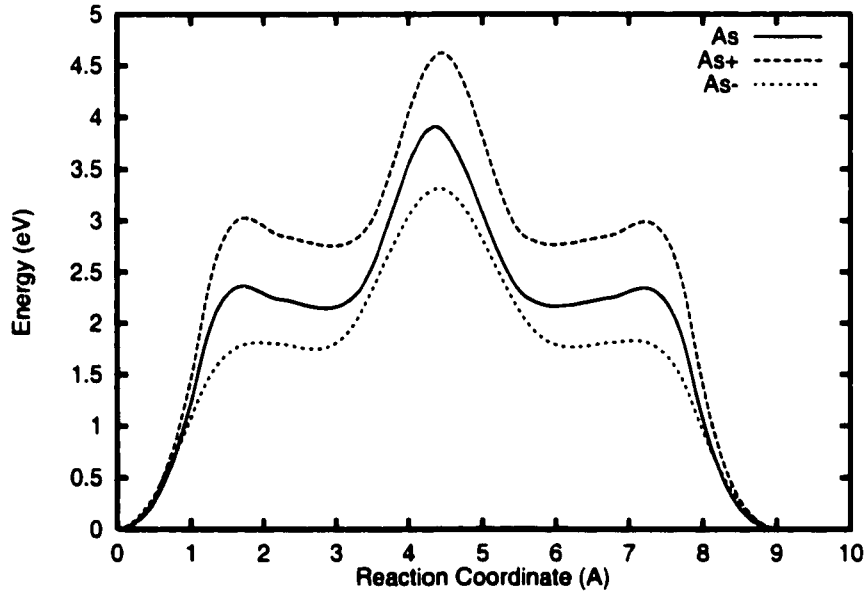


Figure 7.5 : Concerted exchange barrier for neutral, positively charged, and negatively charged As in Si.

7.6 Conclusions

In all cases looked at here, for common column III, IV and V dopants in Si, the DFT result for the CE energy barrier is close to the experimental range for the activation energy of diffusion in Si. In no case is the DFT barrier for CE so different from the experimental activation energy that it must automatically be discarded from further consideration. This might not be true when all possible diffusion mechanisms are analyzed via DFT, in which case it might be seen that vacancy (see chapter 8) or interstitial diffusion have even lower activation energy within the DFT/PW91 description than the CE mechanism. As we have seen previously for both Si self-diffusion and Ge diffusion in Si, CE is a negligible component of the total diffusion constant. However, without further information about the diffusion mechanisms due to interstitials and vacancies, these results for CE compare favorably with experiments, making it a potentially important component of diffusion. CE is a viable mechanism and must be considered in all cases examined here.

The charge state has a significant effect on the barrier height of the CE mechanism. For both B and As diffusion in Si, the CE barrier goes up as the diffusing species goes into its electrically active charge state, by 0.5 eV for B and 0.7 eV for As. The barrier, correspondingly, goes down if the species obtains the opposite charge. This results in barrier height differences of 0.6 eV for B and 1.3 eV for As.

It is unlikely that a donor such as As or an acceptor such as B will either capture an electron or release an electron, respectively. However, the carriers are weakly bound to the dopant atom. For As, the electron it donates is bound by on the order of 50 meV to the As atom. An electron that wanders near the As atom might be temporarily bound and, if another electron also comes by, the lower energy process might be activated. Charge transfer might occur, causing the transition state to have a different charge than the initial or final states. This would add a term to the diffusion constant of As (and B) that is dependent on the local electron density.

Chapter 8

VACANCY ASSISTED DIFFUSION OF DOPANTS IN SILICON

8.1 Introduction

Two relatively common donor dopants used in Si technology are As and Sb. Both are column V elements and thus have one electron that is weakly bound and able to break free to travel freely through the material. Both elements are also physically larger than Si and are experimentally seen to have a significant vacancy mediated component to their diffusion in Si. In the case of Sb, diffusion is thought to be almost entirely due to vacancies, while the picture is not as clear for As [24]. As diffusion is definitely assisted by interstitials and possibly also has a concerted exchange component. Our results from chapter 7 show that concerted exchange has a migration barrier comparable to experimental results and cannot be ruled out a priori. But, for both elements, the vacancy contribution is believed to be the largest component. We examine the vacancy assisted diffusion of both elements out to the third nearest neighbor, since, as previously discussed in the case of Ge (see section 5.5.1), the vacancy must move at least as far away as the third nearest neighbor site to cause net diffusion of a substitutional impurity. As with Ge, arsenic is characterized by an increase in vacancy-As_S binding energy as the $Si_V - As_S$ distance decreases. For Sb, however, the situation is different, with the large Sb atom preferring not to sit on a lattice site, but in the middle of the open space created by the vacancy. In addition, we find that the second nearest neighbor site is no longer a stable site for the vacancy.

8.2 Previous Work

Recent experiments have reported that the vacancy component of As diffusion in Si may be as much as 64% under the right conditions (1100°C) and that of Sb diffusion is even higher, being, for all practical purposes, 100% [24]. Arsenic is thought to diffuse both via interstitials and vacancies in bulk Si, with a concerted exchange component that might be anywhere from 0 to 60% of the total diffusion constant [24]. These studies consist of oxidation and nitridation experiments at two temperatures for four different dopants. Oxidation injects interstitials into Si while nitridation injects vacancies. The profiles of these dopants were measured by SIMS for anneals of 1 hour at 1100°C and 5 hours at 1000°C. This gives enough information to numerically solve the resulting system of equations for the concentration profiles for the fractional contribution of interstitials, vacancies and concerted exchange to the diffusion constant.

A previous DFT study, using LDA for exchange-correlation, calculated the formation energies of As and Sb vacancy complexes [71]. They found that a vacancy binds to substitutional As with an energy of 1.2 eV and to Sb also with an energy of 1.2 eV. They also found an energy of formation of Si_V bound to the substitutional impurity. For both As_S-Si_V and Sb_S-Si_V, they found a formation energy of 2.3 eV. They did not calculate migration barriers.

8.3 Arsenic

Here, we analyze the vacancy component of As diffusion.

Figure 8.1 shows the potential energy landscape as Si_V moves from the nearest neighbor site (nn) of As_S, swaps with As_S and then moves to the 2nd nearest neighbor site (2nn) and the 3rd nearest neighbor site (3nn). The energy of Si_V in these sites is: nn: 0.00 eV, 2nn: 0.82 eV, and 3nn: 0.90 eV. This is still far from the isolated Si_V+As_S energy, which is 1.34 eV (here, the Si_V and As_S geometries were calculated

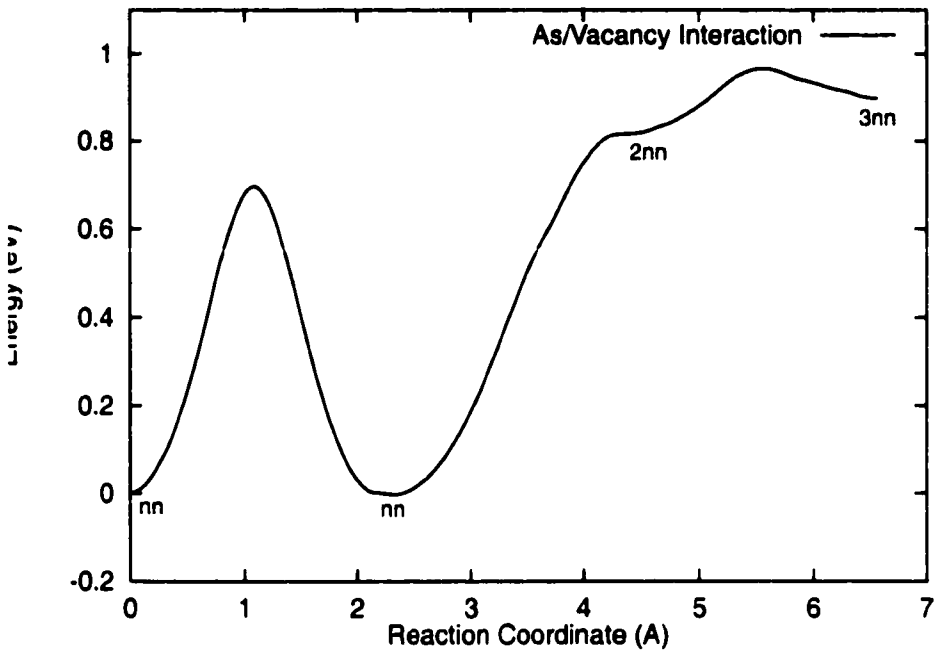


Figure 8.1 : Interaction of As_S and Si_V. The first two minima correspond to Si_V in the nearest neighbor site of As_S and the next two correspond to Si_V in the 2nd and 3rd nearest neighbor sites, respectively.

in separate cells). The barrier to swap As_S and Si_V is 0.70 eV. For Si_V to move from nn to 2nn, the barrier is 0.82 eV and the barrier to move from 2nn to 3nn is 0.15 eV (0.97 eV relative to nn). There is no barrier to return to the nn site from the 2nn site. The 2nn site is a metastable state. It represents an intermediate case between that of Ge and Sb. As seen before for Ge vacancy assisted diffusion in Si, the 2nn site is a distinct minimum. Here, it is only a metastable state, and a displacement towards the As atom will cause it to move into the nn site. The minimum energy path for these processes is very similar to that of Si_V near Ge_S shown in figures 5.5.3 and 5.5.3.

The energy difference between a bound As_S–Si_V pair compared to isolated As_S and Si_V is

$$E_{bind} = E_{\mathrm{As}_S+\mathrm{Si}_V} + E_{ideal} - E_{\mathrm{As}_S} - E_{\mathrm{Si}_V}. \quad (8.1)$$

Calculating this difference leads to the result that As_S effectively traps Si_V by an energy of 1.34 eV.

The formation energy of Si_V in bulk Si is 3.45 eV. The formation energy of the As_S-Si_V pair is this energy minus the binding energy of the pair, which results in a formation energy of 2.11 eV. This agrees well with the value of 2.3 eV found by Nichols *et al.* using LDA [71].

This leads to an overall activation energy of 3.08 eV. Experimental values for the activation energy for As diffusion in Si range from 4.05 to 4.34 eV [42]. So, once again, our value is too low by the typical DFT/PW91 error. This activation energy is also significantly lower, by about 0.8 eV, than the barrier we found for CE of As in Si (chapter 7). This suggests that vacancy-mediated diffusion will dominate for As diffusion in Si, though prefactors need to be calculated to be more conclusive.

8.4 Sb

Sb is larger than As and experimental studies suggest that it diffuses almost entirely via vacancies [3, 24].

The potential energy curve for Si_V diffusion near Sb_S is shown in figure 8.3. The landscape is more complicated for Sb than for As. Again, if a nearest neighbor Si of Sb is removed, the corresponding vacancy geometry is stable and has an energy which we define as 0.00 eV. However, there is an even more stable arrangement, with Sb sitting in the bond centered position between the original Sb_S and Si_V sites. This site, labeled as nn′, has an energy of -0.04 eV. When Si_V is placed in the 2nn site, the structure relaxes to Sb_S with Si_V in the nn site with no barrier. There is no stable 2nn site for this system. Si_V in the 3nn site is stable, and has an energy of 1.07 eV relative to nn. The path for nn to 3nn is shown in figure 8.2.

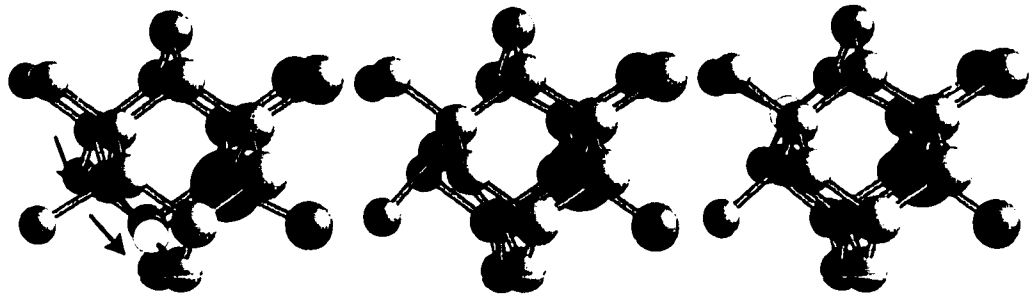


Figure 8.2 : The minimum energy path of Si_V motion near Sb_S. The vacancy moves from the nn site to the 3nn site via a concerted motion of two Si atoms.

The energy of the vacancy far from Sb_S relative to Sb_S-Si_V is 1.58 eV. Thus, Sb acts as an even stronger trap of Si_V than either As or Ge. This is what one would expect, as the larger Sb atom would like more space to relax in. Without the Si_V nearby, Sb_S causes a lot of local strain in the Si lattice. The presence of Si_V reduces this strain, resulting in a strong binding between Sb_S and Si_V.

The formation of the Sb_S-Si_V pair, then, is 1.87 eV relative to Sb_S in an otherwise perfect Si lattice. This is substantially less than that found by Nichols *et al.*, who found a value of 2.3 eV. They did use a smaller cell, which would not account as well for the larger strain introduced by the Sb atom. This is one possible explanation of the discrepancy. In addition, they used LDA while we used GGA for exchange-correlation. Lastly, they did not allow all atoms to relax and did not do a full relaxation of the cell, but rather did several displacements and fit the location of the minimum. Thus, there are some approximations made in their calculation that we have not made.

The energy barrier to move from Si_V in the nn site to the nn′ structure is 0.02 eV and the barrier to leave SbV to return to the nn site is 0.06 eV. Since there is no stable 2nn site, the barrier between nn and 3nn was calculated. This involves a concerted motion of two Si_S atoms moving to fill the nn Si_V site, leaving Si_V in the

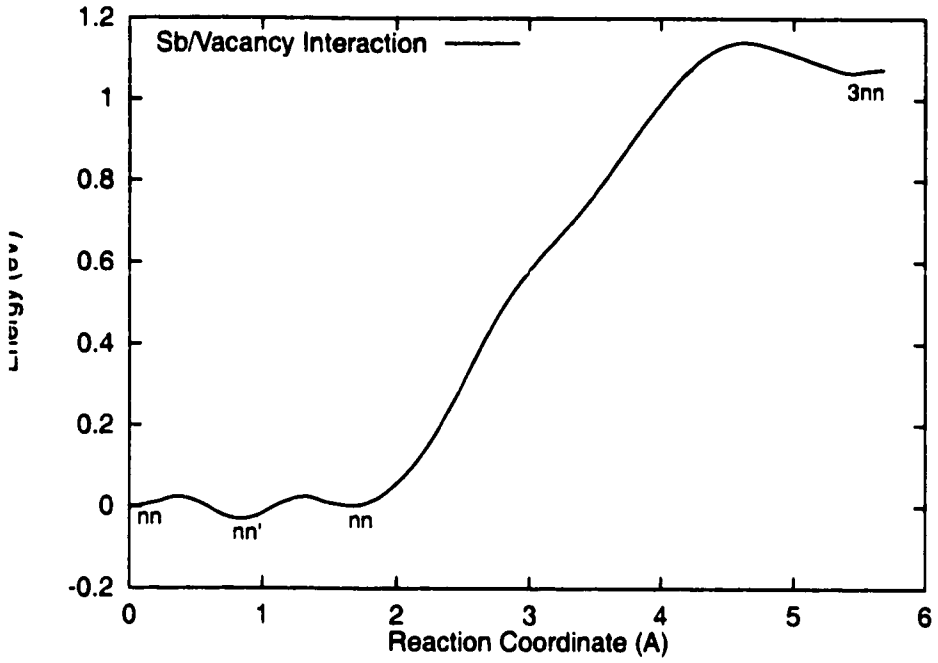


Figure 8.3 : Interaction of Sb_S and Si_V. The first and third minima represent Si_V in the nearest neighbor site of Sb_S. The second minimum is the combined SbV complex. The last minimum is Si_V in the 3nn site. There is no 2nn site for Sb. Placing the system in such a geometry relaxes instead to the nn configuration.

3nn site. The barrier for this process is 1.14 eV. The overall landscape is illustrated in figure 8.3.

The case of Sb continues the trend seen for Ge and As. The 2nn site, being a distinct minimum in Ge and a marginally stable state in As, completely disappears in the case of Sb.

Our overall activation energy for Sb diffusion is 3.01 eV. This is lower than experimental measurements have given, which tend to be around 4 eV (3.89 to 4.05 eV) [42]. The difference here is once again about 1 eV, presumably due to errors in the DFT/PW91 functional.

8.5 Conclusion

We have found the overall activation barrier for As diffusion due to vacancies to be 3.1 eV and for Sb diffusion due to vacancies to be 3.0 eV. Both of these values are about 1 eV lower than experiment. We saw the same discrepancy for both vacancy mediated diffusion of Ge in Si and Si self-diffusion due to vacancies. However, we do see a slight difference between vacancy mediated diffusion of As and Sb, with Sb diffusion being slightly easier by 0.1 eV. This agrees well with the experimental values reported by Fahey *et al.*. They report a spread of activation energies for As of 4.05 to 4.34 eV and for Sb of 3.89 to 4.05 eV. The median activation energy for As is 4.11 eV and, for Sb, 3.98 eV. These too differ by about 0.1 eV. Of course, as mentioned in chapter 7, As diffusion is thought to also be mediated by interstitials and possibly have a large concerted exchange component. Without calculating the prefactors, we cannot say if vacancy mediated diffusion will dominate for either As or Sb. The values we found for CE migration were 3.91 and 4.24 eV for As and Sb, respectively. These energies are much higher than the vacancy activation energies found here. If the prefactors for the various mechanisms for these two elements follow the same qualitative trends as for Ge in Si and Si self-diffusion, then we can expect that the prefactor for CE will not be larger than for vacancies (chapters 4 and 5). If anything, it will be lower. This suggests that the vacancy mechanism will dominate diffusion for both As and Sb, compared to the CE mechanism. To complete this work, prefactors need to be calculated and the interstitial mechanism also needs to be considered. These are subjects for future work.

Chapter 9

GERMANIUM SELF-DIFFUSION

9.1 Introduction

Self-diffusion of Ge has not received the same level of attention as Si self-diffusion has. The reason for this seems obvious: Si holds a much more important place in the semiconductor industry. However, as Ge becomes more common as a component in Si based technology, and the usefulness of Si/Ge superlattices continues to grow, a fundamental understanding of Ge self-diffusion will be necessary to model the growth of these structures. From a theoretical perspective, the similarities between Si and Ge (they are both column IV elements, thus having the same number of valence electrons, and they neighbor one another and have similar properties) suggest an immediate extension of the self-diffusion studies of Si to Ge. That is precisely what is described in this chapter.

9.2 Previous Work

While the amount of research concerning self-diffusion in Ge is not as great as in the case of Si, there have been some previous studies. Frank *et al.* give a summary of the results of Ge self-diffusion measurements done up to 1984 [20]. The enthalpy of activation ranges from 2.95 to 3.14 eV and the prefactors from 7.8 to 44 cm^2s^{-1}. These are mostly tracer diffusion experiments of ^{71}Ge in which the profiles were measured by sectioning, either by grinding or sputtering. The fact that these experiments are all fit well by an Ahrrenius form suggests that only one mechanism is active and the experiments reported by Frank suggest it is a vacancy mediated mechanism.

Haesslein and coworkers conducted perturbed angular correlation spectroscopy experiments in which Ge was doped with ^{111}In probes followed by defect introduction by electron irradiation at 77 K. If defects become trapped at the probes, an electric field gradient is induced at the probe nucleus which causes a nuclear interaction that can be measured. They were able to determine that the onset of long range migration of the neutral vacancy occurs at 200 K and of the positive interstitial at 220 K [73].

Raman scattering by optical phonons was used by Silveira and coworkers [74] to investigate $(^{70}\mathrm{Ge})_n(^{74}\mathrm{Ge})_m$ superlattices annealed at 500°C. These experimental measurements are compared to theoretical calculations within a planar force constant model for phonon frequencies and the bond polarizability approximation for the intensities. They obtain a diffusion constant at 500°C of 5.5×10^{-24} m^{2}s^{-1}, though they do not obtain enough temperature resolved data to extract the activation energy and prefactor for diffusion.

Finally, Fuchs and coworkers used secondary ion mass spectroscopy (SIMS) to measure the diffusion profiles of isotropically enriched layers of Ge at five different temperatures: 543, 586, 605, 636, and 690°C. They found the activation enthalpy to be 3.0 eV and a self-diffusion entropy of about $9k_B$ [75]. To get this entropy, they had to assume $\nu = 8 \times 10^{12}\mathrm{s}^{-1}$ and $a = 5.65$ Åin equation 2.4. Their full diffusion constant is

$$D = 12 \exp^{(-3.0/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \tag{9.1}$$

These numbers fall within the range reported by Frank and coworkers mentioned above.

9.3 Calculations

The calculations reported in this chapter were done as described in section 3.3, except that the energy cutoff used was 174 eV, not the 188 eV used in the Si calculations.

The Ge pseudopotential is smoother ("softer") and fewer plane waves can be used in the basis set.

9.4 Concerted Exchange

The concerted exchange process in Ge is very similar to what is found in Si. The potential energy curve exhibits side minima on both shoulders at an energy of 2.47 eV. The overall height of the curve is 3.57 eV. This is about 1 eV less than the barrier for CE in Si, or for Ge CE in Si.

The prefactor for this process is 4.5×10^{14} s^{-1} corresponding to a diffusion prefactor of 1.1 $cm^2 s^{-1}$. This was found using a displacement of 0.01 Å following the procedure outlined in chapter 2.

9.5 Self-Interstitial

9.5.1 Formation Energy

In Si, the interstitial has three stable geometries. In the order of decreasing stability, they are: the split interstitial Si_I^X, the hexagonal interstitial Si_I^H, and the tetrahedral interstitial Si_I^T. These structures were analyzed for Ge, in which only two distinct structures emerge: Ge_I^X and Ge_I^T. A Ge_I^H structure will relax to Ge_I^T. The relative energetics of these two structures depends greatly on the functional used in the DFT calculations, LDA or GGA. Using the energy of Ge_I^X as the zero of energy, the energy of Ge_I^T is 0.06 eV in LDA, but -0.21 eV in GGA. Thus, unlike in Si, Ge_I^X is not the most stable interstitial structure.

The energy of formation of Ge_I^T relative to the perfect Ge crystal is 2.85 eV using GGA. This is the cost for a Ge atom to move from a bulk lattice site far away and to become an interstitial.

9.5.2 Migration Energy

The migration barrier between these two states was calculated. The barrier to go from Ge_I^T to another Ge_I^T in a neighboring hex channel is 0.71 eV. The barrier to go from Ge_I^T to the nearest Ge_I^X is smaller, 0.51 eV. To go from one Ge_I^X configuration to another on the same lattice site (this is a rotation of Ge_I^X onto a different $\langle 110 \rangle$ direction), the interstitial must pass through a Ge_I^T configuration. However, from any Ge_I^T configuration, there are 4 equivalent paths by which it can become Ge_I^X. The lowest energy pathway, then, for Ge interstitial diffusion is $Ge_I^T \rightarrow Ge_I^X \rightarrow Ge_I^T$. This is different from the case of Si, in which it is no more costly to diffuse down the channels as Si_I^H than to move between Si_I^H and Si_I^X configurations. The potential energy landscape is shown in figure 9.1. As shown in figure 9.2, which illustrates the full path of Ge_I from Ge_I^T to Ge_I^T to Ge_I^X, the transition state for $Ge_I^T \rightarrow Ge_I^T$ is a hexagonal structure (Ge_I^H).

9.5.3 Migration Entropy

The prefactor for the $Ge_I^T \rightarrow Ge_I^X$ process is relatively low, 3×10^{12} s^{-1}. Again, this was found using a displacement of 0.01 Å.

9.5.4 Entropy of Formation

The vibrational entropy of formation of the Ge self-interstitial is found to be higher than for the Si self-interstitial. The vibrational entropy of formation is shown in table 9.1 as a function of the number of atoms included in the calculation for both the local harmonic approximation (LHA) and including the coupling of vibrational modes between atoms (FHA). Also shown is the value of an extrapolation to infinity based upon an exponential function. We find an entropy of $3.70k_B$ for the interstitial in the full harmonic approximation. An extrapolation of the entropy of formation

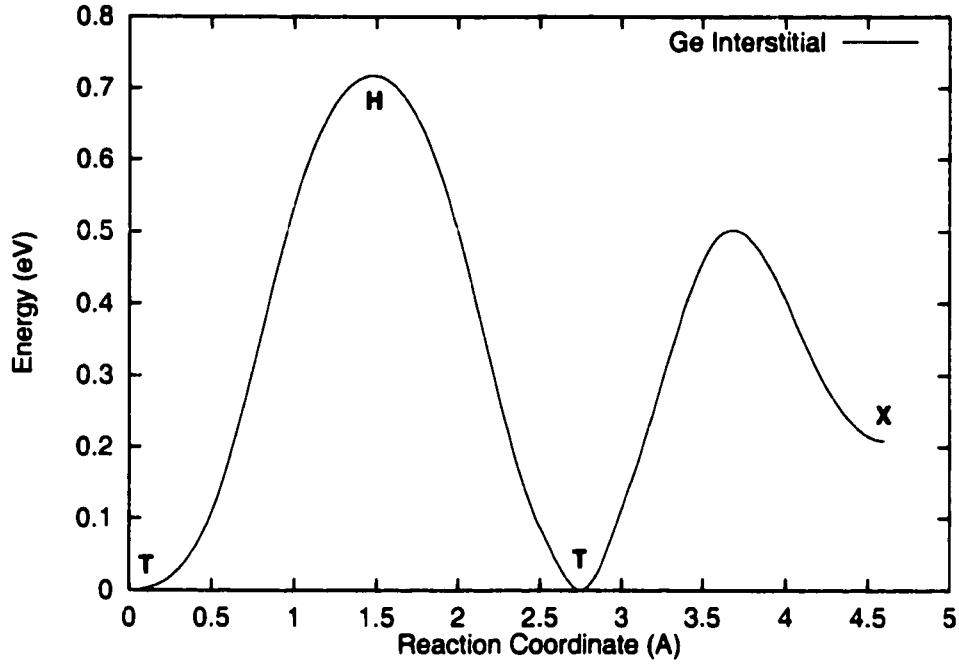


Figure 9.1 : Potential energy landscape of Ge self-interstitial diffusion through Ge crystal. The first and second minima correspond to the Ge interstitial being in the tetrahedral site. The third minimum represents the interstitial in the split [110] site. Diffusion through the crystal is possible by way of $Ge_I^T \rightarrow Ge_I^X \rightarrow Ge_I^T$.

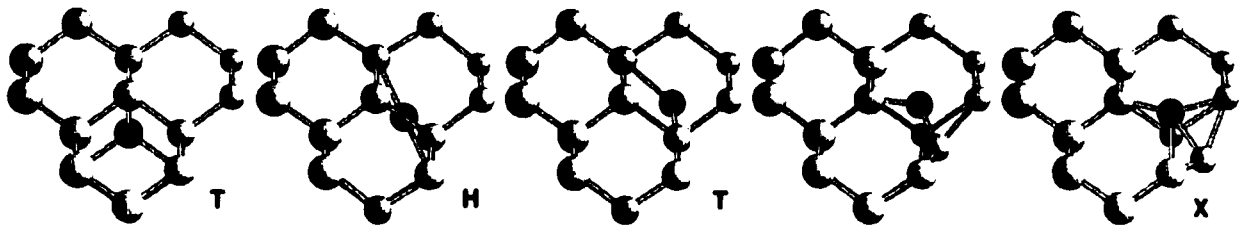


Figure 9.2 : The path in space of Ge_I motion as it goes from the T geometry, through H to T again, and then to X.

Table 9.1 : Entropy of formation for the interstitial and the vacancy using both the local harmonic approximation (LHA) and the full harmonic approximation (FHA).

Interstitial			Vacancy		
Number of atoms included	Entropy (k_B) LHA	FHA	Number of atoms included	Entropy (k_B) LHA	FHA
2	1.36	1.56			
6	2.33	2.43	4	4.30	3.54
18	3.00	3.37	16	6.99	6.43
30	3.25	3.66	28	7.87	7.84
Extrapolation	3.20	3.70	Extrapolation	8.30	9.18

versus included degrees of freedom is shown for both the LHA and FHA. [1]

The vibrational entropy of formation for the Ge self-interstitial is very close to the value we found for the Si self-interstitial, being only $0.14k_B$ higher.

9.6 Vacancy

9.6.1 Migration and Formation Energy

The structure of Ge_V is similar to that of Si, with the four neighbors of Ge_V relaxing inwards by 0.36 Å. The energy of formation of the vacancy, relative to a perfect crystal of Ge, is only 1.88 eV. The barrier to exchange of Ge_S and Ge_V is 0.17 eV. The potential energy profile is given in figure 9.3. The path taken by Ge_V in the lattice is the same as the exchange of Ge_S and Si_V in the Si lattice (see figure 5.5.3).

[1]The functional form used to extrapolate the calculated values of the entropy of formation was $A\exp(Bx) + C$, where A, B, and C are fitting parameters and x is the effective number of atoms included, the number quoted in the text. Convergence to the extrapolated value (the value as x goes to infinity, equivalent to C) is considered to be the value of x for which the fit is within about $0.1k_B$ of C.

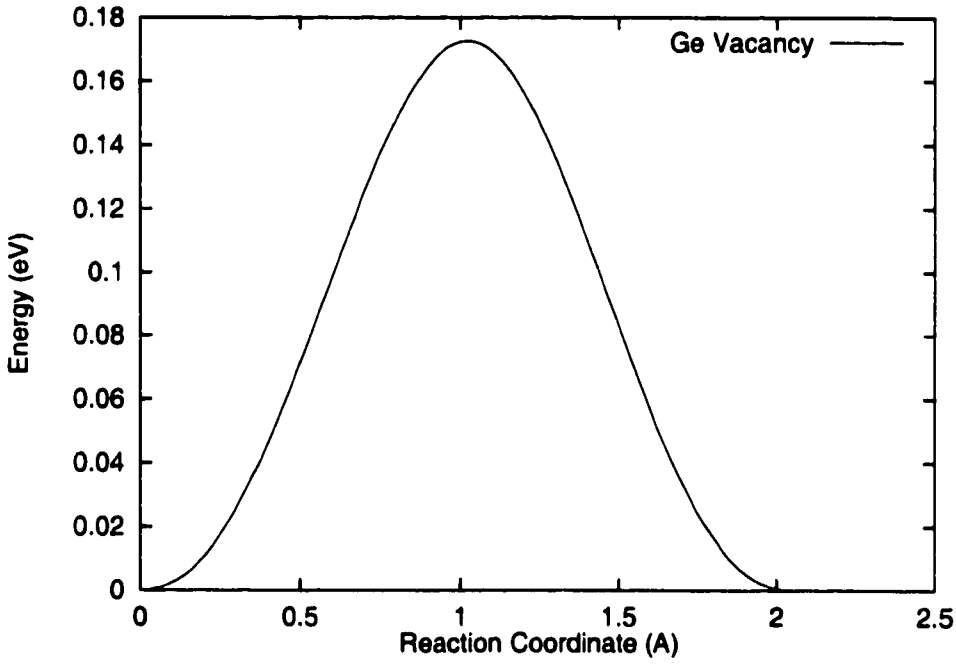


Figure 9.3 : Potential energy landscape of vacancy diffusion through Ge crystal.

9.6.2 Entropy of Migration and Formation

The prefactor for migration for the vacancy to move from the nearest neighbor position to the third nearest neighbor position is 2.4×10^{12} s^{-1}. This was found using a displacement of 0.01 Å.

The entropy of formation of the vacancy, as listed in table 9.1, is $9.18k_B$ in the full harmonic approximation. The entropy of the vacancy is much higher than we found for the vacancy in Si. We find that the Ge vacancy has a vibrational entropy of formation that is $3.75k_B$ higher than in Si. It also converges much more slowly, with the extrapolated value not being met until 70 atoms are effectively included, while for the Si vacancy only 35 atoms need to be included.

Combined with the configurational entropy of $1.1k_B$ for the vacancy, we obtain total entropy of formation for the vacancy of $10.28k_B$.

9.7 Results

To summarize the above calculations, for the vacancy, $E_f = 1.88$ eV and $E_m = 0.17$ eV and for the interstitial, $E_f = 2.85$ eV and $E_m = 0.51$ eV. It is both more costly to form Ge_I than Ge_V, and more costly to migrate through the crystal. The overall activation energy for vacancy diffusion is only 2.05 eV and for interstitial diffusion 3.36 eV. The prefactors for migration are 4.5×10^{14} s^{-1}, 3.0×10^{12} s^{-1}, and 2.4×10^{12} s^{-1} for concerted exchange, interstitial and vacancy motion, respectively, and the entropy of formation for these mechanisms are $1.8k_B$, $5.5k_B$, and $10.3k_B$, respectively.

The diffusion constants that we predict are (assuming $a = 2.5$ Å for the vacancy and concerted exchange and $a = 1.25$ Å for the interstitial):

$$D_I = 0.08\exp^{(-3.4/k_BT)}\,\mathrm{cm^2s^{-1}} \tag{9.2}$$

$$D_V = 15\exp^{(-2.1/k_BT)}\,\mathrm{cm^2s^{-1}} \tag{9.3}$$

$$D_{CE} = 1.1\exp^{(-3.6/k_BT)}\,\mathrm{cm^2s^{-1}}. \tag{9.4}$$

The fractional contribution of each component of diffusion at 1050°C is

$$f_I = 5 \times 10^{-8}, \tag{9.5}$$

$$f_V = 1.0, \tag{9.6}$$

$$f_{CE} = 1.3 \times 10^{-7}. \tag{9.7}$$

The vacancy mechanism completely dominates self-diffusion in Ge.

Our activation energy for vacancies, which dominate diffusion in the DFT description of Ge self-diffusion, is about 1 eV lower than the value found experimentally. The source of this error is probably the same as for Si self-diffusion: the PW91 exchange-correlation functional does not adequately describe the electron-electron interaction, leading to a formation energy that is too low. However, we would expect this to be true for both interstitials and vacancies, and the domination of diffusion by vacancies would still appear. To compare to the experiment by Silveira *et al.*, the diffusion constant due to vacancy diffusion at 500°C is predicted to be 6.3×10^{-17} m^2s^{-1}. If the activation energy of vacancy diffusion were increased by 1 eV, this would become 1.9×10^{-23} m^2s^{-1}, which agrees very well with the results of Silveira, 5.5×10^{-24} m^2s^{-1}. Our prefactor for vacancy diffusion, about 15 cm^2s^{-1}, is close to the 8 to 40 cm^2s^{-1} found experimentally. The good agreement in the case of Ge self-diffusion, while Si self-diffusion prefactors were much too low, may be due to the fact that the Ge self-diffusion experiments were conducted at a low temperature where the harmonic approximation is more appropriate.

9.8 Conclusion

As in the case for Si self-diffusion, as well as Ge diffusion in Si, our overall activation energies are about 1 eV lower than experiment. However, our prefactor for vacancy mediated diffusion agrees well with the experimental prefactor for Ge self-diffusion.

Self-diffusion in Ge is likely be dominated by the vacancy mechanism. Not only does our prefactor for vacancy diffusion agree well with experiment, but the activation energy for vacancy mediated diffusion is over 1 eV lower than for the other mechanisms. for Si self-diffusion did not.

If we adjust our activation energy by 1 eV, the overall diffusivity we calculate agrees very well with the experimental values, within a factor of 2 to 3. As we will see in chapter 10, there is good reason to believe that the DFT value for the formation energy for Si defects is low by about 1 eV. If the same is true for Ge, then our adjusted results here would agree quantitatively with experiment. Such a calculation for Ge is a topic for future work.

Chapter 10

FIXING THE ENERGY OF FORMATION

10.1 Introduction

As has been mentioned throughout this work, plane-wave DFT using the PW91 exchange-correlation functional seriously underestimates the activation energy for diffusion in Si as compared with experiment. Si self-diffusion via interstitials has an activation energy of 4.0 eV under PW91, but recent experiments put this number at 4.68 eV [27] to 4.95 eV [26]. PW91 is low by somewhere between 0.7 and 1 eV.

The activation energy E_a is composed of two parts, the energy of formation E_f and the energy of migration E_m. The PW91 error in E_a could be in the calculated value for E_f, E_m, or both terms. Diffusion quantum Monte Carlo (DMC) calculation have shown that the error is in the PW91 value for E_f for the interstitial. The DMC calculations give a 1.1 eV correction to the PW91 number. The same study finds that PW91 also underestimates E_m for the concerted exchange by 0.98 eV [38]. DMC involves recasting the Schrödinger equation by changing real time into imaginary time, which makes the Schrödinger equation look like a diffusion equation. If the equation is then evolved in time for some initial guess, the long time limit will give the solution for the ground state of the Schrödinger equation. This is in principle exact, but to be computationally efficient, some approximations are made.

The most severe approximation made in DFT, the one approximation that cannot be improved upon by increasing the basis set or the k-point sampling or some other systematic procedure, is the approximation of the exchange-correlation functional. In this work, the GGA functional, as formulated by Perdew and Wang (referenced here as

PW91), was used [4]. This functional uses a purely local description of exchange. This means that exchange between electrons is only considered if they overlap spatially.

Other functionals have been developed that have non-local contributions to the exchange term of the exchange-correlation functional. An example of such a functional is the B3LYP functional [15]. Unfortunately, this functional can only be used in cluster calculations, not in periodic calculations such as those done with Vasp. So, even though better functionals exist, they cannot be used in codes, such as Vasp, that describe defects in bulk structures, such as the interstitial and vacancy calculations done in previous chapters.

What we propose here is to use cluster calculations to estimate the correction to PW91 results one would expect if a calculation using the B3LYP functional were used. That is, we know, for example, what periodic system calculations using PW91 give for E_f for an interstitial. We want to estimate what B3LYP would give for E_f. To do this, we extract clusters around the defect and the equivalent bulk cell and calculate the difference using both PW91 and B3LYP with the Gaussian code [76]. Gaussian treats clusters, not periodic systems, and the energy of a cluster can be calculated with both functionals. The difference in the formation energy between the cluster structures for PW91 and B3LYP will give us an estimate of the correction to E_f or E_m for the periodic system if we could use B3LYP in that case.

It should be noted that, in principle, defect structures could be studied by using large clusters. However, there are two problems with using cluster calculations to study defect structures in bulk materials. First, bulk properties are never described correctly. There is no such thing as a cohesive energy, since, in a cluster, the atoms are never equivalent. There are always surface atoms that differ from interior atoms. So, finding the energy of a defect structure compared to a bulk system of the same number of atoms is difficult. Second, as will be shown, we have found that to get convergence in the difference between PW91 and B3LYP energies for energies of formation, a difference of a difference in energies, we need clusters of nearly 30 atoms. This can

take a long time, on the order of weeks of CPU time, to converge in some cases. The calculations that would be necessary to converge energies of formation would be much longer and, at this moment, are intractable.

10.2 Correction for Concerted Exchange

Calculating the B3LYP correction to the PW91 value for the energy of activation of the concerted exchange process is straight forward. The number of atoms in the initial state and the transition state are equal. Therefore, we can extract a cluster of the both structures and calculate the energy difference of these two clusters with PW91 and B3LYP. The difference of this difference is the estimate of the energy correction for the CE activation energy.

Table 10.1 gives the energy correction found for the CE activation energy as a function of cluster size. Each cluster is a sphere around the bond center of the two atoms involved in the CE mechanism. As can be seen, convergence in the difference of the difference is seen for cluster sizes of 20 atoms. Going to the next largest cluster changes the correction by only 0.1 eV. Extrapolation of these points via an exponential expression gives 0.75 eV as the correction for the CE migration barrier. This agrees quite well with the correction obtained by DMC, 0.98 eV. Part of the difference could be due to a difference in the transition state used in the calculation. If this correction is added to the result previously found for the CE activation energy, we obtain a value of 5.32 eV. The extrapolated curves for this and the other cluster systems analyzed here are given in figure 10.1.

10.3 Correction for Defect Structures

In the following, the superscript c in E^c refers to the cluster calculation, E_I is the interstitial structure energy, E_{bulk} is the perfect crystal structure, E_{Si} is the energy of a free Si atom, and N is the number of atoms in the perfect crystal structure. Also,

Table 10.1 : Energy correction for the concerted exchange activation energy versus cluster size.

Cluster Size	Energy Correction (eV)
8	0.57
20	0.73
32	0.74
Extrapolation	0.75

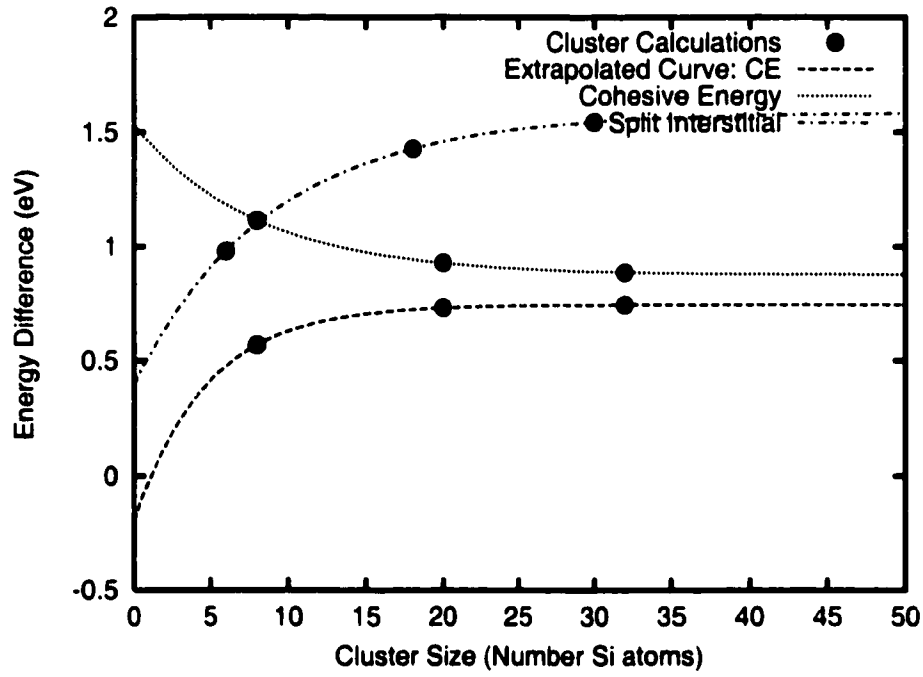


Figure 10.1 : Extrapolation of cluster calculations for the systems analyzed. The points represent the calculations done for different cluster sizes, while the lines are exponential curves fitted to the cluster points. Shown are the curves for the "cohesive energy", the CE migration energy correction, and the split interstitial formation energy correction.

to distinguish between calculations with the PW91 exchange-correlation functional and the B3LYP functional, E is used for PW91 and Σ for B3LYP.

For an interstitial, the energy of formation, E_f, is defined as

$$E_f = E_I - \frac{N+1}{N} E_{bulk}. \tag{10.1}$$

This measures the energy cost to take one atom from a perfect crystal lattice site far away and create an interstitial. We can just as easily measure the formation energy relative to a free atom:

$$E_f' = E_I - E_{bulk} - E_{Si} \tag{10.2}$$

This is the energy cost of bring an atom from the vacuum and placing it into an interstitial site in the crystal. E_f can be written in terms of E_f' as

$$E_f = E_f' - \frac{1}{N} E_{bulk} + E_{Si}. \tag{10.3}$$

The reason for introducing E_f' is because we want to calculate the correction to the formation energy with a cluster, but we cannot use a cluster to calculate E_{bulk} such that we can estimate the cohesive energy, E_{coh} as $\frac{1}{N} E_{bulk}$. By introducing E_f', the reference system is now a free atom, which can be calculated with a cluster. Pictorial definitions of both E_f and E_f' are shown in figure 10.2.

We can calculate E_f' using clusters. The relation is the same as for the plane wave PW91 calculation described before:

$$E_f'^c = E_I^c - E_{bulk}^c - E_{Si}^c. \tag{10.4}$$

Now, this quantity can be calculated for both the PW91 and the B3LYP functionals, as it is a quantity well defined for a cluster. We can define the correction energy, then, as

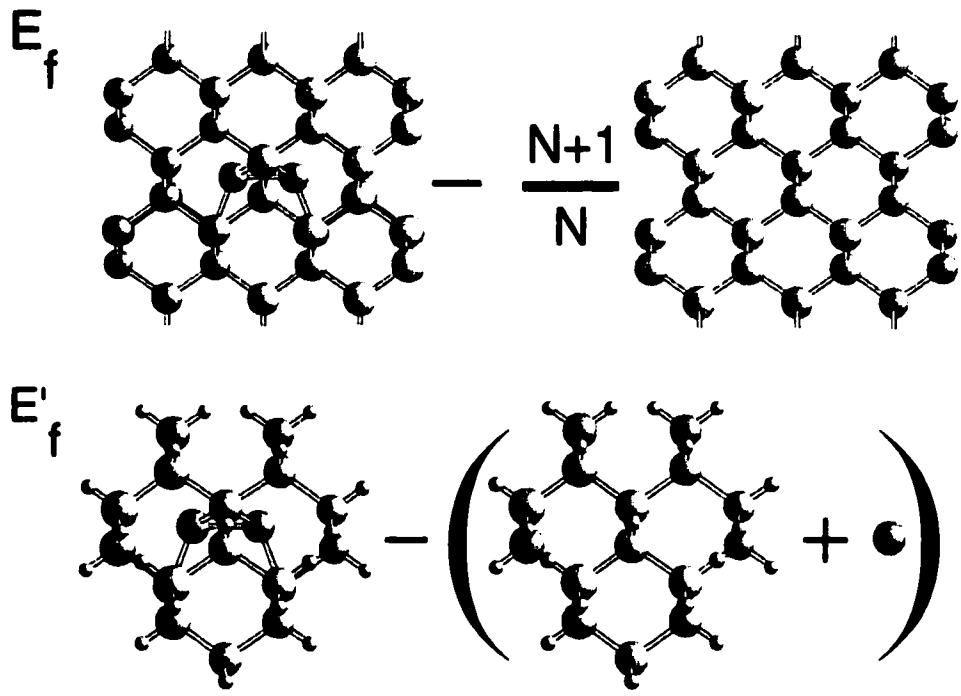


Figure 10.2 : Graphical depiction of the definition of E_f and E_f'. E_f is the formation energy of the interstitial relative to an equivalent number of bulk atoms, while E_f' is the formation energy relative to the same number of bulk atoms and one free Si atom.

$$\Delta E_f^{'c} = \Sigma_f^{'c} - E_f^{'c}, \tag{10.5}$$

where, again, Σ represents a quantity calculated using B3LYP and E a quantity using PW91.

If we now assume that this energy correction, found for the cluster structures, would be the same for the periodic system if we could calculate that system with B3LYP, we then have

$$\Sigma_f^{'} \approx E_f^{'} + \Delta E_f^{'c}. \tag{10.6}$$

Finally, we can then solve for the energy of formation of the periodic system relative to the bulk crystal using B3LYP:

$$\begin{aligned}
\Sigma_f &\equiv \Sigma_f^{'} - \frac{1}{N}\Sigma_{bulk} + \Sigma_{Si} \\
&= E_f^{'} + \Delta E_f^{'c} - \frac{1}{N}\Sigma_{bulk} + \Sigma_{Si} \\
&= E_f + \Delta E_f^{'c} + \frac{1}{N}\left(E_{bulk} - \Sigma_{bulk}\right) + \left(\Sigma_{Si} - E_{Si}\right) \qquad (10.7) \\
&\equiv E_f + \Delta E_f^{'c} + \Delta E_{coh}. \qquad (10.8)
\end{aligned}$$

The quantities Σ_{Si}, Σ_{bulk}, E_{Si} and E_{bulk} refer to the energy of periodic systems. They can be estimated using equivalent cluster structures.

The procedure described above should result in an over estimation of the B3LYP correction to PW91 numbers. It should give us an upper bound. We are extracting a structure that was relaxed via a periodic PW91 calculation and finding the energy of that structure treated as a cluster with PW91 and B3LYP. In neither case is the cluster structure relaxed. If a B3LYP calculation of a periodic system could be done and the structure relaxed, a lower energy would be obtained and the energy correction would decrease somewhat. So, these energy corrections should be viewed as an upper bound to the difference between the DFT/PW91 values for E_f and the "true" values.

10.4 Results of Cluster Calculations

10.4.1 Interstitial

In equation 10.8, we must estimate ΔE_{coh}, which is the difference in cohesive energy predicted by PW91 and B3LYP for the periodic system, with the same difference for the clusters. Both Σ_{Si} and E_{Si} are well defined quantities, but the other term in the definition of ΔE_{coh}, $\frac{1}{N}(E_{bulk} - \Sigma_{bulk})$, is ill-defined, as it assumes each Si atom in the system is equivalent. For a cluster, the surface atoms are not equivalent to the interior atoms and, in principle, one would have to use an infinite cluster, in which the surface atoms compose a negligible fraction of the total number, in order to calculate this quantity. However, we can model the energy of the cluster as

$$E = Ah + Bs \tag{10.9}$$

where h is the number of hydrogen (or "surface") atoms and s is the number of silicon ("bulk") atoms. Fitting the energies we obtain for six different cluster sizes, containing 5, 8, 17, 20, 29 and 32 atoms each, we obtain values for B, the energy per Si atom, of -7874.41 eV using PW91 and -7875.03 eV using B3LYP, giving a difference in B, or, effectively, a difference in the bulk Si energy per atom between PW91 and B3LYP of -0.62 eV. This, coupled with the energy difference between PW91 and B3LYP for the free atom of -1.20 eV, gives a ΔE_{coh} of 0.58 eV.

The difference in formation energy, as defined in equation 10.2, is given in table 10.2 for three different cluster sizes for the Si split interstitial. The complete correction to the periodic PW91 calculations, as described by equation 10.8, is also given. As can be seen, the B3LYP calculations give a much higher value for the formation energy than do the PW91 calculations. In fact, the values calculated over correct for the error we saw previously, in which the PW91 values for E_f were underestimated by 0.7 to 1 eV. Here, we find a correction to the PW91 values of 1.01 eV. This is in good agreement with the DMC calculation which found a correction of 1.12 eV.

Table 10.2 : B3LYP correction to the PW91 energy of formation of the Si split interstitial. The second column is the correction to the energy cost of taking a free atom and placing it in the interstitial site, while the third column is the cost of taking an atom from the bulk and placing it in the interstitial site.

Interstitial Cluster Size	Difference for E_f' (eV)	Difference for E_f (eV)
6	0.98	0.40
18	1.43	0.85
30	1.54	0.96
Extrapolation	1.59	1.01

It is possible that the B3LYP correction found for one interstitial structure could be very different from another. That is, the correction to the hexagonal interstitial might not be the same as for the split interstitial. We have tested this for two different hexagonal cluster sizes, containing 7 and 15 Si atoms. We find the correction to E_f' to be 1.41 and 1.64 eV, respectively. This is slightly higher than the value we found for the split interstitial, but it is not a large difference, being about 0.2 eV for roughly the same size clusters (15 Si atoms here versus 18 Si atoms for the split interstitial cluster). So, the B3LYP correction to the formation energy of interstitial structures is about the same for both important structures and we will just use the split interstitial correction as an estimate for the correction to the energy of formation for all interstitial structures.

One note should be made about basis set convergence. All of the calculations mentioned above used the 6-31G* basis set, which consists of Gaussian functions as the basis for the electronic wavefunction. We have tested a larger basis set to see if the results are dependent on basis set choice. Using the CC-PVTZ basis set, which has more Gaussians than the 6-31G* basis set, we have found both the correction to the CE energy of migration and the split interstitial formation energy E_f'. For the

concerted exchange ΔE_m, we find that the correction using the 6-31G* basis set is 0.57 eV and for the CC-PVTZ basis set it is 0.56 eV. Thus, the 6-31G* basis set gives good corrections. It should be noted that the activation energy E_m, calculated with PW91 for the CE mechanism in the cluster goes from 6.22 eV to 5.97 eV when the basis set is increased. This is a difference of 0.25 eV. A similar difference exists for the B3LYP calculation. Therefore, while the activation energy does change with basis set, the B3LYP correction to the activation energy does not change.

Basis set truncation may be the reason that our calculations with the clusters do not agree on the value of E_f' for the split interstitial using PW91. Using plane waves as basis functions, we have calculated E_f' for the same clusters used in the Gaussian calculation. We find that PW91 calculations of an 18 atom cluster give $E_f' = -0.40$ eV when the 6-31G* basis set is used, but a value of -1.14 eV is obtained with a plane wave basis set. If the two basis sets include enough functions the two cluster calculations should agree on the value of E_f'. If the basis set is increased to CC-PVTZ, the Gaussian calculation gives a value for E_f' of -1.32 eV, which agrees well with the plane wave values. This does not change the result for the energy correction, though, which is 1.45 eV for this cluster size, and differs from the smaller basis set by only 0.02 eV. Thus, while our results are not converged with respect to basis set for the absolute values of E_f', our differences between the B3LYP and PW91 values for E_f' are converged.

As was mentioned previously, in principle the error in the DFT value for E_a could be in either the energy of formation E_f or the energy of migration E_m. We calculated the B3LYP correction to E_m for the Si hex interstitial migration barrier. We used a cluster of 8 Si atoms for both the stable hexagonal interstitial and the transition state. The migration barrier found with plane-wave DFT using PW91 was 0.22 eV. We find a correction using B3LYP of only 0.015 eV, or less than 7% (compared to the nearly 35% correction found for E_f). If we use a larger cluster containing 15 atoms, we find a larger correction, 0.15 eV. This is still small compared to the overall

activation energy. However, it is a large change to the migration energy and further studies need to be conducted to determine this correction more precisely.

It is reasonable to find a smaller correction to E_m because migration of an interstitial atom in the crystal involves less of a rearrangement of charge than does the creation of the interstitial. There is better cancellation of errors in the calculation of migration barriers than in the calculation of formation energy.

10.4.2 Vacancy

The same analysis that was used to find the B3LYP correction to PW91 formation energies of the interstitial can be applied to the vacancy structure. The only difference is that we are comparing a structure of one fewer atoms with the bulk structure, instead of one more. However, the electronic structure of the vacancy is much more complicated, since there are 4 unpaired electrons.

Thus,

$$E_f = E_V - \frac{N-1}{N} E_{bulk} \tag{10.10}$$

$$E_f' = E_V - E_{bulk} + E_{Si}. \tag{10.11}$$

This leads to a correction to E_f of

$$\Sigma_f = E_f + \Delta E_f'^c - \Delta E_{coh}. \tag{10.12}$$

The results of the vacancy cluster calculations are given in table 10.3.

We have also calculated the correction to the vacancy transition state. Here, we have only been able to calculate the correction for one cluster size, a cluster containing 16 atoms. We find a correction of 0.35 eV, a significant correction. Including this correction and the correction to the vacancy formation energy, we obtain an activation energy for vacancy mediated diffusion of 4.56 eV.

Table 10.3 : B3LYP correction to the PW91 energy of formation of the Si vacancy. The second column is the correction to the energy cost of taking an atom from the bulk structure and placing it into the vacuum far from the cluster. The third column is the cost of taking an atom from the bulk structure and placing it in another lattice position far from the generated vacancy.

Vacancy Cluster Size	Difference for E_f' (eV)	Difference for E_f (eV)
4	-0.01	0.57
16	-0.06	0.52
28	-0.10	0.48
Extrapolated	-0.20	0.38

10.5 Future Considerations

Currently, the form of the B3LYP functional makes it impossible to implement in a periodic DFT code. The exact exchange that is included in the B3LYP functional is of infinite range and, while some implementations exist (for example in the CRYSTAL code), the calculations are two orders of magnitude more costly and forces are not available. A possible solution would be to treat some cluster of atoms within the supercell using B3LYP, or some other functional known to describe exchange better than PW91. The remaining atoms in the cell would use PW91 to describe the exchange-correlation energy. This is similar in spirit to the so-called QMMM methods in which a quantum mechanical description is used to describe the central cluster, while the remaining system is described by an empirical potential. Such a method would allow for full relaxations of the defect structures while using B3LYP to describe the interaction amongst those atoms composing the structure.

10.6 Results

We find that the correction to the interstitial formation energy using B3LYP is about 1.01 eV for the split interstitial and that it is about the same for the hexagonal interstitial. The energy of migration does not have a large correction (our calculations give a correction of about 0.15 eV). We find a similarly large correction for the concerted exchange mechanism of 0.75 eV. For the vacancy, we find a correction to the formation energy of 0.38 and a correction to the migration energy of 0.35. The diffusion constant for each of the three mechanisms then becomes

$$D_I = 0.13 \exp^{(-5.05/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \quad (10.13)$$

$$D_V = 2.3 \exp^{(-4.56/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \quad (10.14)$$

$$D_{CE} = 0.54 \exp^{(-5.32/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \quad (10.15)$$

and fractional contributions to total diffusion at 1050° C of

$$f_I = 0.0003, \quad (10.16)$$

$$f_V = 0.981, \quad (10.17)$$

$$f_{CE} = 0.0008. \quad (10.18)$$

Our overall activation energy, which is basically the activation energy of the vacancy mediated mechanism, is still a bit lower than experiment. Another change brought about by the B3LYP correction is that the CE mechanism becomes nearly as important as the interstitial mechanism.

Using the same corrections to the activation energy for Ge diffusion in Si, we now have

$$D_I = 0.02 \exp^{(-5.21/k_B T)} \mathrm{cm}^2\mathrm{s}^{-1} \quad (10.19)$$

$$D_V = 11\exp^{(-4.04/k_BT)}\,\mathrm{cm^2 s^{-1}} \quad (10.20)$$

$$D_{CE} = 2.8\exp^{(-5.29/k_BT)}\,\mathrm{cm^2 s^{-1}} \quad (10.21)$$

and fractional contributions to total diffusion of

$$f_I = 0.00003, \quad (10.22)$$

$$f_V = 0.999, \quad (10.23)$$

$$f_{CE} = 0.001. \quad (10.24)$$

Now, CE contributes more to diffusion than does the interstitial mechanism, but the vacancy mechanism still dominates.

10.7 Conclusion

We have estimated the error in calculating the formation energy of defect structures using PW91 by doing cluster calculations comparing PW91 energies to B3LYP. Our results agree well with the diffusion quantum Monte Carlo calculations of Leung *et al.*, who also found a large correction to E_f but no significant correction to E_m. Their correction to E_f for the interstitial is 1.12 eV and we find 1.01 eV. In addition, we find a correction to the CE migration energy of 0.75 eV, which is very close to their value of 0.98 eV.

The reason PW91 is unable to describe the formation energy of these defect structures, as well as the transition state of the CE mechanism, is unclear. B3LYP includes non-local exchange in the form of Hartree-Fock exchange. PW91 does not have non-local exchange in the same sense. The non-locality of exchange in PW91 is due only to the dependence of exchange on the gradient of the charge density. This only allows exchange between electrons which occupy the same space, while B3LYP allows for exchange between electrons that are "far away" from one another.

The increase in E_f going from PW91 to B3LYP continues the same trend that is seen between LDA and PW91. For relaxed structures, the difference in formation energy of the split interstitial between LDA and PW91 is 0.4 eV. However, if we take the LDA structures and find the GGA energy, the formation energy is 3.95 eV, or 0.65 eV higher than the LDA value and 0.25 eV higher than the relaxed GGA value. So, we can expect that the B3LYP correction to the PW91 value for E_f would decrease, possibly on the order of 0.3 eV, if the structure was allowed to relax.

We have shown that we can use cluster calculations that employ a more accurate description of the exchange-correlation energy to obtain a description of Si self-diffusion that begins to agree quantitatively with experiments.

Chapter 11

NEW DEVELOPMENTS IN THE NUDGED ELASTIC BAND

The work described in this chapter was done in collaboration with Graeme Henkelman. The climbing algorithm for finding the exact saddle point along an NEB described below was Graeme's idea.

11.1 Introduction

Finding saddle points on a potential energy surface is an important problem in solid state physics and physical chemistry. Knowledge of the energy of the transition state gives immediately the likelihood of a process being active at a given temperature and, in the harmonic approximation to transition state theory, the transition state is approximated by the saddle point location. One powerful method for finding saddle points is the nudged elastic band method.

Since its introduction, the nudged elastic band (NEB) method [7] has been applied to a wide range of problems (e.g. [77]). The only inputs to the method are the initial state, the final state, and an initial guess of the minimum energy path (MEP) between the two states. This initial guess for the path is only important insofar that the MEP found will be the one closest to this initial guess. If there is only one MEP connecting the two states, this will be the MEP that is found.

The NEB method has had two shortcomings. First, because of what is termed "kinkiness" in the chain in high force regions, a switching function was needed in order to bring the system to convergence, which could pull the path found by NEB

away from the true MEP. This limitation has recently been lifted by a new definition of the tangent between images [78]. Secondly, the accuracy that the transition state is found depends on the number of images in the chain and the chain spacing. If the spacing was just right for the number of images included, one of the images might land on the transition state and it would be found exactly. However, in general, the image closest to the transition state would lie some distance away, only giving an estimate of the true transition state energy.

The goal of this work is to describe ways in which the NEB method can be modified in order to increase the accuracy with which the transition state is found. Two modifications, both of which change the density of images at the transition state and thus throughout the chain, are described.

In the previous chapters, the NEB method was used, for the most part, without these improvements. As is shown in the examples of Al and Si below, the barrier height estimated via a spline fit to the potential energy curve will generally give the same result as the improvements described. However, if the structure of the transition state needs to be known more accurately, as it was in describing the tetrahedral interstitial and the Si self-interstitial diffusion paths in chapter 4, these improvements were used to get a more accurate geometry. The results of the preceding chapters will not change by using these improvements. However, as will be illustrated, there are systems other than Si for which these improvements are very important.

The importance of these improvements is that one can get more accurate information and extract the true saddle point without doing any more work than without these improvements. There is little to no extra cost in implementing these improvements. The reason that they were not used in the majority of the work described in this thesis is simply because they were made after that work had been done.

11.2 Variable Spring Constants NEB (vNEB)

Generally, when one is searching for reaction paths between two states, the height of the transition state is the crucial piece of information desired. We want more resolution of the minimum energy path near the transition state region than we do in the regions near the initial and final state. If the spring constant is scaled according to the energies of the two images defining the spring, with the spring constant increasing with energy, the density of images at the highest point, the transition state, will increase, increasing the resolution of the barrier region.

We have tried using various types of scaling. The most fruitful has the form:

$$k' = \Delta k \cdot \left(\frac{u_{max} - \overline{u}}{u_{max} - u_{min}} \right) + k_1 \tag{11.1}$$

Here, $\overline{u}$ is the average energy between to adjacent images, u_{min} is the minimum such energy for the entire chain, and u_{max} is the maximum $\overline{u}$. If $n = 1$, then the spring constant k is linearly scaled from k_1 to $k_1 + \Delta k$.

Alternatively, $\overline{u}$ can be set to be the energy of the higher energy image on either side of a chain. With u_{max} set as the highest energy point in the chain and u_{min} as the lowest, the spring constants will be scaled toward the higher energy image defining each spring. This has the advantage that the two images on either side of the saddle point image will be equidistant from the saddle point image, giving at least three images very close to the saddle point. (This is not true if cNEB, described below, is used at the same time.)

In both cases, not only does the increased resolution give a better estimate to the energy and location of the saddle point, but the tangent at the location of the saddle point is better approximated, as the adjacent images are closer and the angle between them will be closer to a straight line. If the separation between the three images closest to the saddle point were infinitesimal, the angle would be exactly 180°, and the calculation of the tangent would be exact. As the images are brought closer

to the saddle point, the approximation of the tangent will become more and more exact.

In the case of very asymmetric barriers, where one of the endpoints is much different in energy than the other, vNEB, as formulated, will result in a very small density of images on one side of the barrier and a very high density on the other. One modification that can be made is to set u_{min} to the higher energy end point. Then, all springs whose energies are below u_{min} will have the minimum spring constant. This will tend to give higher image densities only near the transition state, keeping the spacing of images near both end points the same.

11.3 Climbing Image NEB (cNEB)

The most important factor for determining reaction rates is the saddle point or activation energy. The cNEB method retains the strengths of the NEB method by providing information about the reaction path (ie. the presence of local extrema), but also finds the exact saddle point. Furthermore, this valuable information comes at almost no extra cost.

The cNEB method constitutes a small modification to the NEB method. The idea can be used in conjunction with the vNEB method, or as it has been in this work, with the tNEB method (the NEB method with a stable definition of the tangent along the path and no switching function). At each time step, the image with the highest energy i_{max} is identified. The force on this one image $\mathbf{F}_{i_{max}}$ is modified so that it alone feels only the full force due to the potential, and not the spring forces along the band. Furthermore, the component of this force along the band is inverted

$$\mathbf{F}_{i_{max}} = -\nabla V(\mathbf{R}_{i_{max}}) + 2\nabla V(\mathbf{R}_{i_{max}}) \mid_{\parallel} . \tag{11.2}$$

The parallel force can be written explicitly in terms of the tangent to the path at the highest image

$$\nabla V(\mathbf{R}_{i_{max}}) \mid_{\parallel} = \nabla V(\mathbf{R}_{i_{max}}) \cdot \hat{\tau}_{i_{max}} \hat{\tau}_{i_{max}}. \tag{11.3}$$

Qualitatively, the highest image moves up the potential along the band, and down the potential perpendicular to the band. If the band provides a good enough definition for the reaction coordinate around the transition state, the highest image in the cNEB method will converge to the saddle point. Put in a stronger way, if the cNEB method converges, the highest energy image will be at a saddle point.

As in the vNEB method the images on either side of the transition states will have a different spacing even though the spring constant does not vary along the band. This is because the highest image ignores the spring forces as it moves to the saddle point, compressing the images on one side of the transition states, and expanding the images on the other.

The idea behind the cNEB method can be extended to any or all of the extrema along the reaction path. Typically if there is one maximum along the way, the cNEB method described above does exactly what it desired by finding the saddle point. In cases where there are two maxima as in symmetric processes with a local minima halfway along the path, the cNEB method would be better implemented for both local maxima along the path. Typically though, as few images as possible are used to define the reaction path and moving one of them to the saddle point with the cNEB method is satisfactory.

Both vNEB and cNEB can be used together. This results in an image exactly on the saddle point with several images very nearby. The closer images near the saddle point allow for a more accurate estimate of the tangent at the saddle point, increasing the accuracy of the cNEB algorithm.

11.4 Results

The standard tNEB method has been compared with the two modifications vNEB and cNEB on several test systems. In each case, the total number of force evaluations required for convergence was compared, along with the final estimation of the saddle point energy and the shape of the minimum energy path.

11.4.1 Al/Al(100) system

An EAM type potential [79] was used to describe an aluminum adatom on Al(100) surface. The mechanisms for adatom diffusion, using this potential, have been studied previously [8]. The lowest energy exchange process [80] was chosen to illustrate the different methods being presented.

The regular tNEB method does a good job at mapping out the MEP, and converges in 68 time steps. Figure 11.1 shows the curves calculated by tNEB, cNEB, vNEB and cvNEB. The curves for cNEB and cvNEB show how the climbing method pulls an image exactly onto the saddle point. The number of time steps need to converge was 68, 54, 104 and 85 for tNEB, cNEB, vNEB and cvNEB, respectively.

11.4.2 Si(100)/H_2 system

Density functional theory calculations were done of the dissociation of a H_2 molecule on the Si(100) surface. The DFT code Vasp was used [17]. As in the case of the Al system, we have run the NEB to determine the minimum energy path for various settings of the NEB parameters.

This process illustrates one of the advantages of vNEB. The tail region, where the H_2 molecule approaches the Si(100) surface, is flat and rather uninteresting. vNEB pulls images away from that region, as well as the state where the H_2 is bonded to the surface, towards the saddle point, increasing the resolution of the saddle point area at the expense of areas that are not interesting. The use of vNEB and cNEB

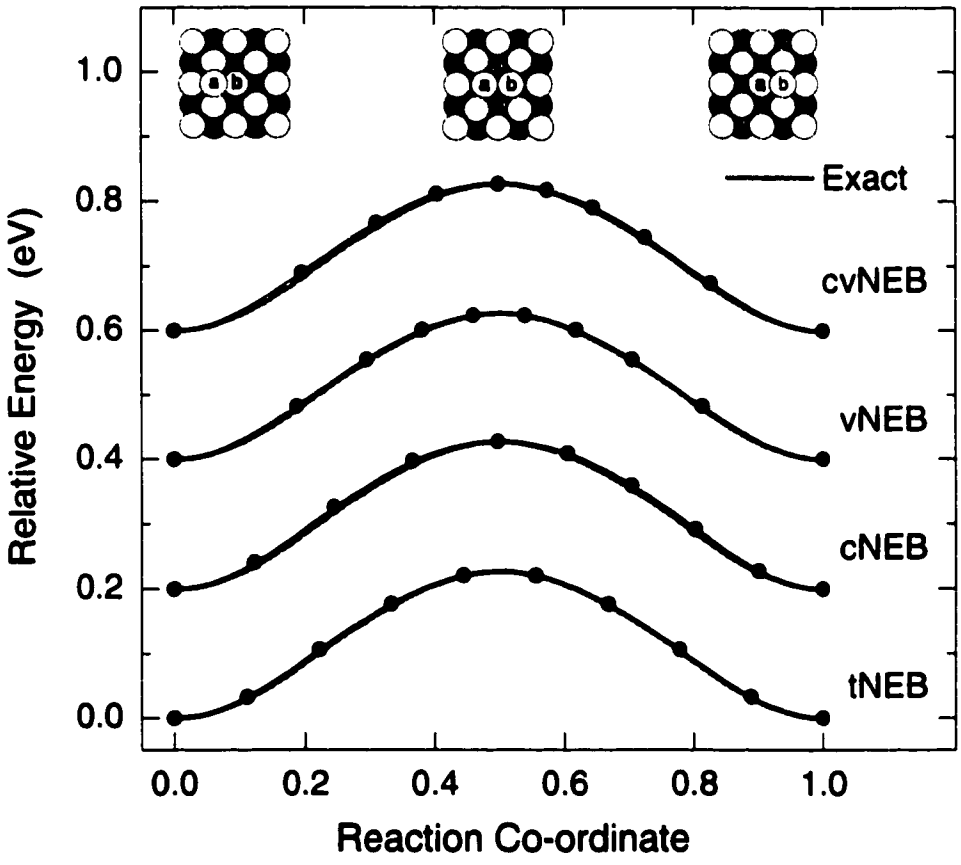


Figure 11.1 : Aluminum exchange barriers.

together using the higher image energy goes even further to concentrate images right around the saddle point.

For all vNEB calculations, u_{min} was set to the higher energy end point, in this case the energy of the free H_2 molecule above the Si(100) surface. In addition, the spring constants were defined not by the average energy of the images defining the springs, but the higher energy image. Two sets of force evaluations are listed for the Si(100)/H_2 system. The first is the number of time steps done with a smaller time step until the system left the high force region of the potential. The second set was done with a larger time step and a more efficient minimization algorithm until convergence was reached, defined here to be 0.05 eV/Å.

The timings for the different methods were 146, 158, 151 and 154 time steps (or, equivalently, ionic force evaluations) for tNEB, cNEB, vNEB and cvNEB, respectively.

The barriers are shown if figure 11.2.

11.4.3 CH_4/Ir(100)

One last system is shown to illustrate how important it can be to use these improvements to find the saddle point energy of a process. The process under consideration here is the adsorption of a CH_4 molecule onto the (100) surface of Ir. Both the regular NEB method and the cNEB method are shown in figure 11.3. As can be seen, the NEB method underestimates the height of the saddle point by nearly 0.1 eV, which is about 50% of the barrier height seen by the CH_4 molecule from the vacuum (that is, from the right of the curve).

11.5 Conclusions

We have made significant improvements to the NEB method that allow a more accurate determination of the saddle point energy of a given process. The climbing image improvement allows the NEB method to find the saddle point exactly, while

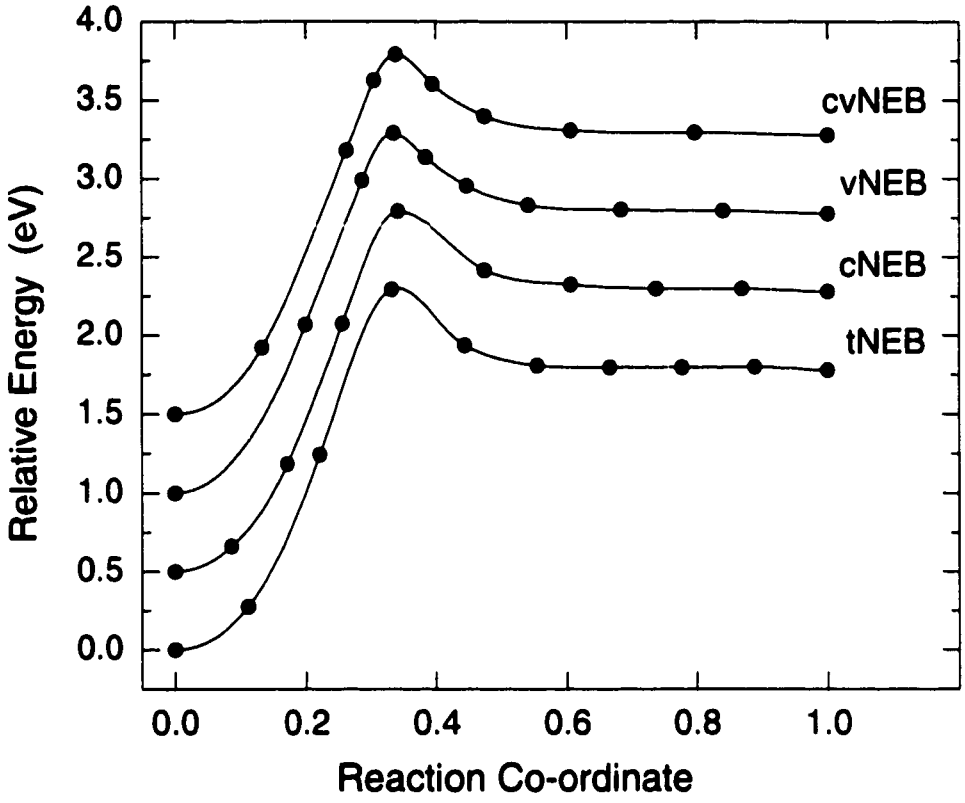


Figure 11.2 : Si(100)/H_2 adsorption barriers.

Table 11.1 : Energy barriers for CE in Si for various dopants and impurities.

Method	Al/Al(100)		H_2/Si(100)	
	Force evaluations	Barrier (eV)	Force evaluations	Barrier (eV)
tNEB	68	0.22704	13+133=146	2.30297
vNEB	104	0.22747	29+122=151	2.29154
cNEB	54	0.22717	27+131=158	2.29211
cvNEB	85	0.22690	27+127=154	2.29114

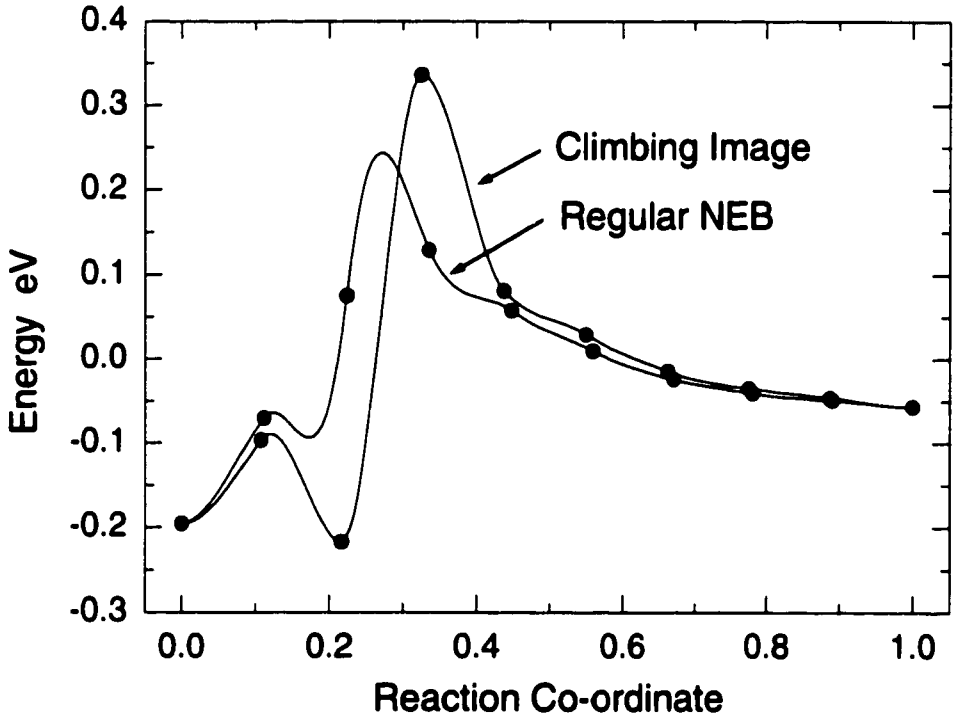


Figure 11.3 : Ir(100)/CH_4 adsorption barriers, illustrating the need for cNEB to accurately find the saddle point.

the variable spring constant improvement gives a much higher resolution of the path near the saddle point. Combined into the cvNEB method, they give a much more accurate description of the saddle point with out a significant increase in the number of force evaluations necessary. Thus, there is no reason not to adopt these changes whenever the NEB method is used.

Chapter 12

ELASTIC SHEET METHOD

The results presented here are published in the *Journal of Chemical Physics*, **111**, 10664 (1999). This work was done in collaboration with Enrique Batista.

12.1 Introduction

Studies of molecules and condensed phases often lead to discussions of charges and multipole moments of individual fragments such as atoms or molecules. Given the continuous electronic density of the system, the question becomes how to identify an atom in a molecule, or a molecule in a cluster or a liquid configuration, for example. Many different partitioning schemes have been proposed. When a calculation of the electronic wavefunction of a system is carried out in terms of an atomic basis functions, it is tempting to assign the electronic density associated with a given basis function to the atom at that site [81]. But, it is important to realize that atomic basis sets are overcomplete and such a decomposition is not unique. In principle, a calculation could be done where all the basis functions are located on one of the atoms in the system which would then lead to an assignment of all the electrons in the system to that one atom.

One compelling way of approaching this problem in a less arbitrary fashion is the decomposition of the charge density proposed by Bader [82]. Here each point in space is assigned to one of the subsystems (e.g. atoms). The dividing surface is chosen to be a zero-flux surface as defined by

$$\nabla\rho \cdot \mathbf{n} = 0 \tag{12.1}$$

where **n** is the surface normal. That is, at every point on a zero-flux surface the gradient of the charge density has no component normal to the surface. Bader has given theoretical arguments as to why this is a good choice for a dividing surface. By using the zero-flux surface, various surface integral terms go to zero when quantum mechanical expectation values for the subsystem are calculated. It can be shown, in particular, that each subsystem defined in this way satisfies a virial theorem [82].

Finding the zero-flux surfaces of the charge density, however, is not trivial. Methods currently employed can, in fact, fail for certain charge density topologies. The method of Stefanov and Cioslowski [83], used, for example, in the Gaussian code, is one example. This method involves fitting the surface with variational trial functions in prolate spheroidal coordinates. It can fail when the zero-flux surface has certain topological features, such as very strong curvature [84]. During a study of molecular multipole moments of water clusters, we found that the method failed on the hexamer. For this reason, we decided to develop an alternative method for finding zero flux surfaces.

12.2 Overview of the Elastic Sheet Method

In finding a zero-flux surface, we want to minimize the gradient of some scalar field - in this case, the charge density - normal to a closed surface. Our method involves defining a set of fictitious particles which essentially give a discrete representation of the surface. Initially, the particles are distributed randomly on some closed surface, such as a sphere, but are then relaxed according to the force acting on them and in the end they are located on the zero-flux surface. The force on these particles has two components. The first is the "real" force, the gradient of the charge density. The second component is an interaction between the particles which keeps the particles distributed evenly on the surface of the sheet. This force is referred to as the "distributing" force. The component of the real force tangent to the sheet surface is

zeroed, as is the component of the distributing force normal to the sheet. This force projection ensures that the distributing force does not interfere with the relaxation of the particles to the zero-flux surface and guarantees that the real force does not affect the distribution of particles within the surface. As a result, the final shape of the sheet will be determined solely by the charge density, while the distributing force will insure that the density of particles on the surface remains more or less uniform. Figure 12.1 illustrates the projection of the forces.

In regions where the gradient of the electronic density is small, one can encounter a problem which we refer to as "kinkiness": the sheet will deform slightly and particles "evaporate" away from the volume defined by the sheet. To counteract this, an additional restoring force is introduced, the nature of which will be discussed below.

The elastic sheet (ES) method can be viewed as an extension of the nudged elastic band (NEB) method for finding minimum energy paths [7]. In the NEB method, a minimum energy path is represented by a discrete set of fictitious particles whose position is optimized by minimizing the perpendicular component of the gradient of the potential under consideration and the parallel component of a spring force between the particles. The spring force causes the particle to be equidistant along the path (when equal spring constants are used). Because only the parallel component of the spring force is kept, the particles relax to the minimum energy path; the spring force only affects the distribution of particles along the path.

12.3 Details of the Elastic Sheet Method

The evolution of the ES is governed by two forces: the real force and the distributing force. The real force is just the gradient of the logarithm of the charge density:

$$\mathbf{f}_{\parallel}^{real} = -\frac{1}{\rho}\nabla\rho \cdot \mathbf{nn} \tag{12.2}$$

The charge density decays exponentially and the real force would as well. To accelerate convergence in regions where the density is changing slowly, we work with the

logarithm of ρ. For finite systems, such as clusters, where the charge density decays to zero, the real force is set to zero at some predefined density contour, which then defines a practical approximation to the zero-flux surface in that direction.

The distributing force acts between the particles that make up the sheet and its purpose is to keep the density of particles on the surface uniform. After testing various types of interactions, we have chosen to use a generalized Lennard-Jones interaction, where the potential energy between particles i and j is given by

$$V_{ij}^{dist} = \begin{cases} 4\epsilon\left(\frac{\sigma^m}{r_{ij}^m} - \frac{\sigma^n}{r_{ij}^n}\right) - 4r_{ij}\epsilon\left(m\frac{\sigma^m}{r_{cut}^{m+1}} - n\frac{\sigma^n}{r_{cut}^{n+1}}\right) + K & r_{ij} < r_{cut} \\ 0 & r_{ij} > r_{cut} \end{cases} \tag{12.3}$$

Here, ϵ is a parameter defining the strength of the interaction, r_{ij} is the distance between i and j, σ is a parameter that characterizes the current spatial distribution of particles, r_{cut} is a cutoff radius, which we define to be some multiple of σ, and K is a constant that makes V go smoothly to zero. m and n are parameters that define the shape of the interaction. We use $\epsilon = 0.5$, $m = 7$ and $n = 6$.

This results in the following force between i and j:

$$\mathbf{f}_{ij}^{dist} = \left[4\epsilon\left(m\frac{\sigma^m}{r_{ij}^{m+1}} - n\frac{\sigma^n}{r_{ij}^{n+1}}\right) - 4\epsilon\left(m\frac{\sigma^m}{r_{cut}^{m+1}} - n\frac{\sigma^n}{r_{cut}^{n+1}}\right)\right]\hat{\mathbf{r}}_{ij} \tag{12.4}$$

which, because of the second term, goes smoothly to zero at r_{cut}. Here, $\mathbf{f}_{ij}^{dist}$ is the distributing force on i due to neighbor j and $\hat{\mathbf{r}}_{ij}$ is the unit vector connecting the two particles. $\mathbf{f}_{ij}^{dist} = 0$ if $r_{ij} > r_{cut}$.

Because the shape of the elastic sheet changes with time, expanding and contracting, the average distance between particles will also change. To keep the magnitude of the distributing force comparable to that of the real force, as well as to keep the distributing force from either diverging or becoming negligibly small, σ is tuned to the current distribution of the particles on the sheet. In our case, we define σ to be the average nearest neighbor distance between the closest 6 neighbors of each particle divided by a parameter α which determines on which side of the potential well minimum the nearest neighbors lie. We use $\alpha = 1$ so the nearest neighbor particles lie just

on the repulsive side of the well. We have found that it is necessary to update σ every time step. Otherwise, the change in the potential between i and j is too sudden, and there can be problems with stability.

As stated above, the various forces need to be projected onto the surface normal. It is very important to have a good estimate of the normal at each particle at each step during the optimization. We have found that a good estimation of the surface normal is, for example, important for keeping the particles from "evaporating" from the elastic sheet. We calculate the normal by first finding three neighbors, j, k, and l, that satisfy the following criteria: Neighbors j and k are chosen from the 6 closest neighbors, denoted by $\{n\}$, so that the angle between r_{ij} and r_{ik} is closest to $2\pi/3$. l is the neighbor remaining among $\{n\}$ such that the angle between r_{il} and both r_{ij} and r_{ik} is closest to $2\pi/3$. More precisely: Defining

$$f_{jk} = \left| \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{ik}}{|\mathbf{r}_{ij}|\,|\mathbf{r}_{ik}|} + \frac{1}{2} \right| \tag{12.5}$$

the particles j, k, and l are chosen from $\{n\}$ such that

$$\begin{aligned} &j \text{ and } k && \text{minimize } f_{jk} \\ &l && \text{minimizes } f_{jl} \text{ and } f_{kl} \end{aligned} \tag{12.6}$$

The normal is then defined as

$$\mathbf{n} = \frac{\mathbf{r}_{jk} \times \mathbf{r}_{jl}}{|\mathbf{r}_{jk} \times \mathbf{r}_{jl}|} \tag{12.7}$$

This normal is then used for the force projections at each particle.

At times, one finds particles that drift away slightly from the rest of the sheet. This then leads to inaccuracies in calculating the normals and these inaccuracies can propagate to nearby particles. To fix this problem, "smoothing" is introduced. Smoothing involves adding a force along the direction of the normal to pull the drifting particle back towards the surface of the sheet. The smoothing force is a function of the projected distance of a particle along the normal from the six nearest neighbors, $\{n\}$:

$$z = \frac{1}{6} \sum_{\{n\}} (\mathbf{r}_j - \mathbf{r}_i) \cdot \mathbf{n} \tag{12.8}$$

This definition assumes a convention where all normal vectors point away from the inside. This is enforced by keeping track of the direction of the normal at each iteration, all the way from the initial sphere. This works since the direction of the normal cannot flip in one iteration (unless the iteration step size is much too large).

The smoothing adjusts the force in the normal direction so as to pull particles along their normal towards the average plane defined by $\{n\}$.

The actual form of the smoothing function is:

$$s = \begin{cases} -\frac{1}{2}\left(1 - \cos(\frac{\pi|z|}{\beta\sigma}\right) & \text{if } 0 < |z| < \beta\sigma \\ -1 & \text{if } |z| > \beta\sigma \end{cases} \tag{12.9}$$

s is negative if $0 < z < \beta\sigma$. Here, β is a parameter that determines the strength of the smoothing function. The smaller the choice for β, the faster the smoothing force is turned on. We have been using values between 0.5 and 1.0 for β. σ is the same parameter in the definition of the distributing force.

The form of the smoothing force is

$$\mathbf{f}_i^{switch} = s \times k \times z\mathbf{n} \tag{12.10}$$

where k is a parameter that determines the strength of the smoothing force. We have used a value between 50 and 100 Å^{-2}.

The final force on particle i is then, after zeroing the appropriate components and adding the smoothing force

$$\mathbf{f} = \mathbf{f}_{\|i}^{real} + \mathbf{f}_{\perp i}^{dist} + \mathbf{f}_i^{switch} \tag{12.11}$$

where

$$\mathbf{f}_{\|}^{real} = -\frac{1}{\rho}\nabla\rho \cdot \mathbf{nn} \tag{12.12}$$

$$\mathbf{f}_{\|}^{dist} = -\nabla V^{dist} \cdot \mathbf{nn} \tag{12.13}$$

$$\mathbf{f}_{\perp}^{dist} = -\nabla V^{dist} - \mathbf{f}_{\|}^{dist} \tag{12.14}$$

$$\mathbf{f}^{switch} = s \times k \times z\mathbf{n} \tag{12.15}$$

The sheet is minimized with these forces using a minimization method based on the velocity Verlet algorithm [85] where the component of the velocity perpendicular to the force is zeroed at each iteration and the entire velocity vector is zeroed if $\mathbf{v} \cdot \mathbf{f} < 0$. When the forces on the sheet are minimized, the shape of the sheet then defines the zero-flux surface.

12.4 Integrating the Subspace defined by the Zero-Flux Surface

Once the zero-flux surface is known, one would like to calculate various properties of the subspace so defined. Most importantly, one would like to know the total charge contained within the surface. Another property that may be of interest is dipole moment, or higher multipole moments. It is, therefore, important to be able to tell which points on the charge density grid lie within the region enclosed by the dividing surface.

This is accomplished by starting with some reference point, $\mathbf{R_0}$, that is known to lie within the region. This can, for example, be the location of the atom around which the elastic sheet has expanded. At each charge density point, $\mathbf{R}_i$, a line is then drawn from the reference point through the charge density point,

$$\hat{\mathbf{r}} = \frac{\mathbf{R}_i - \mathbf{R_0}}{|\mathbf{R}_i - \mathbf{R_0}|}. \tag{12.16}$$

Of the particles defining the elastic sheet, the M particles closest to the line $\hat{\mathbf{r}}$ are found. We are currently using $M = 20$. The location of these particles is projected on to a plane perpendicular to $\hat{\mathbf{r}}$. Then, using a method by D. F. Watson [86], we triangulate in this plane the particles M to obtain the connectivity among particles. This connectivity is then used to reconstruct the surface of the ES locally as a collection of triangles in three-dimensional space. For each triangle, the equation

$$\alpha_1 \mathbf{V}_1 + \alpha_2 \mathbf{V}_2 + \alpha_3 \mathbf{V}_3 = \mathbf{R_0} + \lambda \hat{\mathbf{r}}. \tag{12.17}$$

is solved. Here, the $\mathbf{V}_i$ are the coordinates of the vertices of the triangle being

examined, λ measures the distance the triangle is from $\mathbf{R}_0$ along $\hat{\mathbf{r}}$, and α_i are the barycentric coordinates of the triangle. For a point to lie within the triangle, $\sum_i \alpha_i = 1$, so the constraint $\alpha_1 = 1 - \alpha_2 - \alpha_3$ is enforced. This gives a set of three linear equations with three unknowns. Solving this system for α_i, the line $\mathbf{R}_0 + \lambda\hat{\mathbf{r}}$ will cross the triangle defined by $\mathbf{V}_i$ if and only if

$$0 \leq \alpha_i \leq 1, \forall i. \tag{12.18}$$

By solving this system of equations for each triangle in the triangulation, it can be determined which triangle the line crosses and at which point in space the line and the ES intersect:

$$\mathbf{R}'_i = \alpha_1\mathbf{V}_1 + \alpha_2\mathbf{V}_2 + \alpha_3\mathbf{V}_3 \tag{12.19}$$

The charge density point $\mathbf{R}_i$ is then inside the sheet if

$$|\mathbf{R}_i - \mathbf{R}_0| < |\mathbf{R}' - \mathbf{R}_0|. \tag{12.20}$$

For each point on the density grid, this method can be used to determine whether it is located in the region enclosed by the dividing surface.

This method will not work for complex surfaces where the line $\hat{\mathbf{r}}$ connecting a charge density point to the reference point can cross the elastic sheet more than once. More elaborate methods need to be used in such cases.

12.5 Results

We have applied the above algorithm to partitioning of the valence electron densities obtained in DFT/PW91 pseudopotential calculations of Si crystal, bulk ice, and water clusters containing from 2 to 6 water molecules. The Si structures studied include the bulk bond (Figure 12.2), the bond between a pair of atoms which have been rotated in the Si crystal (Figure 12.3), and the bond between the two atoms forming a dumbbell interstitial (Figure 12.4). This last structure is especially complex, being composed

of a total of 5 local maxima, leading to 5 different zero-flux surfaces to describe the valence charge density in the bonding region.

The bond in a perfect Si crystal (Figure 12.2) should have an integrated charge density of 2 electrons From the region enclosed by the converged elastic sheet using 2000 particles which is shown in figure 12.2 the integrated charge density amounts to 1.977 electrons, within 1.1% of the expected value of 2.

Figure 12.3 shows the zero-flux surface of the bond between two atoms that are rotated in the bulk. This configuration is metastable and is found along the minimum energy path of the concerted exchange process proposed by Pandey [21]. The integrated charge inside this surface is 2.097 electrons, showing that this bond has a slightly enhanced electronic density as compared with the perfect crystal.

A much more complicated topology of the zero flux surface is associated with the split interstitial (also known as the dumbbell interstitial). The valence electron charge density for this bond is composed of five local maxima: a central maximum and four symmetric satellite maxima. Figure 12.4 shows the composition of these five surfaces. The total integrated charge density of the combined surfaces is 1.42 electrons, with 0.86 electrons in the central maximum and 0.14 electrons in each of the satellite maxima. These zero-flux surfaces are especially complex, with the central maximum exhibiting a very sharp cusp in the [100] direction and each of the satellite maxima containing two sharp points. As long as the smoothing force is included, the elastic sheet method describes these surfaces very well.

In a study of molecular multipole moments in water clusters, we used the elastic sheet method to identify Bader 'molecules' in the clusters. The partitioning of the hexamer cluster is shown in Fig. 12.5. This figure illustrates how well the calculated subregions fill space. The six sheets were calculated separately. The final partitioning is cut so the inside is visible. There is some space in the very center of the partitioning that is not accounted for by the method (the size of the particles hides that fact). However, most of the space of the system is accounted for by the subspaces. The total

integrated valence density of the six H_2O molecules inside the calculated surfaces is 47.96 electrons, so the error is only 0.04 electrons out of 48.

The molecular multipole moments of the various H_2O molecules were calculated. Due to the effect of the electric field from neighboring molecules, the dipole moment increases from 1.86 D [87] in the gas phase to 2.47 D in the hexamer and 2.74 D in ice [88]. When the electric field was evaluated at a typical intermolecular distance from the cluster molecules, the multipole expansion using multipoles obtained from the Bader partitioning converged to the field obtained from the full electron density at the hexadecapole. We also compared calculations using the elastic sheet method with calculations based the method of Stefanov and Cioslowski [83] using the Gaussian94 code. While the calculation of the zero-flux surfaces for the hexamer did not converge, some of the other water clusters did converge and the calculated molecular dipole moments then agreed to within 1% with our results from the elastic sheet method.

12.6 Discussion

During the development of the elastic sheet method, we have encountered several problems that had to be over come. The biggest of these was the form of the distributing force. As was mentioned before, we tune σ such that the nearest neighbors of a particle always lie on the repulsive side of the potential well. This repulsive pressure can in some cases force particles out of the sheet at sharp features such as points and cusps and into the neighboring zero-flux surfaces, forming wing-like structures. One might think that a simple solution to this would be to choose an interaction resulting in attraction between nearest neighbors. When we tried this, however, holes formed in the sheet and the particles tended to clump together.

The addition of a 'smoothing' force solved this problem. By adding a restoring force that tends to make the sheet locally flat, the particles are not allowed to escape from the dividing surface. Not only does this lead to smoother shape, but it also

helps the long term convergence. Without the smoothing force, the best estimate for the integrated charge of the bond in the Si crystal was 1.88 electrons. By adding the smoothing force, the integrated charge is 1.97 electrons, very close to the exact value of 2.00.

When the calculation is started, the particles are placed at random on a sphere centered on the atom of interest. The random placement of the particles will often cause very large forces between them, so the interaction force is scaled down in the first few steps of the minimization, until the particles have reached a reasonable distribution within the sphere.

12.7 Conclusions

We have developed an elastic sheet method for finding the zero-flux dividing surfaces of the charge density. The method has been applied to a study of the electron density around interstitials in Si and in analysis of multipole moments of water molecules in water clusters. The discrete representation of the zero-flux surfaces obtained with the elastic sheet method can be used to calculate the integrated charge enclosed by the surface. The method should have more general applicability. The elastic sheet algorithm can be applied to any system where zero-flux surfaces of some scalar field are needed. It is possible to extend the method to higher dimensional systems, where it might, for example, be used to find the transition state dividing surface of a potential energy surface or even a free energy surface.

The computer program for carrying elastic sheet calculations is available on request. The input for the calculation is simply the charge density evaluated on a uniform grid in three-dimensional space. An interpolation formula is then used to evaluate the charge density at any point in space, as well as the gradient of the charge density.

12.8 Acknowledgments

This work was supported by NSF grant number CHE-9710995. We gratefully acknowledge helpful discussions with Graeme Henkelman.

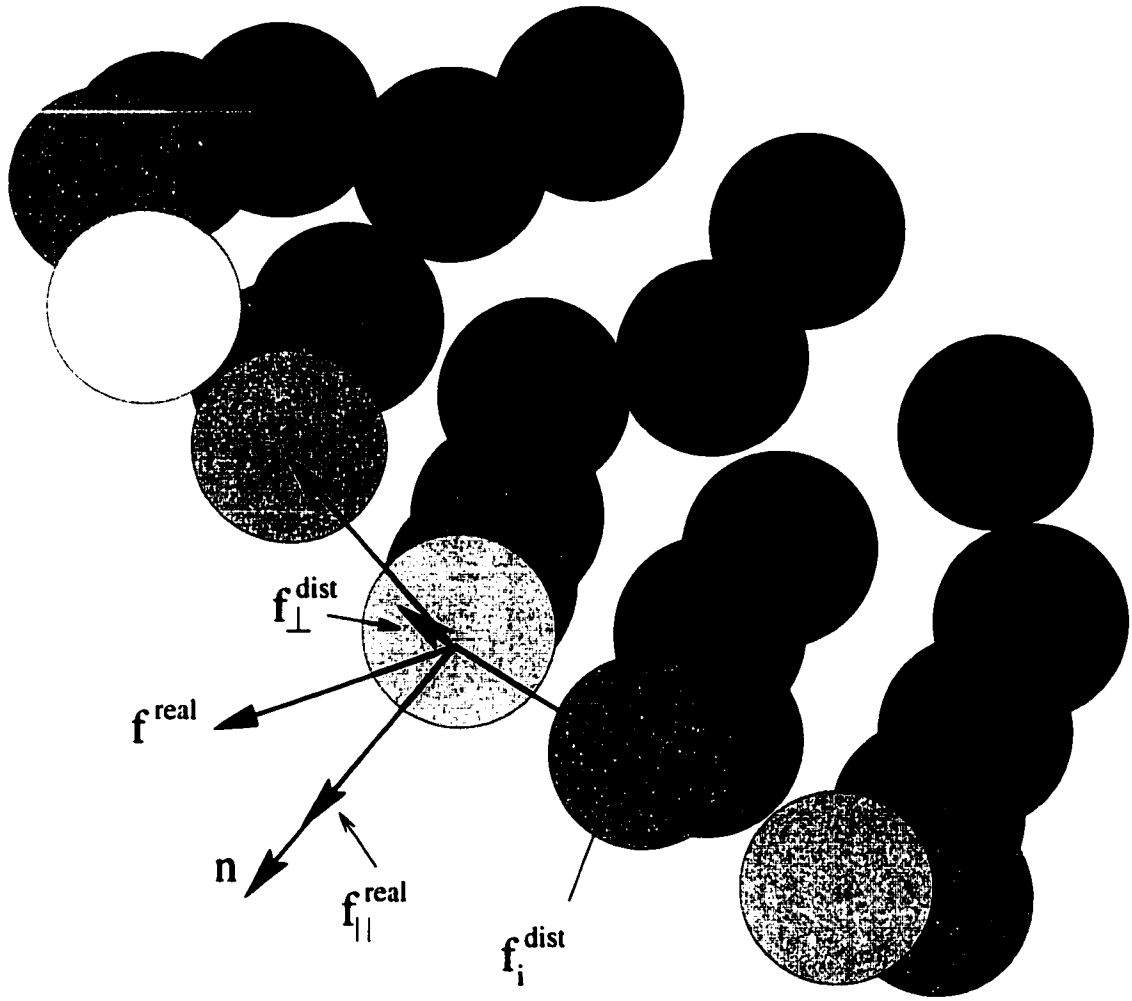


Figure 12.1 : Forces acting on a particle in the elastic sheet. The particles move in response to the normal component of the real force (gradient of the electronic charge density) and the component of the distributing forces in the local tangent plane. The first acts to move the particles to the zero-flux surface while the second acts to keep the particle distribution nearly uniform.

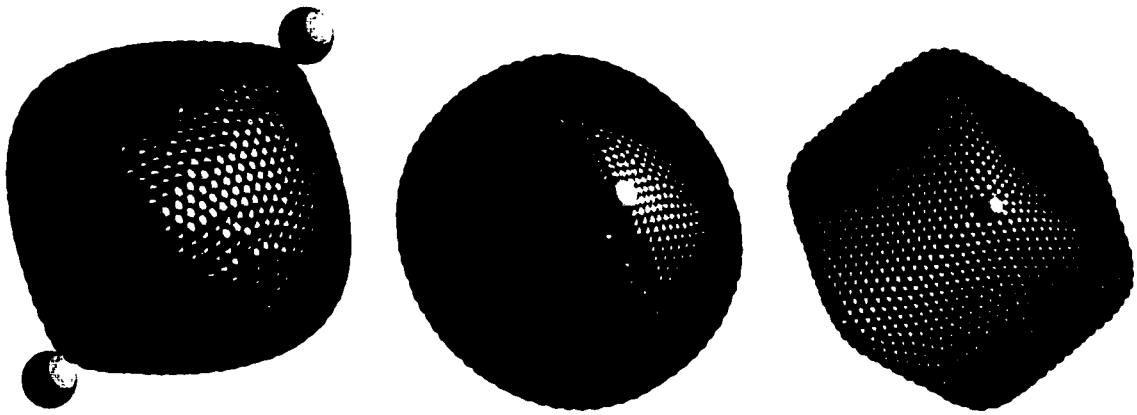


Figure 12.2 : Evolution of the elastic sheet around the valence electron density of a bond in a Si crystal. Starting with a perfect sphere, the first snapshot shown (left) is taken after 200 iterations, the second after 1000 iterations, and third (right) after convergence to the zero-flux surface after 10,000 iterations. The larger spheres indicate the position of the Si atoms. The integrated charge of the enclosed volume is 1.976 electrons. The calculation took 34 minutes on a Pentium 400.

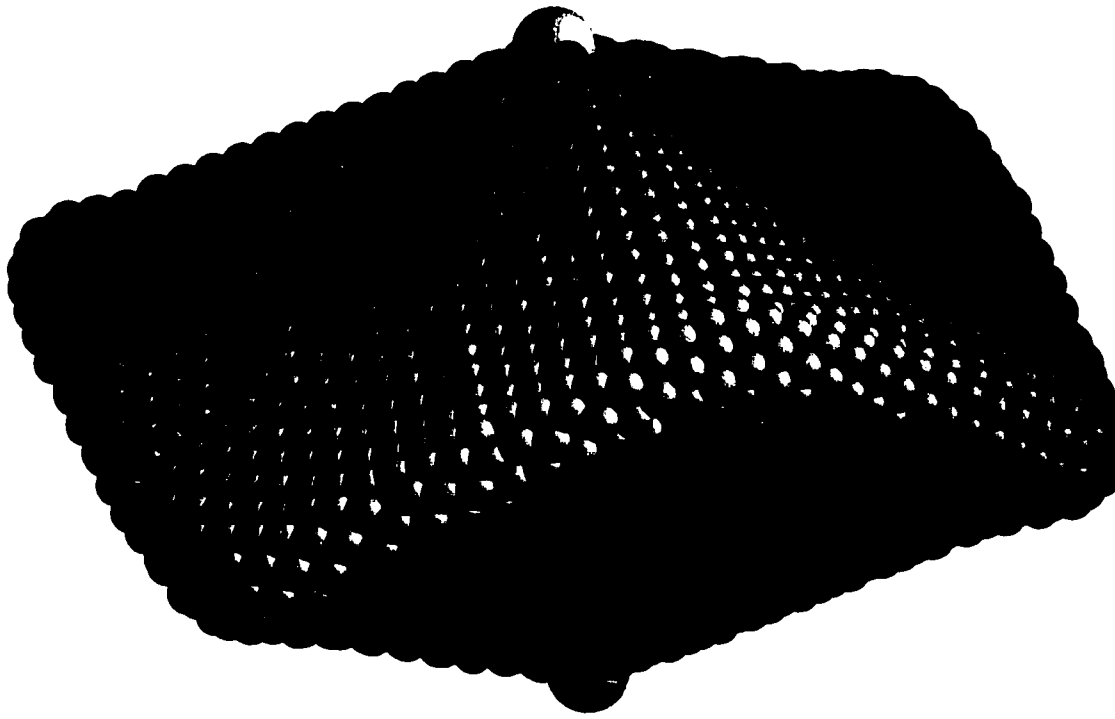


Figure 12.3 : Zero-flux surface for the bond between two atoms in the metastable state found along the minimum energy path of the concerted exchange Si diffusion process proposed by Pandey. The shape is very similar for the bond in the perfect crystal. The total integrated charge, is larger for this bond, with the surface enclosing 2.097 electrons. The larger spheres show the location of the Si atoms.

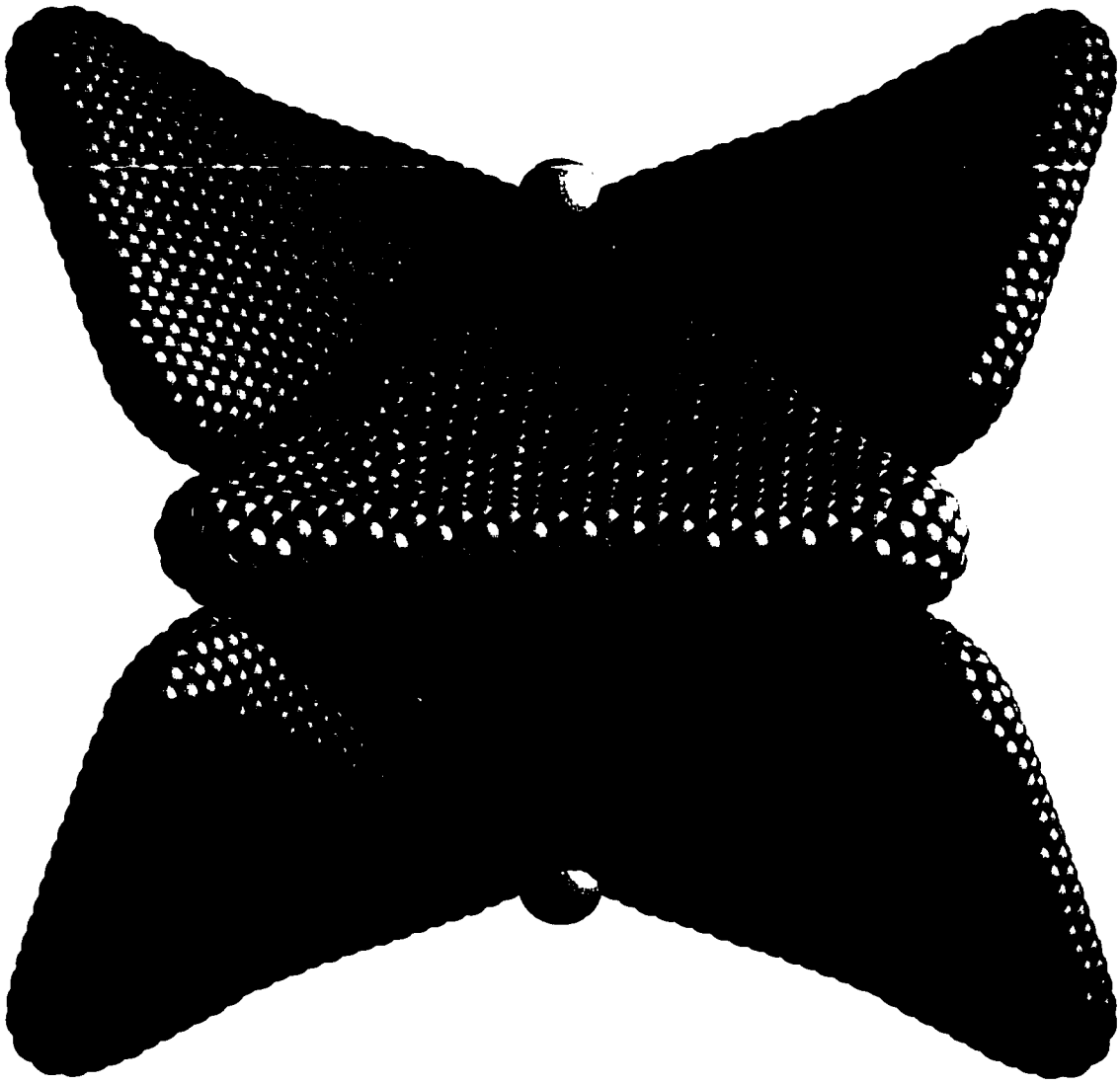


Figure 12.4 : Zero-flux surface for the bond between the two Si atoms forming a split interstitial configuration in a Si crystal. The valence charge density is decomposed into regions by zero-flux surfaces. The figure shows all five regions from the [100] Si crystal direction. The integrated charge of the central region is 0.86 electrons, while each of the satellite regions contains 0.14 electrons, giving a total of 1.42 electrons in the bond. The larger spheres indicate the location of the two Si atoms.

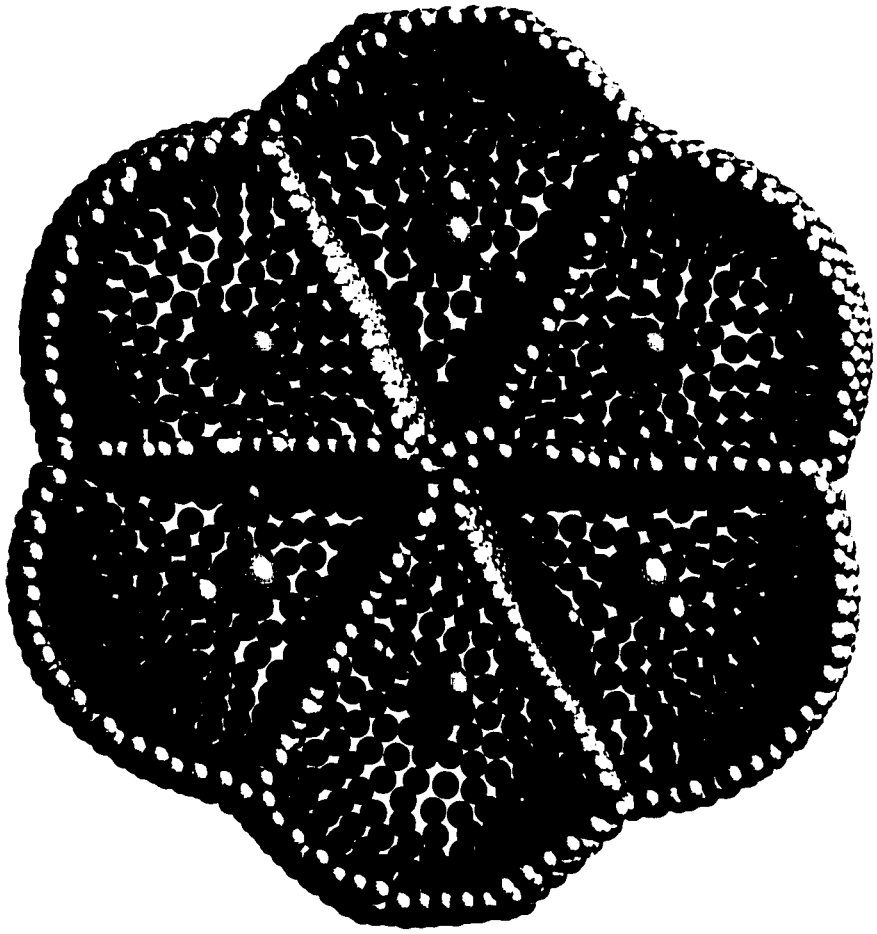


Figure 12.5 : Zero-flux surfaces for the six molecules in the water hexamer calculated from the valence charge density. Each of the surfaces was calculated separately. Of the total 48 valence electrons in the cluster, 47.96 are accounted for by the six subregions enclosed by the elastic sheets. The decomposition of the cluster charge density enabled calculation of the molecular multipole moments. The molecular dipole moment was found to be 2.47 D, up by 33% from the gas phase value. The outer boundary was chosen to be the $\rho = 0.001 \text{electrons/Å}^3$ contour. The water molecules are also shown. Note the bending of the zero-flux surfaces near the hydrogen atoms.

Chapter 13

CONCLUSION

We have studied diffusion in Si and Ge, in particular Si self-diffusion, Ge diffusion in Si, and Ge self-diffusion. We have been able to calculate the four quantities that determine the diffusion constant: the energy of formation and migration and the entropy of formation and migration. While we do not have quantitative agreement with experiment, we do have a basic description of diffusion in Si and Ge that agrees with experiments.

The DFT/PW91 energy of activation we find is consistently low compared to experiment in all of our calculations by about 1 eV. This is true for both interstitial and vacancy activation energies, and for all systems considered, not only Si and Ge self-diffusion but dopant diffusion in Si as well. This strongly suggests that the discrepancy is due to some error made in all of the calculations. The common component for all of these calculations is the formation energy of the defect. That is, all mechanisms involve the formation of a Si interstitial or vacancy. They differ in the binding and migration of these defects near the foreign atom. This means that the migration energies are quantitatively much more accurate than the formation energies. This is born out by the fact that we find a small B3LYP correction to the migration energy of the interstitial, but a large correction to the formation energy. This is also seen by the DMC study [38]. Diffusion equation simulations of B diffusion in Si using the migration energies we calculate agree well with experiment [5]. These same simulations also use our entropies of migration.

Smargiassi and Car [37] have shown that at temperatures of 500 K, the harmonic approximation gives very accurate entropy of formation for the Si vacancy. However,

they also show that at higher temperature (1000 K), anharmonic contributions to the entropy of formation are $2 \pm 1 k_B$. This is one source of the errors in our prefactors. In addition, Clark and Ackland [29] saw correlated hops in finite temperature simulations of the Si interstitial. Such correlated hops are not accounted for in transition state theory and dynamical corrections to TST need to be considered to obtain more quantitative agreement with experiment. Finally, there should also be anharmonic contributions to the entropy of migration. As mentioned above, the harmonic prefactors calculated here reproduce experimental profiles of B diffusion. However, it is possible that a common correction to the prefactors would give better quantitative agreement with experiment while not changing the results of those simulations.

In addition to errors in the entropy introduced by our use of the harmonic approximation, similar errors could exist in the energy of formation. We have assumed that the energy of the state is the minimum energy structure energy, which may not be true at finite temperature. All of these contributions need to be calculated in order to obtain better agreement with experiment.

The good agreement we obtain with the experimentally determined prefactor for Ge self-diffusion indicates that the harmonic approximation is good enough at 500° C, the temperature at which the Ge experiments were conducted. This conclusion is consistent with the simulations of Smargiassi and Car [37] who found that anharmonic contributions to the entropy of formation of the Si vacancy at 500 K were negligible. The discrepancy between our calculations and the Si self-diffusion experiments is most likely because those experiments are carried out at a high temperature, 1000° C. This is because of the higher activation energy of Si self-diffusion. Viewed in light of the calculations of Smargiassi and Car, the reason we agree with experiments for the prefactor of Ge self-diffusion but not for Si self-diffusion is because anharmonic contributions to the prefactor are not active at the low temperatures of Ge self-diffusion experiments but are active for the Si self-diffusion experiments. Since we neglect these contributions, we agree with the Ge self-diffusion data, but not the Si

self-diffusion data.

We have shown that we can use cluster calculations that employ a more accurate description of the exchange-correlation energy to obtain a description of Si self-diffusion that begins to agree quantitatively with experiments. Our prefactors are low in every case except for Ge self-diffusion. To address this discrepancy, calculations taking into account the anharmonicities of the interactions amongst atoms need to be done. However, though we do not agree quantitatively in every aspect with experiment, we do agree qualitatively. We find that Ge diffusion in Si is dominated by vacancies more than Si self-diffusion, as one would expect for the larger Ge atom in the confining spaces of the Si lattice. In addition, our PW91 prefactors and activation energies agree qualitatively with Ural and not with Bracht, suggesting that the analysis of Ural is more accurate. This is the true strength of these kinds of studies: we can help interpret experiments that are difficult to analyze.

Of course, it would be preferred to agree quantitatively as well as qualitatively. There are definite studies that need to be conducted that would result in more accurate values for both the prefactor and the activation energy. We have shown how the value of the activation energy can be more accurately determined. Free energy calculations of atomic motion pathways would include both anharmonic contributions to the entropies as well as revealing any temperature dependence in the free energy barrier. These studies need to be done in the future.

Therefore, the next step is to do the theoretical calculations necessary to get better quantitative agreement with experiment. The B3LYP correction to the vacancy migration energy needs to be tested with larger clusters. Also, anharmonic contributions to the entropy of formation and migration need to be found. These two studies would greatly clarify the source of the quantitative disagreement between theory and experiment.

There are also many systems that can be studied. We have begun to do some of them here. To complete the picture of As diffusion in Si, the interstitial mechanism

needs to be analyzed. Prefactors need to be calculated for both As and Sb diffusion. Charge states need to be considered. There is still a wealth of information that needs to be gathered to get a complete theoretical picture of diffusion in Si.

Experimentally, experiments that can some how determine the contribution of concerted exchange to self-diffusion in Si as well as Ge diffusion in Si need to be done. Though we find in all cases that CE contributes very little to diffusion, after applying the B3LYP corrections, we find that in some cases, it contributes as much as or more than interstitial mechanisms. Experiments have found that interstitial diffusion in some of these systems might be as high 30 to 40%. If our results are qualitatively correct in that the CE contribution to diffusion is the same as interstitials, then it may be an important mechanism for diffusion.

BIBLIOGRAPHY

[1] P. C. Kelires and J. Tersoff, Physical Review Letters **63**, 1164 (1989).

[2] J. Tersoff, Physical Review Letters **74**, 5080 (1995).

[3] J. D. Plummer, M. D. Deal, and P. B. Griffin, in *Silicon VLSI Technology : Fundamentals, Practice, and Modeling* (Prentice Hall, Englewood Cliffs, New Jersey, 2000).

[4] J. P. Perdew, in *Electronic Structure of Solids*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.

[5] W. Windl *et al.*, in preparation (unpublished).

[6] E. Wigner, Trans. Faraday Soc. **34**, 29 (1938).

[7] G. Mills, H. Jónsson and G. K. Schenter, Surf. Sci. **324**, 305 (1995); G. Mills, H. Jónsson and K. W. Jacobsen, Ch. 16 in B. J. Berne, G. Ciccotti and D. F. Coker, eds., *Classical and Quantum Dynamics in Condensed Phase Simulations* (World Scientific, 1998).

[8] G. Henkelman and H. Jónsson, Journal of Chemical Physics **111**, 7010 (1999).

[9] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).

[10] D. Chandler, in *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987), p. 70.

[11] R. LeSar, R. Najafabadi, and D. J. Srolovitz, Physical Review Letters **63**, 624 (1989).

[12] M. C. Payne *et al.*, Reviews of Modern Physics **64**, 1045 (1992).

[13] R. O. Jones and O. Gunnarsson, Reviews of Modern Physics **61**, 689 (1989).

[14] J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77**, 3865 (1996).

[15] A. D. Becke, Phys. Rev. A **38**, 3098 (1988); C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988); S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[16] J. P. Perdew, M. Ernzerhof, and K. Burke, Journal of Chemical Physics **105**, 9982 (1996).

[17] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); **49**, 14251 (1994); G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 16 (1996); Phys. Rev. B **55**, 11169 (1996).

[18] D. Vanderbilt, Physical Review B **41**, 7892 (1990).

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

[20] W. Frank, U. Gösele, H. Mehrer, and A. Seeger, in *Diffusion in Crystalline Solids*, edited by G. E. Murch and A. S. Nowick (Academic Press, Orlando, 1984), p. 63.

[21] K. C. Pandey, Phys. Rev. Lett. **57**, 2287 (1986).

[22] R. J. Borg and G. J. Dienes, in *An Introduction to Solid State Diffusion* (Academic Press, Inc, London, 1988).

[23] F. Morehead, N. A. Stolwijk, W. Meyberg, and U. Gosele, App. Phys. Lett. **42**, 690 (1983).

[24] A. Ural, P. B. Griffin, and J. D. Plummer, Journal of Applied Physics **85**, 6440 (1999).

[25] M. Jaraiz, G. H. Gilmer, J. M. Poate, and T. D. de la Rubia, Applied Physics Letters **68**, 409 (1996).

[26] H. Bracht, E. E. Haller, and R. Clark-Phelps, Physical Review Letters **81**, 393 (1998).

[27] A. Ural, P. B. Griffin, and J. D. Plummer, Physical Review Letters **83**, 3454 (1999).

[28] P. E. Blöchl *et al.*, Physical Review Letters **70**, 2435 (1993).

[29] S. J. Clark and G. J. Ackland, Physical Review B **56**, 47 (1997).

[30] K. Kato, J. Phys.: Condens. Matter **5**, 6387 (1993).

[31] W.-C. Lee, S.-G. Lee, and K. J. Chang, J. Phys.:Condens. Matter **10**, 995 (1998).

[32] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, Phyisca **127B**, 401 (1984).

[33] Y. Bar-Yam and J. D. Joannopoulos, Physical Review B **30**, 2216 (1984).

[34] M. Tang, L. Colombo, J. Zhu, and T. D. de la Rubia, Phys. Rev. B **55**, 14279 (1997).

[35] M. Nastar, V. V. Bulatov, and S. Yip, Phys. Rev. B **53**, 13521 (1996).

[36] H. Bracht, N. A. Stolwijk, and H. Mehrer, Physical Review B **52**, 16542 (1995).

[37] E. Smargiassi and R. Car, Physical Review B **53**, 9760 (1996).

[38] W.-K. Leung, R. J. Needs, and G. Rajagopal, Physical Review Letters **83**, 2351 (1999).

[39] U. Gnutzmann and K. Clausecker, Applied Physics **3**, 9 (1974).

[40] P. Fahey, S. S. Iyer, and G. J. Scilla, Applied Physics Letters **54**, 843 (1989).

[41] U. Södervall, Quantitative Applications of Secondary Ion Mass Spectroscopy; Solid-State Diffusion and Mass Fractionation Studies, ph.D. Dissertation, Chalmers University of Technology, Göteborg, Sweden.

[42] P. M. Fahey, P. B. Griffin, and J. D. Plummer, Reviews of Modern Physics **61**, 289 (1989).

[43] F. Schaffler, Semicond. Sci. Technol. **12**, 1515 (1997).

[44] D.-S. Lin, T. Miller, and T.-C. Chiang, Physical Review B **45**, 11415 (1992).

[45] H. Oyanagi, K. Sakamoto, R. Shioda, and T. Sakamoto, Jpn. J. Appl. Phys. **33**, 3545 (1994).

[46] A. Ikeda *et al.*, Surface Science **385**, 200 (1997).

[47] R. Gunnella *et al.*, Phys. Rev. B **54**, 8882 (1996).

[48] L. Patthey *et al.*, Physical Review Letters **75**, 2538 (1995).

[49] H. W. Yeom *et al.*, Surf. Sci. **381**, L534 (1997).

[50] F. Liu and M. G. Lagally, Physical Review Letters **76**, 3156 (1996).

[51] F. K. LeGoues *et al.*, Physical Review Letters **64**, 2038 (1990).

[52] D. E. Jesson, S. J. Pennycook, J. M. Baribeau, and D. C. Houghton, Phys. Rev. Lett. **68**, 2062 (1992).

[53] X. Y. Zhu and Y. H. Lee, Physical Review B **59**, 9764 (1999).

[54] M. A. Leskovar, ph.D. Dissertation, University of Washington Department of Physics, 1998.

[55] W. F. Egelhoff, CRC Crit. Rev. **16**, 213 (1990).

[56] E. Kim, C. Chen, T. Pang, and Y. H. Lee, Physical Review B **60**, 8680 (1999).

[57] J. Tersoff, Physical Review B **39**, 5566 (1989).

[58] J. Hunter and W. Reinhardt, Journal of Chemical Physics **97**, 1599 (1992).

[59] P. Dorner, W. Gust, P. Predel, and U. Roll, Philos. Mag. **49**, 557 (1984).

[60] G. Hettich, H. Mehrer, and K. Maier, in *Defects and Radiation Effects in Semiconductors, Inst. Phys. Conf. Ser.*, edited by J. H. Albany (Institute of Physics, London, 1979), p. 500.

[61] J. Tersoff, Physical Review Letters **61**, 2879 (1988).

[62] J. Tersoff, Physical Review B **37**, 6991 (1988).

[63] J. Tersoff, Physical Review B **38**, 9902 (1988).

[64] B. C. Bolding and H. C. Andersen, Physical Review B **41**, 10568 (1990).

[65] D. E. Jesson, S. J. Pennycook, and J. M. Baribeau, Physical Review Letters **66**, 750 (1991).

[66] D. E. Jesson *et al.*, Physical Review Letters **70**, 2293 (1993).

[67] N. Ikarashi, K. Akimoto, T. Tatsumi, and K. Ishida, Physicsal Review Letters **72**, 3198 (1994).

[68] J. Z. Tischler *et al.*, Physical Review B **51**, 10947 (1995).

[69] W. Windl *et al.*, Physical Review Letters **83**, 4345 (1999).

[70] A. Antonelli, S. Ismail-Beigi, E. Kaxiras, and K. C. Pandey, Physical Review B **53**, 1310 (1996).

[71] C. S. Nichols, C. G. V. de Walle, and S. T. Pantelides, Physical Review B **40**, 5484 (1989).

[72] P. Kringhoj and A. N. Larsen, Physical Review B **56**, 6396 (1997).

[73] H. Haesslein, R. Sielemann, and C. Zistl, Physical Review Letters **80**, 2626 (1998).

[74] E. Silveira, W. Dondl, G. Abstreiter, and E. E. Haller, Physical Review B **56**, 2062 (1997).

[75] H. D. Fuchs *et al.*, Physical Review B **51**, 16817 (1995).

[76] *Gaussian 98* (Revision A.7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery,

R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.

[77] B. P. Uberuaga *et al.*, Physical Review Letters **84**, 2441 (2000).

[78] G. Henkelman, B. Uberuaga, and H. Jónsson, in preparation (unpublished).

[79] A. F. Voter and S. P. Chen, Mat. Res. Soc. Symp. Proc. **82**, 2384 (1987).

[80] P. J. Feibelman, Physical Review Letters **65**, 729 (1990).

[81] R. S. Mulliken, Journal of Chemical Physics **23**, 1833 (1955).

[82] R. F. W. Bader, in *Atoms in Molecules - A Quantum Theory* (Oxford University Press, Oxford, 1990).

[83] B. B. Stefanov and J. Cioslowski, Journal of Computational Chemistry **16**, 1394 (1995).

[84] Gaussian 94 User's Reference, page 32.

[85] H. C. Andersen, Journal of Chemical Physics **72**, 2384 (1980).

[86] D. F. Watson, Computers & Geosciences **8**, 97 (1982).

[87] T. Dyke and J. Muenter, Journal of Chemical Physics **59**, 3125 (1973).

[88] E. R. Batista, S. S. Xantheas, and H. Jónsson, Journal of Chemical Physics **109**, 4546 (1998).

VITA

- BA in Physics. University of Idaho. May 1994
- MS in Physics. University of Washington. March 1996
- PhD in Physics. University of Washington. June 2000