Defects in Crystalline Silicon: Integrated Atomistic and Continuum Modeling

by

Talid R. Sinno

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

~~Jan 1998~~ February 1998

Author .....
Department of Chemical Engineering
January 23, 1998

Certified by .....
Robert A. Brown
Dean of Engineering and Warren K. Lewis Professor of Chemical Engineering
Thesis Supervisor

Accepted by .....
Robert E. Cohen
St. Laurent Professor of Chemical Engineering
Chairman, Departmental Committee on Graduate Students

Defects in Crystalline Silicon: Integrated Atomistic and Continuum Modeling

by

Talid R. Sinno

Submitted to the Department of Chemical Engineering
on January 23, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

One of the most crucial measures of silicon wafer suitability for device manufacture is the defect distribution within the silicon material. Lattice defects are responsible for a host of device failures, especially as the feature length of microelectronic devices decreases. Minimization of silicon defects is one of the most pressing issues in silicon technology today. This thesis presents a computational framework for modeling microdefect formation during crystal growth. Microdefect formation is initiated by the transport and aggregation of point defects. Experimentally observed microdefects are essentially mesoscale phenomena, existing in the region of $O(10^{-6})$ m and typically consisting of up to $O(10^{8})$ point defects. The objective of this thesis is the prediction of microdefect distributions within macroscopic crystals as a function of crystal growth operating conditions.

The computational framework consists of two distinct approaches. Point defect equilibrium and transport properties are investigated quantitatively using an empirical interatomic potential in an optimized molecular dynamics algorithm. Atomistic simulation is also used to compute thermodynamic properties of small point defect clusters. These results, combined with phenomenological models for larger clusters are used to obtain a quantitative estimate for point defect aggregation rates. A generalized phenomenological framework based on the solution of coupled species conservation and rate equations is then used to compute defect distributions in macroscopic crystals, with the microscopic properties computed by atomistic simulation as input. Realistic thermal fields and crystal geometries are incorporated from simulations employing large-scale heat transfer models of actual crystal growth systems.

The model is first used to investigate the dynamics of the Oxidation-Induced Stacking-Fault Ring (OSF-ring) which appears in Czochralski (CZ) grown silicon crystals after thermochemical processing. The OSF-ring has been studied extensively because it exhibits unique and reproducible dynamics where it changes radial position depending on the crystal growth rate and the thermal environment during growth. Numerous studies indicate that microdefects which lie inside the ring region are formed by vacancies while those that lie outside the ring are self-interstitial in na-

ture. This observation provides a direct link between the location of the OSF-ring and the point defect distribution pattern within a crystal and allows us to model OSF-ring dynamics without explicitly considering the nucleation/aggregation problem.

An optimal set of point defect thermophysical properties is found by fitting the preexponential coefficients of the diffusion coefficients and equilibrium concentrations to a single set of experimental OSF-ring data. The model is used to successfully predict OSF-ring dynamics for a wide range of crystal growth conditions. An asymptotic analysis of the point defect balance equations correctly predicts the scaling of the OSF-ring location with the ratio of the crystal growth rate to the axial thermal field at the melt/solid interface, or V/G, which has also been demonstrated experimentally.

The complete defect dynamics model is used to investigate the size distribution of vacancy clusters in crystals grown under vacancy-rich conditions, with the OSF-ring at the periphery of the crystal. All of todays commercial CZ crystals are grown in the vacancy-rich regime because of the higher pull rates allowed and because interstitial-type dislocation loops found outside the OSF-ring are too large to allow reliable device manufacture. The control of void defects is one of the most pressing issues in silicon crystal growth because of their well-documented detrimental effects on device performance. All point defect thermophysical properties are maintained as computed with the OSF-ring model. The predictions of void size distribution as a function of crystal growth conditions are in quantitative agreement with various experimental measurements. The void density is predicted to depend linearly on the cooling rate of the crystal across a wide temperature range, which scales as the product of the crystal growth rate and the thermal gradient in the relevant portion of the crystal. The fact that this relationship is observed experimentally supports the hypothesis that simple homogeneous nucleation is a significant mechanism for void growth. The radial non-uniformity of the void distribution is investigated using a sequence of one-dimensional axial slices of the same crystal. The results demonstrate that the void size distribution shifts towards smaller voids as the OSF-ring location is spproached. In the region immediately next to the OSF-ring, interstitials exist in sufficient quantities to severely retard void formation, leaving a high number of free vacancies frozen into the crystal. These are presumed to provide the free volume needed for the formation of silicon dioxide precipitates which in turn provide the nuclei for the stacking-faults in the OSF-ring region.

The model in this thesis provides a self-consistent, quantitative description of microdefect formation during silicon crystal growth. Fitting to a single set of experimental data leads to a set of thermophysical properties that appear to be applicable to a very wide range of growth conditions. The results presented for the formation of void defects provide one of the first indications that quantitative defect engineering in silicon crystal growth is feasible with our present understanding.

Thesis Supervisor: Robert A. Brown
Title: Dean of Engineering and Warren K. Lewis Professor of Chemical Engineering

Acknowledgments

As I sit down to express my thanks to the many people that made the completion of this thesis possible, I realize that its not going to be as easy as I thought. So many people have affected my life in the last few years and I'm sure I will not be able to mention them all here. Firstly, I would like to greatly thank my advisor, Professor Robert Brown, for all his input, support, and guidance. I arrived at MIT without much in the way of a clear idea of what I was going to do. Bob quickly changed that by leading me into an area I knew nothing about but which, upon some consideration, seemed very interesting. During the course of my thesis work, Bob has always provided me with motivation, ideas and encouragement. Above all, Bob's perpetual insistence on the highest quality work has instilled in me a sense of self-criticality. I would also like to thank the members of my thesis committee, Professors Klavs Jensen, Gregory Rutledge, and Sidney Yip for input during committee meetings and for reading the various documents associated with my stay at MIT. I thank Professor Yip for making my introduction into the realm of atomistic simulation so enjoyable. I would also like to thank Professor Dimitrios Maroudas for his invaluable help and friendship during my early years here as I struggled to get started (and he struggled to get out of here).

The last four years of this thesis were undertaken as part of a research consortium sponsored by four industrial organizations which have greatly contributed to the direction and success of the work in this thesis. These organizations are MEMC Corporation, SEMATECH, Shin-Etsu Handotai Company, and Wacker Siltronic. I thank these companies for their financial support, and for their input on many issues, especially in guiding my attention towards problems of current interest. Much of the computational resources used in the generation of the results in this thesis were provided by their generosity. I would like in particular to thank Erich Dornberger for all his help and especially for his friendship over the last four years.

Thanks and acknowledgements must go out to the multitude of friends I have had the luck to have with me here at MIT. First, I wish to thank my officemates

and colleagues, Howard Covert, Tatsuo Mori, Tony Caola, Garrett Sin, and Aleksei Lipchin, who had to endure my use of the office as a social center, bike workshop, and changing room. The countless pots of (often burnt) coffee will never be forgotten. My endless discussions on every topic under the sun with Howard have definitely left their mark on my outlook on life (I mean that in a good way, I'm pretty sure) and when we part ways, will be sorely missed.

In assessing the various contributions to my relatively lengthy stay at MIT, it is very obvious to me that two people are mainly responsible: Shahin Ali and John Tolsma. Strange as it may seem, I thank these two for their close and unconditional friendship and for all the wonderful times we have shared over the years. I can think of few days where the three of us did not spend at least some time together and even fewer that were not better for it. Sha's unique brand of advice on many matters has and will continue to be very useful. Thats not to say that many other people were not part of my day to day life outside the confines of the office. Many thanks to Fred von Gottberg for being a great friend and even for a part of this thesis! His and Antonia's latest role as my homeless shelter providers is greatly appreciated. Thanks to Melody Swartz for her friendship and for making my years that much more colorful. Thanks to Sandeep "Dancing Queen" Patel, an indispensible member of the nightclub crew. Thanks to John Konz for overseeing my entry into bike racing. My efforts at reciprocation for this have unfortunately been much less healthy. I would also like to thank Arline Benford and Carol Phillips for making this department a much more pleasant place - I have enjoyed every conversation we have had. Many others have helped me from afar: my old and dear friends Tory Corbett, Terry Carino, Sanjay Motwani, Eric Bosco, Bassam Sarkis, Dino Loizides - thank you.

Without question, the most loving support has come from my family who have never had anything but words of encouragement throughout my stay at MIT. I give my thanks and love to my parents, Rabih and Najat Sinno, and to my brother Wael, all of who have had to deal with many outbursts and tales of doom and gloom during my many ups and downs. Thank you so much for everything you have done for me and given me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the defining aspects of modern times has been the development of the microelectronic device. These devices are critical components in an enormous variety of applications, ranging from the obvious, such as computers and other electronic equipment, to modern household appliances, such as ovens and alarm clocks. Common to all types of electronic devices is the semiconducting substrates on which they are manufactured; of these, crystalline silicon is by far the most important and commonly used. Thus, the properties of silicon have been studied extensively, allowing for the rapid development of silicon processing technology. Despite the wealth of knowledge currently available, silicon wafer production technology is not well understood because many of the advancements in crystal quality that have been achieved have been the result of trial and error experimentation in the absence of a physical understanding of the relevant microscopic processes.

The evolution of large, complex microelectronic devices coupled with increasing miniaturization has been the driving force for improvements in the quality of single crystal silicon. Shimura (1989) and Lawrence and Huff (1982) describe the process of device evolution in terms of the number of components per device. The age of very-large-scale integration (VLSI) saw devices with about $10^5 - 10^7$ components and was ushered in by the early 1980s. Ultra-large-scale integration refers to current technology and describes devices with up to 10^9 components (Shimura, 1989). Despite improvements in miniaturization, the transition from VLSI to ULSI has resulted in

chip size increasing from about 40 mm^2 well over 100 mm^2 which greatly reduces the number of devices that can be fabricated on a single wafer. Thus, there is a strong tendency to grow crystals with larger and larger diameters to keep device processing costs down by allowing for a larger number of devices to be fabricated on each wafer. Today's wafers are mainly 200 mm in diameter while the next generation of wafers will be 300 mm (von Ammon, 1996).

Wafer suitability for microelectronic device fabrication is determined by several different measurements that include surface uniformity, perfection of the underlying lattice, impurity concentrations, and crystal electrical properties such as resistivity (Huff and Goodall, 1996). Increasing device size increases the requirements on surface uniformity, while decreasing device feature length imposes additional constraints on the allowable size of lattice perturbations. For a given concentration of defects, losses (i.e device failures) are greater if fewer, more expensive devices are fabricated on the wafer surface. Crystal quality will therefore be of utmost importance in the coming years as these trends continue to decrease the tolerances on wafer specifications.

The fabrication of electronic devices and circuits on a silicon substrate involves a long and complex sequence of processes, both physical and chemical (Ghandhi, 1983). The first steps involve the reduction and purification of quartzite (silica) into pure polycrystalline silicon, which is subsequently melted and used to grow single crystals, usually in cylindrical ingot form. The ingots are then sliced into wafers about 500 μm thick; about 30 % of the crystal material is lost in this process. The cutting process results in a significant amount of surface damage and non-uniformity. This is corrected with chemical etching which removes about 15-75 μm of wafer surface. Etchants are usually a combination of nitric and hydrofluoric acids. The wafer surface is then chemically and mechanically polished to ensure a high degree of flatness, which is required for device manufacture. This process removes an additional 10-50 μm of material from the wafer surface (Ravi, 1981).

After these processing steps, the silicon wafer is ready for oxidation, which results in a silicon dioxide layer on the surface of the wafer. Oxidation is usually performed

thermally, by annealing the wafer in a steam ambient at a temperature ranging from 1050 K - 1450 K. The resulting oxide barrier is used to direct the subsequent diffusion of dopants into the silicon substrate. These dopants are used to affect locally the electronic properties of the silicon lattice so build features such as junctions and transistors (Ravi, 1981). Examples of electrically active dopants are boron, phosphorous, and arsenic. These dopants alter the electronic properties of the lattice by either donating an electron (the group V elements) or providing a hole (the group III elements).

One of the most crucial measures of wafer suitability for device manufacture is the defect distribution within the silicon material. In this thesis, the term defect is used to label any irregularity in the silicon lattice, either due to the misalignment of silicon atoms or the introduction of impurities during processing. Lattice defects are responsible for a host of device failures, especially as the feature length decreases. Minimization of silicon defects is one of the most pressing issues in silicon technology today. There are very specific guidelines for mapping out progress in this area. Current wafer surface specifications limit defect concentrations to 0.17 cm^{-2} for defects 0.12 μm and larger. By 2004, it is expected that wafers will contain no more than 0.05 cm^{-2} of defects with size 0.04 μm or larger. Similar improvements in iron impurity levels are expected; current tolerances are at 1×10^{11} cm^{-3} and by 2004 are expected to be below 1×10^{10} cm^{-3} (Huff *et al.*, 1995).

This thesis is entirely devoted to the study of the formation processes of lattice defects in silicon, both at the microscopic and macroscopic levels. Although defect profiles are ultimately determined by all stages of processing, the focus here is on crystal growth. The thermophysical conditions during crystal growth determine the essential features of the resulting defect profiles in the crystal, while the subsequent thermochemical annealing steps affect defect profiles on a secondary level.

1.1 Crystal Growth Techniques

There are currently two crystal growth methods with which most of todays single crystal silicon is grown; these are the Floating-Zone (FZ) and the Czochralski (CZ) techniques (Ghandhi, 1983). Of these two methods, CZ growth is the most prevalent in the production of wafers for device manufacture; the FZ technique which allows for the production of purer material is generally used in the production of the high-resistivity material used in power transistors.

In the CZ method shown schematically in Figure 1-1, polycrystalline silicon is melted in a crucible and a seed single crystal with the required crystallographic orientation is used to pull an ingot at a controlled rate from the melt (Teal and Little, 1950; Brown, 1988). Specification of the thermal environment is one of the most important issues in CZ growth (Lin and Benson, 1987). This is especially important in the growth of large-diameter crystals (larger than 100 mm) because of radial cooling. In general, as the radius of the crystal increases more heat is trapped in the center portion of the crystal which leads to greater radial temperature gradients. Such temperature gradients lead to significant thermal stresses which can lead to dislocation formation (Billig, 1956). Dislocations generated at high temperature can easily multiply and destroy the single crystal. To avoid such a scenario, the crystal thermal environment is controlled with the use of carefully placed heat shields which reduce surface cooling rates. The crystal diameter is controlled by by adjusting both the growth rate and the power supplied to the melt heater. Crystal diameter, heater power, and growth rate are related by the heat balance at the melt/solid interface, where the latent heat released by the solidification process is an important factor (Lin and Benson, 1987).

Chemically, one of the defining features of the CZ method is the use of a quartz liner within the crucible (Ravi, 1981), which reacts with the silicon to introduce oxygen into the melt; this is incorporated into the solid as the crystal grows. The degree of incorporation is dependent on the flow pattern in the melt due to the rotation of the crucible and crystal (Shimura, 1989). It is also dependent on the segregation

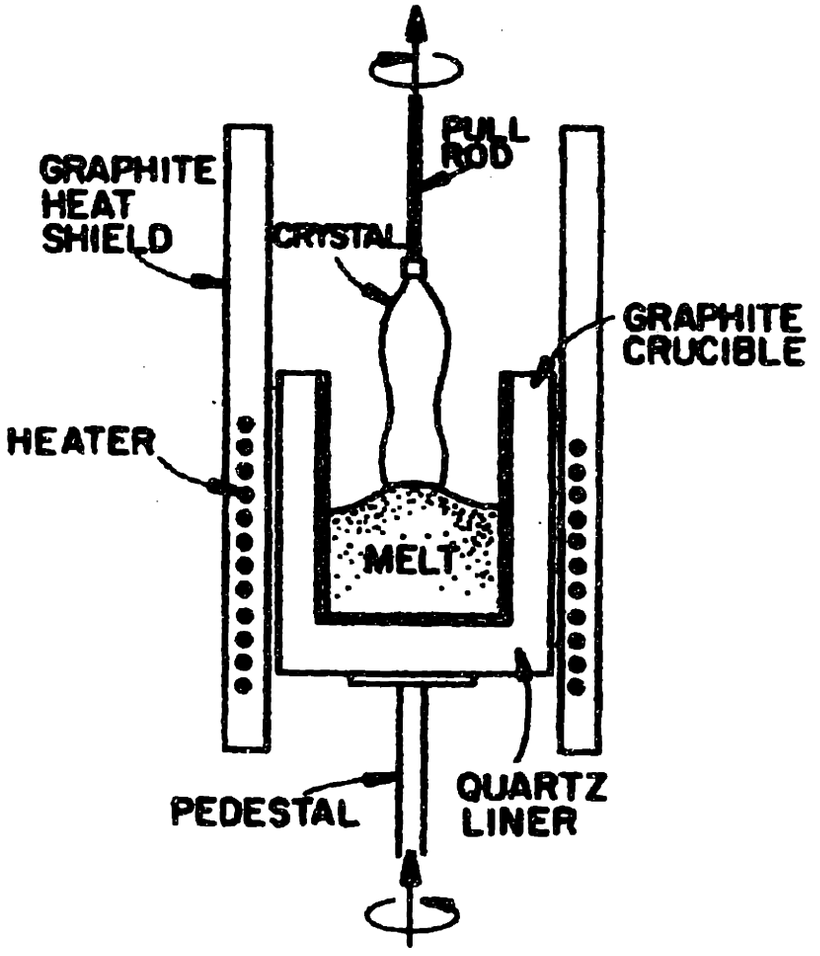

Figure 1-1: Schematic of the Czochralski crystal growth system. Taken from Brown (1988).

coefficient for oxygen which describes the relative affinity of oxygen in the melt and solid phases (Shimura and Huff, 1985) and the degree of oxygen evaporation from the melt surface during growth (Lin and Benson, 1987; Chaney and Varker, 1976; Kaiser and Kech, 1957). Oxygen is typically incorporated in concentrations of $O(10^{18})$ cm^{-3} becomes higher than the solubility limit as the crystal cools during growth. The supersaturation leads to the formation of oxide precipitates (Inoue *et al.*, 1982).

In FZ growth, a polycrystalline rod of silicon is converted into a single crystal by moving a molten melt zone of silicon from one end of a polysilicon rod to the other; the initial zone being created by contact with a single crystal seed. The heating is usually provided by an RF induction coil (Keck *et al.*, 1954) which radially surrounds the melt zone as shown in Figure 1-2. FZ silicon is not subject to crucible contamination and thus contains oxygen concentrations two orders of magnitude lower than those typically found in CZ silicon, which can be as high as 5×10^{18} cm^{-3}. This is a major reason for the different appearance of the defect structures typically found in the two types of crystals. Despite the advantages of lower impurity levels in FZ crystals, a serious limitation of FZ wafers arises because of the difficulty of growing large diameter crystals (Keller and Muhlbauer, 1981). The diameter of FZ crystals is limited by the stability of the floating zone which is primarily destabilized by gravity and is held together by the forces of melt surface tension and melt/solid interface tension (Brice, 1973). As the zone becomes larger, the stabilizing effects of surface tension rapidly diminish relative to gravity and the zone becomes unstable. To overcome this, many researchers have developed techniques to maximize the size of FZ crystals without introducing steep thermal fields which lead to the formation of bulk defects. The reader is referred to Keller and Muhlbauer (1981) for a discussion of the various techniques used to grow large diameter FZ crystals.

1.2 Defects in Crystalline Silicon

Two major classes of crystal defects exist in silicon; *native* defects which are caused by misplaced silicon atoms and *impurity* defects formed by the presence of foreign atoms.

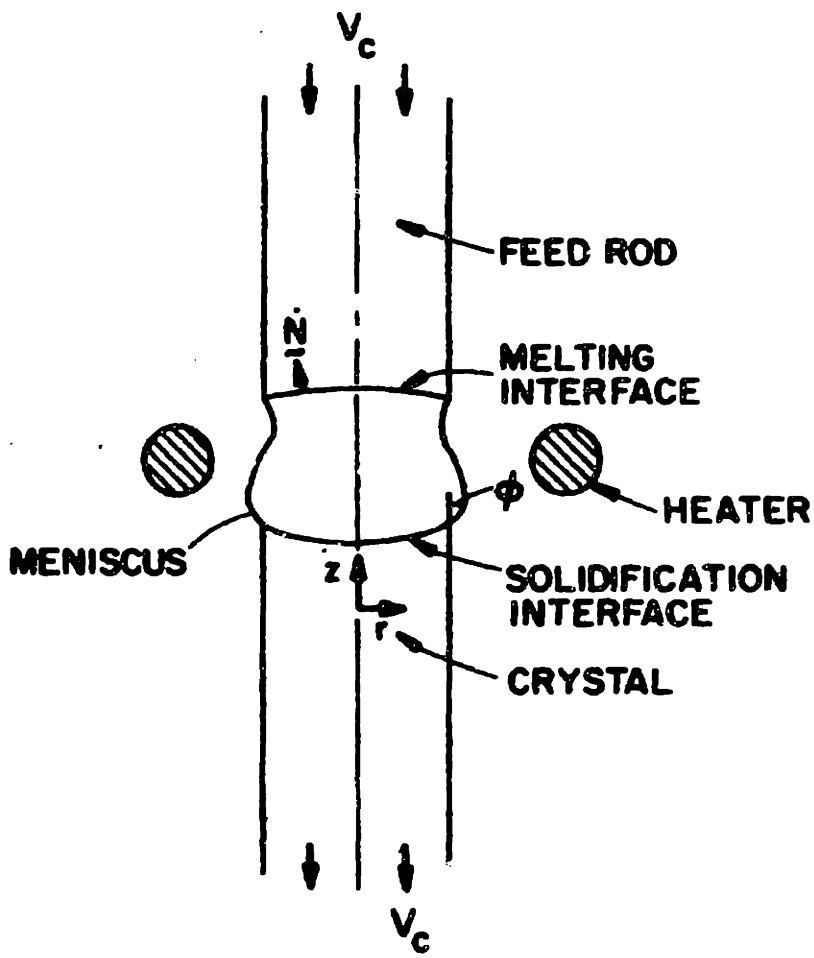

Figure 1-2: The floating-zone crystal growth system. Taken from Brown (1988)

The most basic native defect is the point defect. The two fundamental native point defects are the *self-interstitial* and the *vacancy*, which correspond to an excess silicon atom placed on an off-lattice site, and a missing silicon lattice atom, respectively, as shown in Figure 1-3. The native point defect disrupts the lattice on the length scale of several lattice spacings or on the order of 1–10 Å. Thermodynamics predicts a finite equilibrium concentration of native point defects at finite temperatures for any crystal (Girifalco, 1973). These point defects do not adversely affect crystal properties if they exist in dilute concentrations equivalent to the equilibrium value or lower. When present in supersaturation, however, native point defects lead to

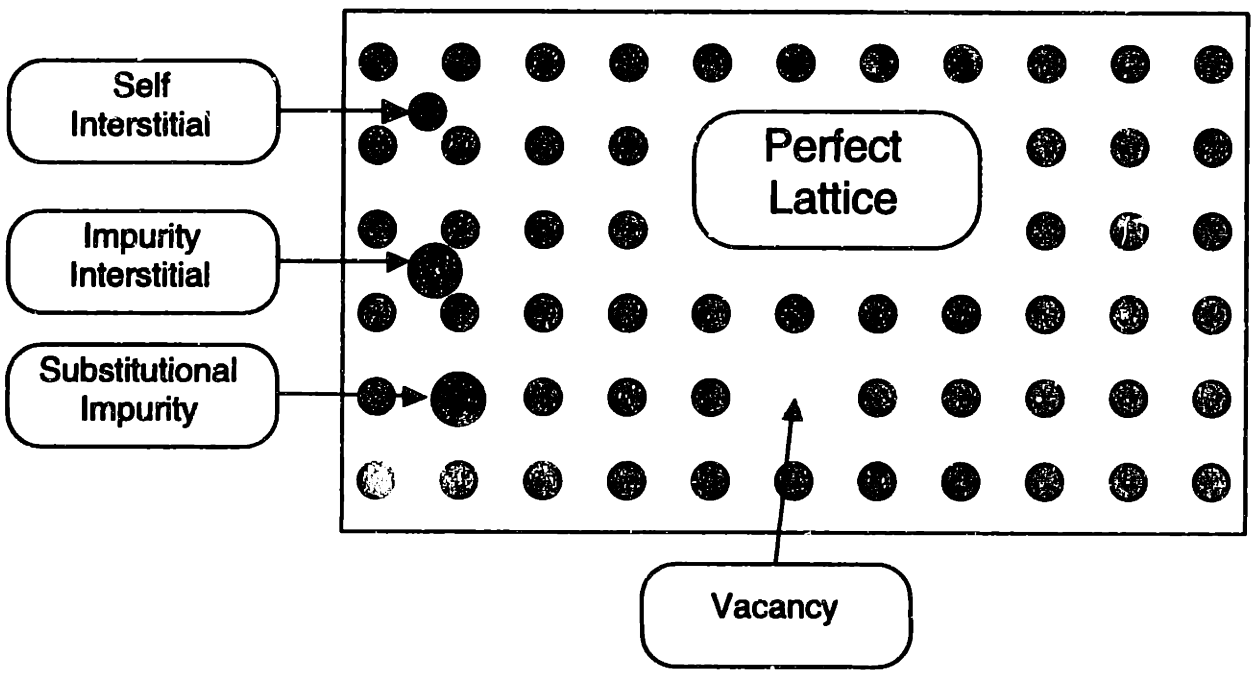

Figure 1-3: Schematic representation of various point defect species commonly found in crystalline silicon

a variety of larger, more complex defects. These structures can be classified either as line-defects (bulk dislocations), or three-dimensional agglomerates (microdefects). The term microdefect is used here to refer to any irregularity in the crystal that is not

a point defect or a line dislocation. In silicon growth by either FZ or CZ methods, it is now possible to produce single crystals with zero bulk dislocation density (Dash, 1958; Zulehner, 1989). Microdefect formation, however, is still not fully controllable and is being studied actively throughout the industry.

Microdefects have been observed experimentally by various methods such as x-ray topography following copper decoration (Roksnoer and van den Boom, 1981), and transmission electron microscopy (Petroff and de Kock, 1975). Microdefects range in size, composition, and homogeneity to produce a host of features such as striations, or "swirls", microscopic dislocation loop clusters (Petroff and de Kock, 1975) and large voids (Itsumi *et al.*, 1995). Typically, these structures are less than one micron in size, and thus are composed of 10^4–10^8 individual point defects. It is now accepted that microdefects are primarily the result of "condensation" of supersaturated point defects and impurities into a new phase.

Impurity atoms within the lattice complicate the picture even further. Not all impurities in silicon crystals are unintentional, however, as dopants like boron and phosphorous frequently are used to modify the electrical properties of the silicon substrate used in electronic device fabrication (deKock and Wijgert, 1980). Figure 1-3 represents schematically the two states that impurity point defects can occupy within the silicon lattice; either occupying a silicon lattice site (substitutional) or an off-lattice position (interstitial). Of the uncontrolled impurities, oxygen and carbon are the most common. These atoms have been shown to be responsible for the formation of various defect structures in addition to modifying the distribution of native defects. Examples of oxygen related defects include Stacking-Faults (Hasebe *et al.*, 1989) and SiO_2 microprecipitates (Inoue *et al.*, 1982). Carbon has been observed to dramatically affect microdefect profiles and is thought to act as a heterogeneous nucleant (Abe and Harada, 1983b).

1.3 Modelling Strategy

The goal of this thesis is to present a computational framework for modeling microdefect formation during crystal growth. The most formidable aspect of such a task lies in the vast range of length and time scales which a comprehensive picture of defect formation must encompass. Microdefect formation is initiated by the transport and aggregation of point defects which are characterized by length and time scales $O(10^{-10})$ m and $O(10^{-12})$ s, determined by the lattice parameter and crystal vibrational frequencies respectively (Girifalco, 1973). Experimentally observed microdefects are essentially mesoscale phenomena, existing in the region of $O(10^{-6})$ m and typically consisting of up to $O(10^{8})$ point defects. The objective of this thesis is the prediction of microdefect distributions within macroscopic crystals as a function of crystal growth operating conditions, where length scales are $O(10^{-1})$ m and time scales $O(10^{2})$ s as determined by crystal ingot geometry and growth rate.

Clearly, it is not possible to employ uniformly a single model across the entire range of length and time scales discussed above. Hence a computational approach must necessarily be multiscale in order to be tractable. At the atomic level, point defect thermophysical properties, i.e. their equilibrium, transport, and kinetic properties are direct functions of electronic and atomic processes. These properties are investigated either computationally with atomistic simulation or experimentally with various techniques such as dopant diffusion-profile measurements. In the case of atomistic simulation, the physical description of the atomic/electronic system is usually in the form of an interatomic potential that describes the force field exerted by each atom on the others. The structure of this force field is determined by the electronic structure of the bonds that link the atoms in the system. There are generally two levels of detail that an atomistic simulation can consider; *ab initio* methods solve the quantum mechanical Shrödinger equation and explicitly compute electronic energies (see for example, Kelly and Car (1992)), while empirical interatomic potentials lump the effects of electron dynamics into analytical functions (Stillinger and Weber, 1985; Dodson, 1987). The latter approach is computationally much less expensive and al-

lows for the investigation of more complex defect structures and is the approach taken in this thesis. A discussion of the relative benefits and drawbacks of these different approaches is presented in Chapter 3.

Results of atomistic simulation provide the first estimates for point defect thermophysical properties which are used as input to continuum scale models for calculation of the point defect distributions within the macroscopic crystal. Such models typically consist of numerical solution of balance equations for point defects using finite element or finite difference methods, and account for point defect transport, reaction and aggregation during crystal growth.

The link between point defects and large microdefects ideally needs to be considered externally to both of the above approaches. Technologically important microdefects typically exist on a mesoscale of $O(10^{-6})$ m. Although it is currently possible to obtain detailed thermodynamic information for small point defect clusters containing $O(10^2)$ point defects, coarser nucleation models must be employed for the larger ones. Such models generally consist of free energy descriptions derived from phenomenological models; for example, treating microdefects as spherical entities and writing their total free energies as the sum of surface and bulk contributions.

It is important to note that detailed nucleation models are not always required in order to explain some aspects of microdefect behavior. An example of such a case will be discussed in detail in Chapter 5, where a native point defect picture is used in conjunction with a simple homogeneous nucleation model to describe the behavior of a complex defect phenomenon known as the Oxidation-Induced Stacking Fault Ring (Hasebe *et al.*, 1989), which has been shown to contain large stacking faults containing self-interstitials and oxygen precipitates. In this case, the model is not able to predict the composition and structure of the defects within the ring, but is able to describe well its position in response to changes in the macroscopic operating conditions.

1.4 Thesis Objectives and Outline

The goal of this thesis is to formulate and implement a computational scheme suitable for quantitative analysis of microdefect formation within a silicon crystal free of bulk dislocations. Such an analysis, if quantitatively predictive, has the potential of significant commercial value because certain microdefects are now posing limitations on the miniaturization of electronic devices. The scheme consists of numerical simulations on both atomistic and continuum scales. Figure 1-4 summarizes the various components of the modeling framework in this thesis in the context of *defect engineering*, which is defined here as the control of microdefect distributions in a crystal by optimization of the crystal growth operating conditions. Model components that

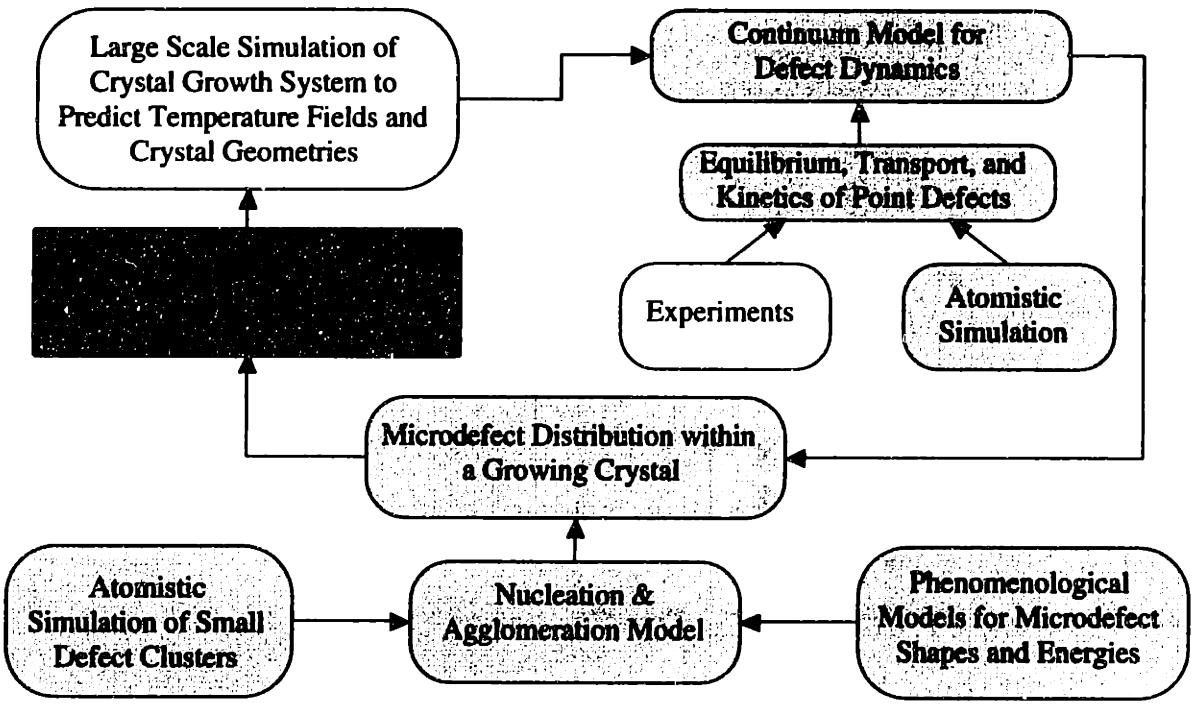

Figure 1-4: Multiscale modeling framework for analysis of defect dynamics in silicon.

are important to this overall goal but not explicitly a part of this thesis are shown in

a lighter shade. These are the simulation of the heat-transfer and melt/solid interface geometry within the crystal growth system which is used to provide temperature data for input to the continuum defect models, and experimental measurements of point defect properties with are used to validate our atomistic simulation results.

Point defect equilibrium and transport properties are computed using an empirical interatomic potential in an optimized molecular dynamics algorithm that allows for the direct calculation of these temperature dependent properties without assumptions except that of the validity of the potential. A model for the temperature dependence of the point defect recombination rate also is developed based on these calculations. Atomistic simulation is also used to compute thermodynamic properties of small point defect clusters. These results, combined with phenomenological models for larger clusters are used to obtain a quantitative estimate of point defect aggregation rates. A generalized phenomenological framework is then used to compute point defect distributions in macroscopic crystals, with the point defect microscopic properties as input. Realistic thermal fields and crystal geometries are incorporated from simulations employing large-scale heat transfer models of actual crystal growth systems.

A review of the current microdefect literature is presented in Chapter 2, where measurements in both FZ and CZ crystals are described. The second section of Chapter 2 discusses the various qualitative pictures that have been proposed to explain observed microdefect distributions. Chapter 3 describes atomistic simulation methods for estimating point defect dynamic properties. A review of the different classes of potentials and their results in the context of point defects in silicon also is presented. The results from the simulations carried out in this thesis are compared to current estimates within the literature, obtained from other theoretical work as well as experimental measurements.

Chapter 4 discusses the implementation of the molecular dynamics method on a parallel computer, in order to make feasible the investigation of point defect cluster thermodynamics. Clusters of point defects require much larger simulations to allow for adequate structural relaxation. The computational intensity of the molec-

ular dynamics method when complex interatomic potentials are employed requires the use of parallel architectures if this is to be a tractable computation. Free energy computations are used to determine aggregation rates for small clusters as a function of temperature. Clustering kinetics are also investigated for larger systems by monitoring the evolution of the cluster size distribution in time.

Chapter 5 describes the application of continuum-scale modeling to the study of a particular microdefect, the Oxidation-Induced Stacking-Fault Ring (OSF-ring). This defect structure appears after wafer oxidation of CZ-grown material in the form of an annular ring of stacking-faults which is observed to move radially with small changes in the operating conditions during growth, such as the thermal environment and crystal pull rate. Point defect distributions are linked to the location of the OSF-ring by using a highly simplified homogeneous nucleation model. The second part of the chapter discusses the development of a closed-form asymptotic solution to the model which is shown to lead to detailed information on the mechanism that produces OSF-ring dynamics. This analysis provides confirmation that our model for the OSF-ring is essentially correct. Chapter 6 builds on the model of Chapter 5 by including an explicit representation of point defect nucleation and aggregation. The results of this model are compared directly to measurements of void concentrations and sizes found in commercial silicon wafers. The model is used to probe the effects of thermal history on the size and density of microdefects as well as provide an explanation for the origin of the OSF-ring.

Chapter 2

Microdefects in Silicon

One of the most challenging aspects of developing a computational framework for predicting microdefect formation as a function of crystal growth operating conditions is the difficulty in obtaining precise information about the structure and formation mechanisms of microdefects. Over the last thirty years, many techniques for investigating the different defects that appear immediately after crystal growth (the *as-grown* state) have been developed. These have led to a wealth of valuable, but often inconclusive data. The purpose of this chapter is to give the reader a historical perspective on the gradual evolution of our understanding of microdefect formation and structure, leading up to our current understanding of these phenomena. This is important because many of the original issues in microdefects growth are still unresolved and debates in the current literature often stem from data presented more than a decade ago.

It is generally agreed that the source of most observable microdefects is the condensation of supersaturated point defects, both intrinsic (self-interstitials and vacancies) and extrinsic (impurity related), as the crystal cools during growth. Whether the nucleation process in each case is homogeneous, or assisted by heterogeneous sites, is less well understood. There are generally two types of experimental research performed on microdefects; the first focuses on the elucidation of the structure and composition of microdefects, while the second is concerned with the characterization of microdefect distribution response to changes in the growth conditions, such as the

crystal geometry, pull rate, and thermal history. The results of both of these studies are critical for any modeling effort that aims to be quantitative.

This Chapter presents an overview of our current understanding of microdefect origins in silicon single crystals grown by the floating-zone (FZ) and Czochralski (CZ) methods. While, the CZ method is used almost exclusively for electronic device fabrication, the FZ method is important to us because FZ crystals are chemically purer, and exhibit fewer oxygen related microdefects. Furthermore, microdefect data in FZ crystals is useful for the isolation of phenomena that are related only to native point defects. The behavior of intrinsic point defects is essentially the same in both FZ and CZ crystals; however, the much higher oxygen concentration in CZ crystals (10^{17-18} cm^{-3} in CZ as compared to 10^{15-16} cm^{-3} in FZ crystals) leads to significant differences in some of the observed microdefect structures. With this in mind, the FZ and CZ cases are are discussed separately, but whenever appropriate, the similarities are pointed out. In general, because FZ crystals are chemically simpler, microdefects found there will also be found in CZ crystals.

The first part of the Chapter summarizes the results of experimental observations in the literature, focusing mainly on commonly observed defects that have been studied in detail. These are presented in roughly the order in which they were discovered over the last three decades. The end of the section provides a summary of the current state of our understanding of microdefects and their relevance to the semiconductor industry.

The second part is devoted to the discussion of the numerous qualitative models that have been proposed to explain the origin of these structures. Issues such as the source of supersaturated intrinsic point defects, possible nucleation mechanisms, and the role of impurities are all considered in these models. These models have been postulated to explain the experimental observations but many of them are contradictory and some of them are now known to be incomplete or even incorrect. The goal of this discussion is to provide perspective on the wealth of data in the literature and thus isolate the most likely scenario upon which the modeling approach in thesis can be built.

2.1 FZ Crystals

Crystals grown by the FZ method are characterized by high purity unless intentionally doped. This usually results in a simpler microdefect distribution profile as compared to CZ. The main contaminants in FZ crystals are carbon and oxygen which are usually found in levels below $1 \times 10^{16}\ cm^{-3}$ in todays crystals. These concentrations are significantly lower than the high temperature solid solubility of both species and little homogeneous precipitation of either can be expected to result during FZ growth.

2.1.1 A- and B-defects

Ever since microdefects in a dislocation-free silicon crystals were described by Plaskett (1965) , there has been a major effort to understand their origin and composition. The microdefects observed by Plaskett using copper-decoration, in both FZ and CZ crystals, exhibited a characteristic circular distribution in the plane perpendicular to the crystal growth direction that has led to the term "swirl defects". In longitudinal sections, these defects appeared in a striated distribution. Figure 2-1 shows a radial cross-section of an FZ crystal containing A-defects in a striated pattern (Föll and Kolbesen, 1975).

Later, deKock (deKock, 1970) investigated swirls using x-ray topography following copper decoration and established that the swirl region consists of two main kinds of defects, distinguished by their different sizes, which are now termed as A- and B-defects. Electron microscopy subsequently showed that the larger A-defects were perfect dislocation-loops of interstitial type, whereas the smaller B-defects could only be observed with copper decoration (Föll and Kolbesen, 1975). This led to the conclusion that B-defects exhibited low stress fields and were three-dimensional aggregates of supersaturated self-interstitials because A- and B-defects were generally observed to appear together. There is still controversy as to whether these defects are nucleated homogeneously (Föll and Kolbesen, 1975) or heterogeneously with the aid of carbon impurity (Vysotskaya *et al.*, 1990). The interstitial nature of these defects, however, is well established by transmission electron microscopy. Figure 2-2, taken

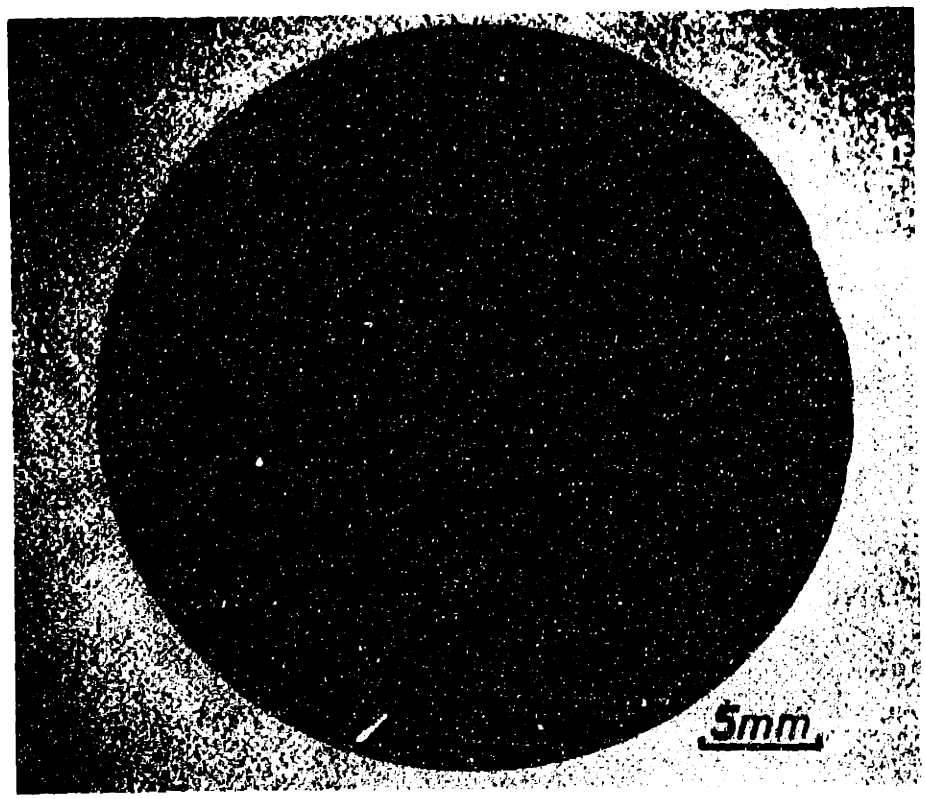

Figure 2-1: Striated distribution of A- and B-defects in a radial cross-section of and FZ crystal. Taken from Foll et al. (1983)

from Foll et al. (1977), shows the complex structures that these dislocation loops can exhibit, and also shows their enormous typical size which is far larger than the 0.2 μm device linewidth of todays devices.

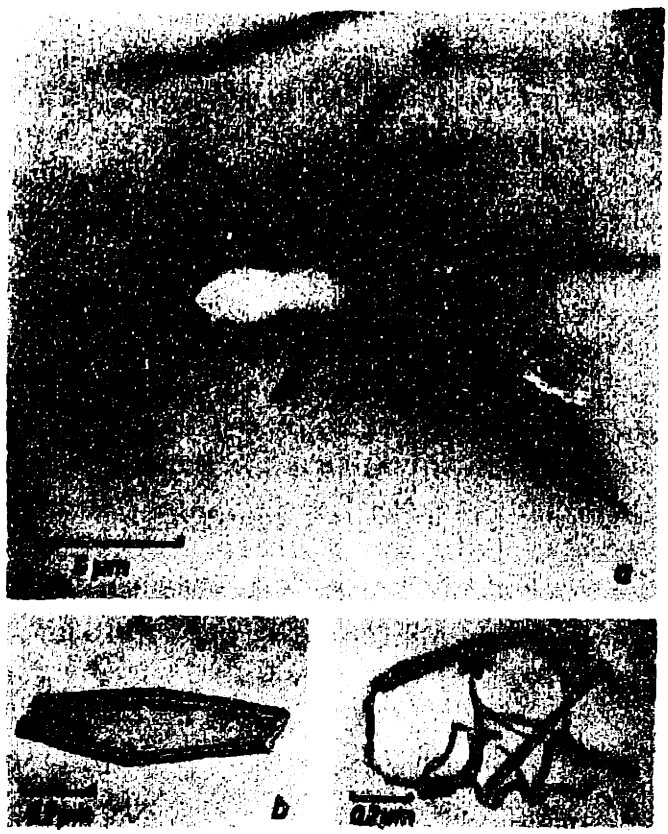

Figure 2-2: Interstitial dislocation loops linked to A-defects. Taken from Foll et al. (1977)

The striated pattern that A- and B-defects appear in is thought to be due to the inhomogeneous distribution of point defects caused by micro-fluctuations in the melt/solid interface (Ravi, 1981). These fluctuations are the result of the fact that the melt/solid interface does not advance smoothly but in steps of several atomic layers. This phenomenon will be discussed in more detail in Section 2.5 as a possible source for point defects in the lattice.

2.1.2 D-defects

Further investigation of FZ crystals with copper decoration followed by electron microscopy showed yet another, fundamentally different, kind of microdefect (Roksnoer and van den Boom, 1981). These defects did not appear in striated distributions and only appeared at crystals grown at higher pull rates (> 6 mm/min for a 23 mm diameter crystal). These defects were quickly associated with vacancies because they responded to changes in growth conditions in an opposite fashion to A- and B-defects. It was subsequently found that changes in the crystal pull rate and diameter strongly influenced the distribution of these defects (Abe and Harada, 1983b) as well as swirl defects. Figure 2-3 is an axial section x-ray topograph that shows the response of swirl defects and D-defects to changes in the crystal growth rate. At higher growth rates (4 mm/min), swirls disappear at the surface of the crystal, while D-defects extend from the center outwards. At lower pull rates, D-defects eventually disappear into the center of the crystal, making way for the swirls. Note the apparently defect-free regions between the swirl and D-defect areas and near the radial periphery of the crystal. The apparent lack of defects in this region has led to the term "neutral zone". Microdefect distribution was also found to depend strongly on crystal diameter. For wider crystals, the pull rate at which D-defects appeared and swirls disappeared was lower. Figure 2-4 (Abe and Harada, 1983b) shows the region in two-dimensional parameter space for which swirl defects are seen in FZ crystals. It shows that as the radius of the crystal increases, the pull rate at which swirls vanish decreases monotonically.

2.1.3 Other FZ defects

Closer inspection of the neutral zone between the swirl and D-defect regions with secondary ion mass spectroscopy (SIMS) measurement of the copper decoration level shows that this region does in fact contain very small defects, termed I-defects (Abe and Harada, 1983b). These defects possess the same decoration intensity as the swirl defects (which is ten times higher than the D-defect level) but very little else can be said about their composition; they have been linked to oxygen in the crystal and

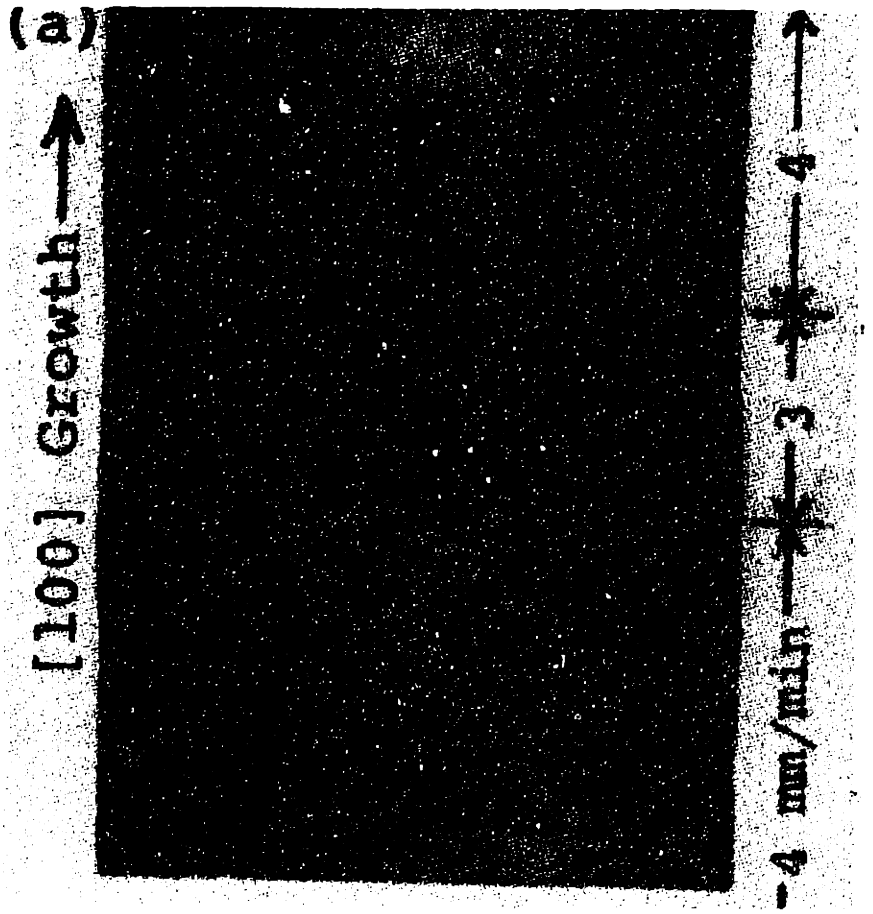

Figure 2-3: Swirl defect and D-defect dependence on changes in the pull rate. X-ray topograph of an axial cross-section of FZ crystal taken from Abe et al. (1983)

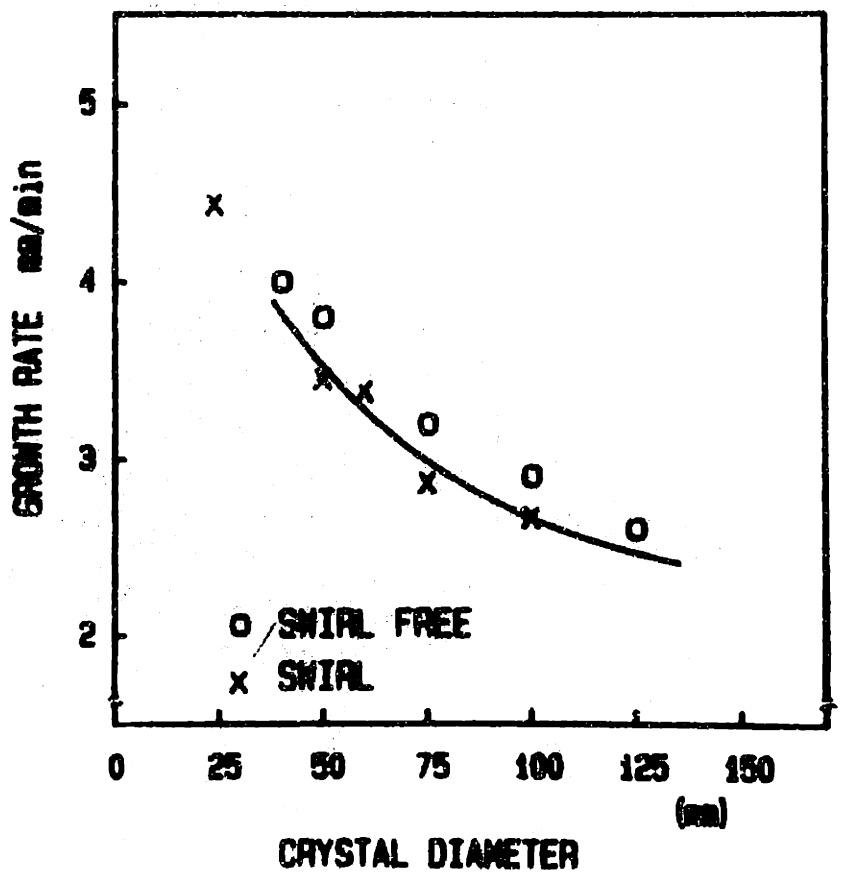

Figure 2-4: Swirl defect dependence on crystal diameter and pull rate. Taken from deKock et al. (1974)

could be oxygen microprecipitates (Abe and Harada, 1983b). Due to their very small size, very few other studies have attempted to characterize them.

2.1.4 Summary of Microdefect Behavior in FZ Crystals

The results of the experiments described in Section 2.1.2 led to the conclusion that the controlling parameters for determining whether A,B, or D defects appeared in FZ crystals, are the crystal thermal field (which is very strongly affected by the radius of the crystal due to changes in the heat transfer profile), and the pull rate during growth. Because point defects were known to be responsible for microdefect formation, it was further reasoned that the crystal growth conditions directly affected the balance between the two intrinsic point defect species and this effect was responsible for the observed microdefect distribution after growth. This important conclusion is the basis for most quantitative models of microdefect behavior today. The work of Voronkov (Voronkov, 1982) was the first to semi-quantitavely explain the behavior of A,B and D defects in terms of point defect dynamics. Voronkov's model is discussed in more detail in Chapter 5, in the context of the point defect model developed in this thesis.

While many conclusions about the nature of these microdefects in the literature are in conflict, a majority (including the work of Voronkov (1982)), essentially put forth the following picture. B-defects are small three dimensional aggregates of interstitials which form at least partially due to interstitial supersaturation. A-defects are also interstitial related consisting of large, extrinsic perfect dislocation loops. These are possibly formed by the collapse of B-defects. In vacancy-rich regions, D-defects appear which are poorly characterized but due to their opposing behavior to changes in the operating conditions, are thought to be vacancy related. Whether oxygen plays a part in the formation of these defects is still not well understood. In crystals, where both vacancy and interstitial rich regions exist simultaneously, a neutral zone appears in between, where only very small, possibly oxygen related, defects are observed following decoration. No defects are observed with direct methods.

Despite the relative chemical purity of FZ crystals, microdefects in FZ suffer from

relative obscurity relative to their CZ counterparts. The Czochralski growth technique is far more prevalent in the electronic device industry and the bulk of industrial research is directly almost exclusively on defects in CZ crystals. There is very little quantitative data available describing the concentration and size distribution of microdefects in FZ crystals. Our interest in them will be mainly for the purposes of confirmation of our models developed for CZ microdefects.

2.2 Microdefects in CZ Crystals

In order to discuss coherently the multitude of microdefects in CZ crystals it makes sense to first consider those which are analogs of species found in FZ crystals, and then those that only appear in CZ crystals. This partitioning can be useful because it gives information on the possible formation processes that control certain microdefect species. In particular, microdefects that appear in both crystals in similar distributions can generally be assumed to result primarily from intrinsic point defect dynamics, while those observed only in CZ crystals can be assumed to result primarily from the high oxygen levels. This inference can be made because the equilibrium concentrations of intrinsic point defects are determined by identical thermodynamics in both growth methods, while oxygen concentration is the single largest chemical difference between crystals grown by the two methods.

2.2.1 A-defects and the Swirl Region

CZ silicon crystals, under certain growth conditions, often exhibit similar defects to the ones described above in FZ crystals. Swirl-distributed defects have been observed and, like in FZ crystals, they disappear for higher growth rates (deKock and Wijgert, 1980). These can be also be assumed to be related to self-interstitial supersaturation (Ravi, 1981). This region of CZ crystals is commonly referred to as the A-region and exhibits large pits which have been found to be dislocation loops similar to the ones found in FZ crystals (Sadamitsu *et al.*, 1993). The extent of the A-region with respect to the crystal growth rate and thermal environment is also the same as found

in FZ crystals.

Oxygen does not appear to play a major role in this part of the crystal. This can be explained by noting that oxygen precipitation should be suppressed in areas of interstitial excess because of strain effects. The precipitation of oxygen in silicon leads to significant local compressive stress because SiO_2 requires more than double the volume per silicon atom than the perfect crystal contains. Thus, in regions of self-interstitial supersaturation, oxygen precipitation should be retarded.

2.2.2 D-defects

D-defects are of much greater technological importance than A-defects in CZ silicon because all commercial crystals are grown so that the D-defect region extends across the entire radius of the crystal. There are several reasons for this, the most important being the fact that the large dislocation loops observed in the A-region of the crystal lead to device failure because the defect size can be larger than the device feature length. Another consideration is that in order to grow the crystal exclusively in this region, the pull rate has to be very low, which is commercially undesirable. For these reasons, the investigation of D-defects represents a major aspect of the work in this thesis. D-defects exhibit much of the same properties in CZ as they do in FZ crystal growth (Yamagishi *et al.*, 1992). D-defects have been directly linked to a particular form of device performance degradation by several researchers (von Ammon *et al.*, 1996). In particular, the so-called Gate-Oxide Integrity (GOI) of electronic devices is impaired in crystal regions where the D-defect density is high. By contrast, the interstitial-dominated A-region in CZ crystals exhibits excellent GOI characteristics.

Because of their technological importance, several methods have been developed to detect D-defects. These methods measure fundamentally different properties but none are capable of delineating the entire microdefect population. The detection limit for D-defects with any technique is currently about 50 nm in size. It is important to note however, that smaller microdefects are still detrimental to device performance and simulation is an excellent method for estimating the concentrations of defects of all sizes.

Yamagishi et al. have developed an etching process by which D-defects could be studied in more detail than previously possible with the copper decoration technique. This etching leads to the delineation of wedge-shaped flow-patterns which originate at tiny etch pits; these structures are now called Flow-Pattern Defects (FPDs). Figure 2-5 shows a photograph of these defects, after wafer etching.

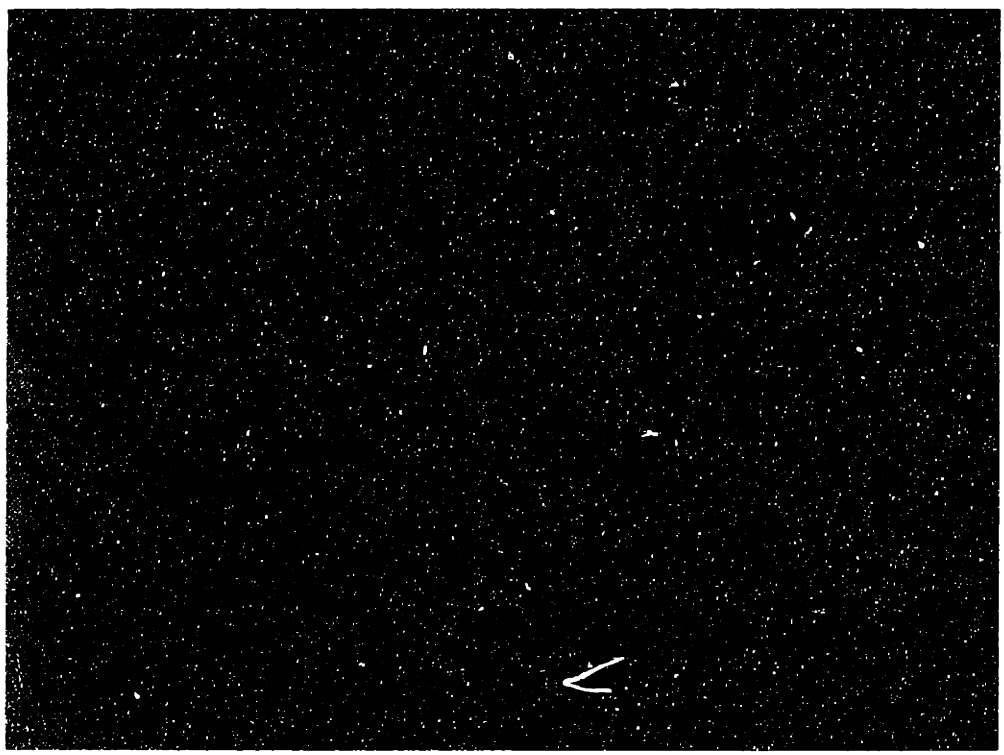

Figure 2-5: Flow Patterns after etching in a region of high D-defect density. Optical micrograph taken from Yamagishi et al. (1992)

Another study, using a wafer cleaning technique known as SC1 (Kern and Puotinen, 1970) led to the observation of particles on wafer surfaces when a laser particle counter was used (Ryuta *et al.*, 1990; Morita *et al.*, 1992). These defects were termed Crystal Originated Particles (COPs) and are now known to be linked to the presence of FPDs. However, they found in concentrations about two orders-of-magnitude

higher than FPDs; this fact casts doubts on whether these two methods are measuring the same D-defect. Yet another technique for evaluating the presence of D-defects is Light Scattering Tomography (Moriya and Ogawa, 1983; Moriya, 1989; Hourai *et al.*, 1995). Defects found with this technique are known as Light-Scattering Tomography Defects, or LSTDs. These defects are observed in similar concentrations to COPs and are distributed in similar patterns to both FPDs and COPs.

The work of Takano et al. (Takano *et al.*, 1995a) sheds some light on the nature of the various kinds of D-defects. They showed that the densities of all three microdefect types are highly correlated in their response to changes in the thermal history during growth. Figure 2-6 shows the relationship between concentrations of FPDs and LSTDs for crystals grown under different cooling conditions. While the actual concentrations are very different, the two microdefect levels are obviously correlated in a linear fashion. Furthermore, Dornberger et al. (1997) have shown that the FPD density is directly correlated to the GOI value for any crystal. These observations have consolidated the results of numerous studies on GOI-related defects and essentially lead to the following conclusion: FPDs, COPs, and LSTDs are all D-defects and are responsible for the degradation of the gate-oxide integrity of electronic devices.

While these investigations have led to different methods of detecting D-defects, conclusions as to D-defect origin have been speculative at best. Until recently, most researchers have thought that D-defects in CZ crystals were large, amorphous SiO_2 precipitates (Wada *et al.*, 1982; Itsumi *et al.*, 1992), fundamentally different to the D-defects in FZ crystals which were thought (but never proven) to be vacancy clusters. However, recently, an important experimental study has presented strong evidence based upon transmission electron microscopy (TEM), that D-defects in CZ crystals are essentially microvoids caused by the condensation of supersaturated vacancies (Itsumi *et al.*, 1995; Itsumi *et al.*, 1996). These voids were observed to appear in octahedral geometries with facets oriented along the $< 111 >$ planes of the bulk crystal, which is in accord with the fact that $< 111 >$ surfaces exhibit the highest density and therefore the lowest energy in diamond lattices. Such voids have been estimated to be about $100-300\ nm$ in diameter and pose serious problems for device

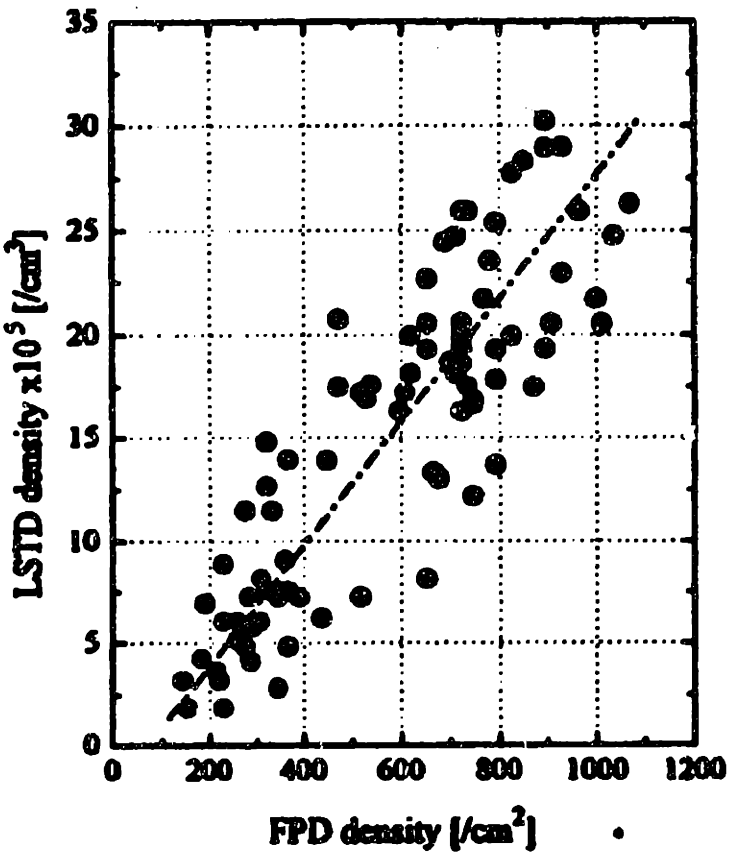

Figure 2-6: Correlation between the FPD and LSTD response to crystal growth thermal history.

reliability. The Itsumi study also found that small quantities of silicon dioxide are usually associated with void formation. Whether the oxide is incorporated during nucleation or after growth is not yet known.

2.2.3 D-defect Response to Thermal History During Growth

D-defects (also called GOI-defects because of their detrimental effects on that property) in CZ crystals are one of the most intensely researched defects at this time. In order to better understand the relationship between their distribution and the crystal growth conditions, several researchers have designed experiments to correlate their density with the thermal history of the crystal during growth (Dornberger and von Ammon, 1996; Takano *et al.*, 1995a). These studies are now being used to develop quantitative models for D-defect formation, and such a model is the subject of Chapter 6.

Crystals grown with different cooling profiles have led some researchers to speculate that there is a critical temperature region in which most D-defect nuclei grow. This temperature region is not entirely agreed upon, with estimates scattered across a wide range, 1173 K-1323 K (Dornberger *et al.*, 1996b), 1353 K-1423 K (Takano *et al.*, 1995b), 1273 K-1423 K (Hourai *et al.*, 1995), above 1273 K (Sueoka *et al.*, 1994), above 1473 K (Nakajima *et al.*, 1994), and 1073 K-1473 K (Abe and Takeno, 1992). Takano et al. (1995) have grown several crystals with different cooling rates in the temperature range 1353 K-1423 K, and then computed corresponding D-defect concentrations in the as-grown crystals. There is clearly a very well defined dependence of the final defect density on the cooling rate; Figure 2-7 shows a plot of the FPD density versus the cooling rate in the temperature range.

Dornberger et al. (1997) have performed similar experiments using a wider range of cooling rates in the temperature range 1173 K-1323 K, which they believe to be the relevant range. These results also indicate a high degree of correlation between GOI related defect density and cooling rate in the specified temperature range. The fact that such a region exists can be explained in terms of a homogeneous nucleation mechanism which is dependent on the presence of high point defect supersaturation.

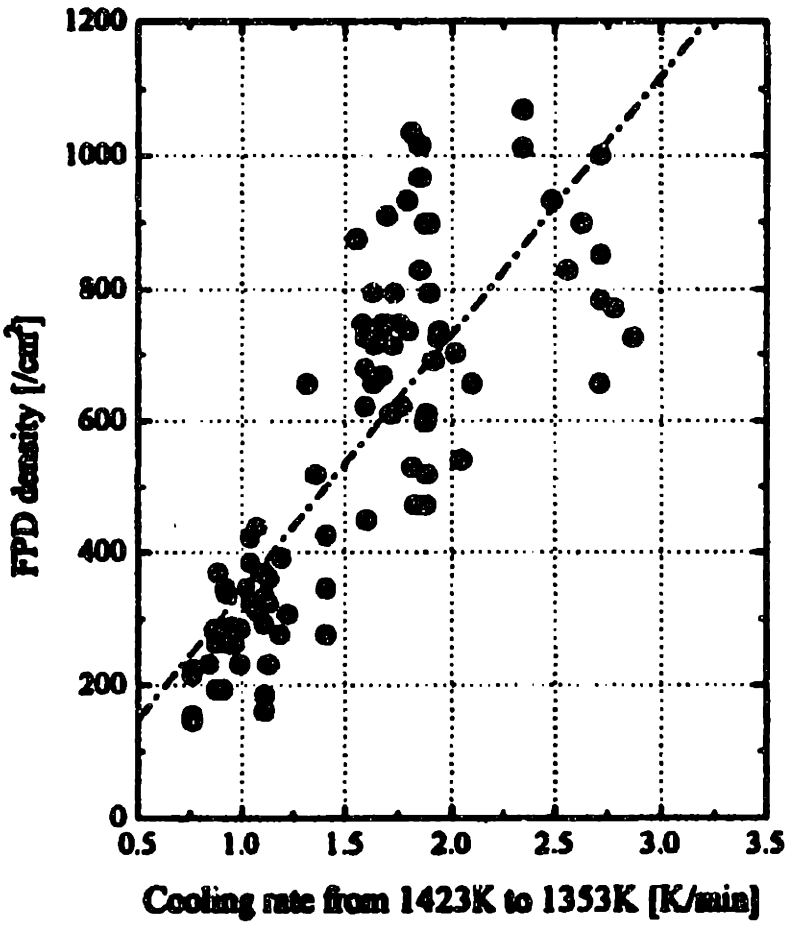

Figure 2-7: Correlation between FPD density and cooling rate in the temperature range, 1353 K-1473 K. Taken from Takano et al. (1995).

This point would correspond to the upper temperature in the ranges given above. The lower limit then corresponds to the point where aggregation is halted due to the decreased mobility of point defects as the crystal cools, or the depletion of the supesaturated monomers.

2.2.4 Current Status of D-defect Research

The technological importance of D-defects for advancement of device miniaturization has been emphasized because it represents the leading edge of crystal growth research. The sizes of these defects are rapidly approaching the feature length of electronic devices at which point they will lead to complete device failure. Furthermore, single devices are becoming extremely complex, and therefore expensive, making the costs associated with device failure prohibitively high. These factors provide the motivation for the research described in Chapter 6 of this thesis.

The basic assumption in our approach will be that vacancy supersaturation is the main driving force for D-defect formation, and that other species such as oxygen only play secondary roles. Specific questions that will be addressed are whether we can predict quantitatively the distribution of D-defects for a given set of operating conditions. Whether this is possible with a native point defect model will be an indication of how close to the real physical picture our assumption is. Evidence for the fundamental role of native point defects in specifying the microdefect profiles in both CZ and FZ crystals will be presented later in this Chapter and in detail in Chapter 5.

Of particular importance is the structure of the thermal field imposed within the crystal (Winkler and Sano, 1994) and the issue of whether D-defect nucleation and growth is indeed limited to a fairly narrow temperature range. This issue has tremendous potential value for the reduction of these detrimental defects with appropriate specification of the heating environment during the crystal growth process. The use of heat shields to influence the thermal history of a crystal is already an essential component in commercial crystal growth and this approach for reducing D-defects in CZ crystals would likely be the most expedient from a cost point of view.

There is some evidence that the relevant portion of the cooling profile is not as well defined as hoped. An indicator of this fact is that the agreement between the different estimates for the boundaries of this range are not in agreement. If all estimates are combined, we find that temperatures ranging from near the melting point all the way down to about 1173 K have been associated with D-defect evolution. Takano et al. (1995) investigated the correlation between FPD and LSTD densities and dwell time in the temperature range 1685 K - 1473 K and found positive correlation, albeit subject to large uncertainties; their results are shown below in Figure 2-8.

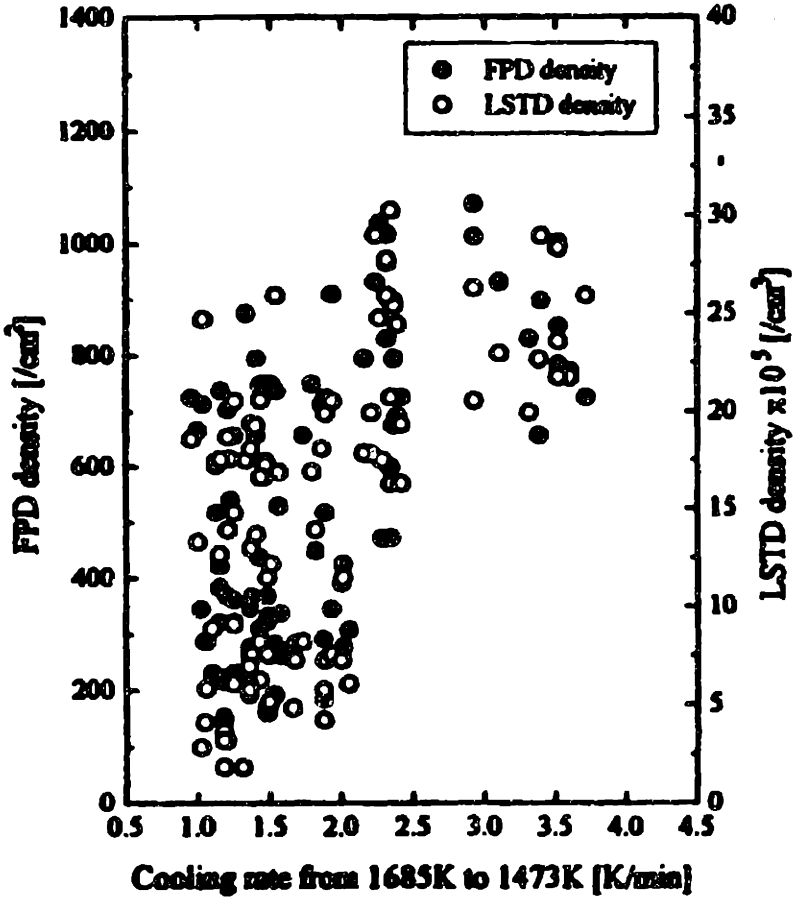

Figure 2-8: Correlation between FPD density and cooling rate in the temperature range, 1473 K-1685 K. Taken from Takano et al. (1995).

2.2.5 Oxygen Related Microdefects

Oxygen is introduced in CZ-grown crystals by the dissolution of the quartz crucible into the silicon melt during growth. CZ crystals contain oxygen at levels up to 10^{18} cm^{-3} which is greater than any other impurity. Typical specified oxygen concentrations in CZ crystals are between 5.0×10^{17} cm^{-3} and 8.0×10^{17} cm^{-3}. The role of oxygen in silicon is extensive, affecting crystal properties in both positive and negative ways and there are many methods used to control its levels such as gas flow control, magnetic fields, and the thermal environment (Izunome *et al.*, 1995; Kakimoto *et al.*, 1996; Togawa *et al.*, 1995)

Following growth, CZ wafers are heat treated to form oxide precipitates. These precipitates are used as internal metal impurity collectors, or *getters* (Tan *et al.*, 1977; Rozgonyi and Pearce, 1978). If the initial oxygen concentration during growth is too small, too few oxide precipitates form and metal impurities are uncontrolled. On the other hand, excessive oxygen levels cause too much precipitation which then interferes with device characteristics, leading to junction current leakage and other effects. Other uses for oxygen precipitates include the trapping of any line dislocations formed during wafer processing.

The precipitation of oxygen during post-growth annealing is well studied because of the usefulness of these gettering sites. Typically, a low temperature anneal at about 750 C is performed first to nucleate oxide precipitates. A high temperature anneal at about 1000 C follows to allow the nuclei to grow. The resulting density of precipitates is a function of the initial oxygen concentration (Zulehner and Huber, 1982), the presence of other impurities such as carbon, and thermal history of the crystal during growth. The role of carbon in oxygen precipitation is thought to be via the provision of heterogeneous nucleation sites.

Much less well characterized is the formation of oxide precipitates in the as-grown crystal. It has been put forth, that simultaneous supersaturation of oxygen and vacancies leads to the formation of very small oxide precipitates during crystal growth. The only way to observe the distribution of these "latent" nuclei is by subjecting the

crystal to post-growth annealing treatments. These experiments show non-uniform distributions of oxygen precipitation in different regions of the crystal which implies the presence of nuclei in the as-grown crystal. Inoue et al. (1982) have concluded that small oxide precipitates, about 1 *nm* in size exist in the as-grown crystal in high concentrations of about 10^{10} cm^{-3}. Evidence for the generation of oxygen precipitates during growth also comes from the observation that a thin layer of oxide is frequently associated with the octahedral voids described in Section 2.2.2.

The conclusion as far as oxide precipitation is concerned is that if they are formed during growth, they are very small, roughly about one to several *nm*. Oxygen has a high solubility in silicon and requires a large amount of free volume in order for precipitation to proceed. These facts have two main implications; the first is that oxygen precipitation proceeds at low temperatures, which the crystal spends little time at during the growth process. The second is that native point defects should play a significant role in the precipitation of oxygen during crystal growth by providing strain-releasing mechanisms. Section 2.5.2 describes the various mechanisms by which this can occur.

2.2.6 The Oxidation-Induced Stacking-Fault Ring

Separating the D-defect and A-defect regions in CZ wafers is an annular region that appears to be defect free in the as-grown crystal, but that develops a large concentration of extrinsic (self-interstitial-type) stacking faults after wafer oxidation in a steam ambient (Hasebe *et al.*, 1989; Abe and Harada, 1983a). The position of the oxidation-induced stacking-fault ring (OSF-ring) is entirely dependent on the conditions during crystal growth and not on the post-growth annealing treatment. This implies that oxygen-related nuclei, too small to be observed, are formed during growth which then develop during the thermochemical treatment of the wafers. This theory is supported by the fact that TEM has located tiny oxygen precipitates (silicon dioxide) at the center of the stacking-faults (Ravi, 1981). Thus, once these are formed during CZ growth, they provide nucleation sites for the self-interstitials that are injected across the silicon/silicon dioxide interface during the subsequent oxidation (Hasebe *et al.*,

1989). The injection process across the interface is driven by the same compressive stress model described in Section 2.2.1. Figure 2-9 shows a transmission electron micrograph of the stacking faults that appear in the OSF-region after annealing. These are interstitial-type because they represent an extra layer of atoms inserted into the lattice. They are oriented along the dense $< 111 >$ planes which is the minimum energy orientation. Figure 2-10 shows an X-ray topograph of a wafer with an OSF-ring

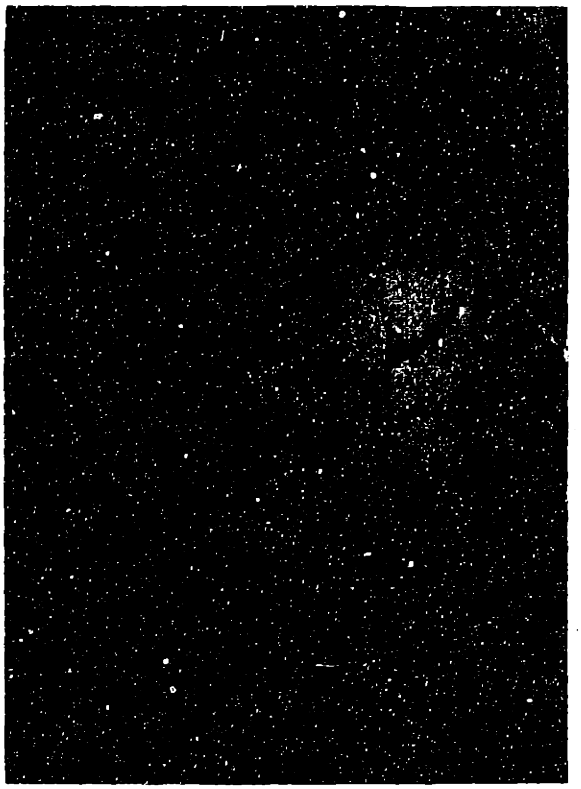

Figure 2-9: Transmission electron micrograph of an oxidation-induced stacking fault in an annealed wafer. Stacking faults A and B are bulk stacking faults, while C and D are stacking faults that terminate at the surface. Taken from Ravi (1981).

after wet oxidation (Hasebe *et al.*, 1989).

The OSF-ring is unique to CZ crystals and is never observed in FZ crystals even when they are doped with oxygen levels comparable to those found in CZ crystals. This anomaly has not yet been explained and seems to indicate that oxygen is not

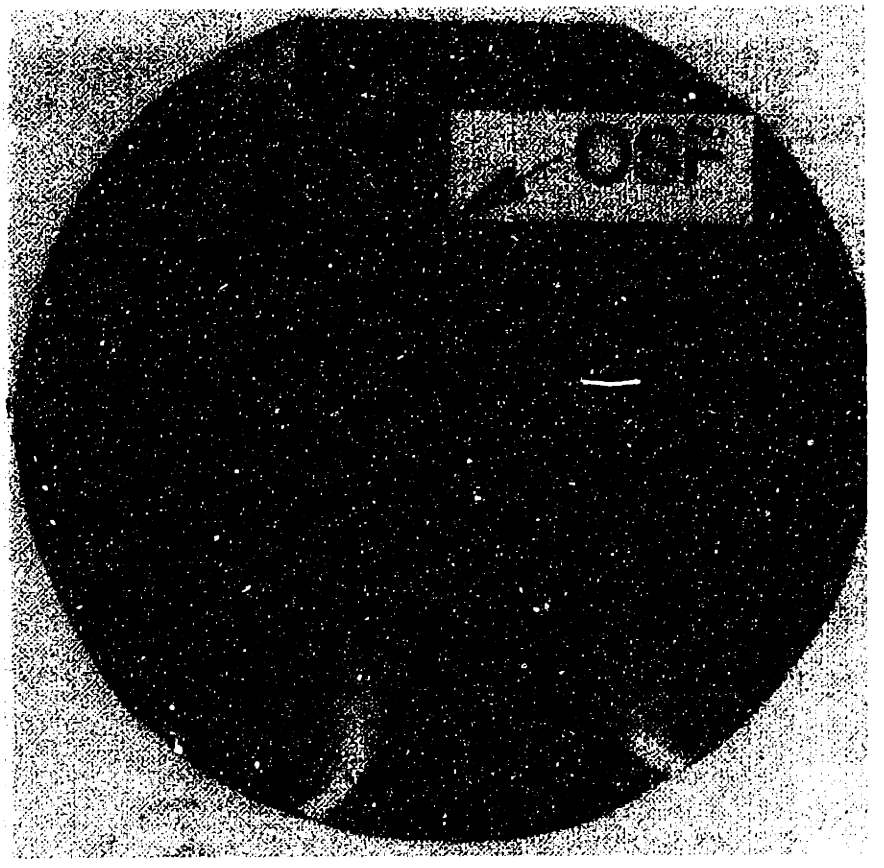

Figure 2-10: X-ray topograph of a CZ wafer with an OSF-ring appearing after wet oxidation processing. Taken from Yamagishi et al. (1992)

the only factor in the formation of the OSF-ring. In FZ crystals, the neutral zone between the A/B defect region and D-defect region contains only tiny defects which can only be observed indirectly with copper decoration. The common feature of the neutral zone and the OSF-ring, however, is that they both divide regions that contain microdefects associated with excess self-interstitials (exterior) and vacancies (interior). In other words both of these regions are primarily due to the dynamics of *native* point defects *during* crystal growth. This observation serves as the ansatz of the continuum model in Chapter 5 of this thesis; it essentially links the OSF-ring (and the neutral zone) to the distribution of intrinsic point defects.

The unique dynamics of the OSF-ring have led many researchers to attempt to explain the mechanisms that lead to its formation and predict its position in response to changes in the crystal growth rate and thermal history (Brown *et al.*, 1994; Habu and Tomiura, 1996; Wijaranakula, 1990). Interest in the OSF-ring is not driven by commercial pressure; crystal growers know empirically the conditions required to grow a crystal with the OSF-ring safely at the periphery. However, it has long been realized that understanding this very significant defect would lead to progress in understanding more subtle, and technologically important features.

The OSF-ring is observed to move radially inwards as the crystal pull rate is reduced, until at a certain pull rate, it disappears into the center of the crystal. This pull rate is a function of the thermal environment, particularly the axial temperature gradient near the melt/solid interface (Dornberger and von Ammon, 1996). Dornberger et al. (1996) have performed the most comprehensive experiments to date for characterizing the response of the OSF-ring to changes in the pull rate and axial temperature gradient. Figure 2-11 shows an X-ray topograph of an axial cross-section of a CZ crystal grown in a slowly varying pull rate.

A major portion of this thesis is directed at studying the OSF-ring. The motivation for doing this arises from the realization that the OSF-ring provides us with an opportunity to study point defect dynamics in detail without the use of difficult diffusion experiments. Specifically, we will be attempting to develop a model for predicting the radial location of the OSF-ring for a given set of operating conditions.

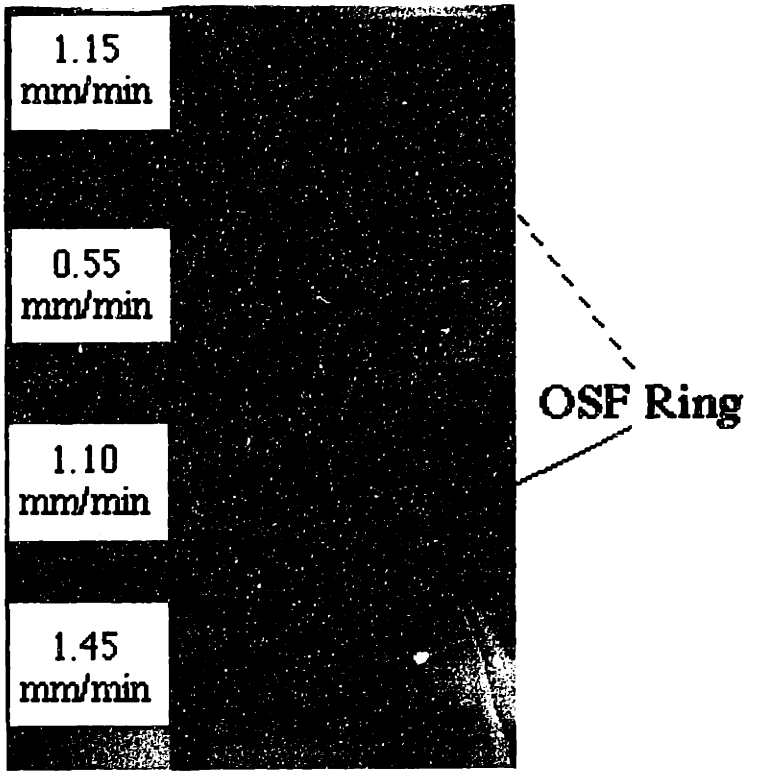

Figure 2-11: X-ray topograph of an axial cross-section of a CZ crystal showing the response of the OSF-ring to changes in the pull rate.

The starting point for the analysis is linking the OSF-ring to the distribution of native point defects within the crystal; attempting to model stacking fault growth explicitly from the beginning is not a reasonable goal because of the lack of a precise mechanism. Before quantitative prediction can be possible, detailed thermophysical property information for point defects is required; this is the goal of Chapter 3 which describes our attempts at computing point defect properties using atomistic simulation. The goals of this study will be to determine if indeed native point defect distributions are the sole factor in determining the location of the OSF-ring. If this is the case, we need to find the actual mechanism responsible for the behavior shown in Figure 5-1. Several authors have put forth potential mechanisms but none have been quantitatively successful so far (Habu and Tomiura, 1996; Wijaranakula, 1990; Brown *et al.*, 1994).

2.3 Effects of Impurities and Dopants on Microdefects

One method by which the composition of microdefects can be investigated is by intentionally doping crystals with impurities and observing the effect on microdefect density. The interaction between self-interstitials, vacancies and impurities has been extensively studied in dopant diffusion experiments (Zimmermann and Ryssel, 1992; Stolwijk *et al.*, 1986) and by directly observing doping effects on microdefect distributions (Abe and Harada, 1983b). Such experiments give information about the reactions that take place between impurities and native point defects and will be discussed in more detail in the context of point defect thermophysical properties in Chapter 3. Generally, the presence of certain impurities will strongly reduce the mobility or concentration of a particular point defect type; the observed effect on different microdefects can then used to infer their composition. Impurities can also strongly affect the nucleation rate of microdefects by reducing strain barriers. Large impurity atoms can facilitate the condensation of vacancies while small atoms oc-

cupying lattice sites can help ease the strain caused by locally high self-interstitial presence. The following is a summary of the effects of a few representative impurities and dopants common in both FZ and CZ crystals.

Carbon is a common impurity in both FZ and CZ crystals, introduced in the purification of the polysilicon feed material, and, at levels of about $5 \times 10^{16}\ cm^{-3}$, D-defects are suppressed while swirls are enhanced (Abe and Harada, 1983b). Clearly, carbon upsets the point defect balance in favor of self-interstitials. The exact mechanism for this process is not well understood. Carbon atoms however, are smaller than silicon, and thus provide a source of free volume which would aid in self-interstitial aggregation but effectively raise vacancy energy.

Nitrogen has a suppressing effect on both types of defects. This effect has been explained by speculating that nitrogen enhances the recombination reaction between interstitials and vacancies which depletes both species in equal numbers by the following reactions (von Ammon *et al.*, 1996):

$$N_2(Interstitial) + V \rightarrow N_2(Substitutional) + I \rightarrow N_2(Interstitial) \qquad (2.1)$$

The presence of nitrogen therefore effectively lowers the energy barrier against interstitial-vacancy recombination and has the net result of lowering point defect supersaturation.

Boron doping affects microdefect distribution in the same way as carbon, also increasing the self-interstitial presence in the crystal, moving the OSF-ring towards the center of the crystal in the case of CZ growth. In FZ crystals, the swirl region is observed to extend further into the crystal and survives at higher pull rates than shown in Figure 2-4. The mechanism by which boron alters the point defect balance is under investigation because boron is an important electrical dopant in CZ crystals (Ghandhi, 1983). Current possibilities under investigation are the formation of boron-point defect pairs which deplete native point defects or reduce their mobility.

Recent atomistic simulations (Rasband *et al.*, 1996a) have computed estimates for the equilibrium concentrations of boron-interstitial and boron-vacancy pairs which provide a possible mechanism for altering the point defect balance in a crystal. These

are represented by

$$B + I \longleftrightarrow BI \tag{2.2}$$

and

$$B + V \longleftrightarrow BV \tag{2.3}$$

Each of these reactions deplete the available supply of point defects, and, depending on the relative stability of the boron-point defect pairs can lead to a shift in the resulting point defect concentrations.

Antimony doping exhibits the opposite effect to boron and carbon doping. Here, the D-defect region is observed to extend across the entire radius of the crystal, even for lower pull rates that previousiy led to swirl defect generation. In this case the larger covalent radius of the antimony atoms stabilizes the formation of vacancy aggregates by consuming the excess free volume and lowering their free energies. Self-interstitial precipitation is suppressed for the same reason. The fact that most impurity effects can be explained by a simple free volume argument is fairly good evidence that swirl defects are linked to self-interstitial excess while D-defects are mainly composed of vacancies.

2.4 Summary of Observable Microdefects in Silicon

The previous sections have described some of the major microdefects that have been observed in FZ and CZ silicon. Some these are of current technological importance while others have been largely eliminated from commercial crystals. All are important from the point of view of understanding microdefect formation at a fundamental, quantitative level. The following table provides a summary of microdefect formation in both FZ and CZ crystals.

The rest of this thesis will focus on two particular microdefect structures; the OSF-ring and the D defect, which we interpret as a vacancy aggregate. The OSF-ring is well characterized inasmuch as we know the composition of the stacking faults

Defect Type and Reference	Detection Technique	Size Range	Characteristics
A-defect (Plaskett, 1965) (deKock, 1970)	Copper Decoration followed by x-ray topography Also visible with TEM.	0.1-1 μm	Interstitial type dislocation loops. Appear in striated or swirl patterns.
B-defect (Föll and Kolbesen, 1975) (deKock, 1970)	Copper Decoration followed by x-ray topography	$< 0.1\ \mu m$	3-D interstitial aggregates possibly heterogeneously nucleated by carbon.
D-defect Roksnoer and van den Boom (1981)	Copper Decoration followed by electron microscopy.	N/A	Three dimensional vacancy-related aggregates.
COP defect (Ryuta *et al.*, 1990)	NH_4OH/H_2O_2 Etching	$\approx 0.1\ \mu m$	D-defect detected by laser particle counting.
FPD defect (Yamagishi *et al.*, 1992)	Secco Etching	$\approx 0.1\ \mu m$	A type of D-defect in wedge-shaped flow patterns.
LST defect (Yamagishi *et al.*, 1992)	Laser Scattering	$\approx 0.1\ \mu m$	Considered a D-defect.
Bulk OSF (Hasebe *et al.*, 1989)	Wet oxidation followed by etching and x-ray topography	$O(1\ \mu m)$	Extrinsic stacking faults nucleated around an oxide precipitate.

Table 2.1: Summary of microdefects found in as-grown FZ and CZ silicon single crystals.

that appear in the ring region and how the radius of the ring responds to changes in the crystal pull rate and temperature field. What is not known is the mechanism that leads to its formation; i.e the specific atomic environment that leads to the nucleation and growth of interstitial-type stacking faults. Also not known is the mechanism that governs OSF-ring dynamics. There is strong evidence that intrinsic point defects are the main factors here but no model has yet succeeded in producing quantitatively predictive results.

Point defect dynamics are investigated in detail using atomistic simulation in the following Chapter. The goal is to compute thermophysical properties that are sufficiently accurate for use in a continuum scale model for OSF-ring dynamics. The advantage of this approach is internal consistency with respect to experimental data. Thermophysical property data derived from experiment must be indirectly extracted and is subject to large uncertainty. Chapter 5 aims to develop a model that can explain the behavior of the OSF-ring for a wide range of crystal growth operating conditions based on these estimates. We also seek to understand mathematically the observation that the OSF-ring position seems to be a function of the conditions near the melt/solid interface of the growing crystal.

Our goals for modeling D defect formation are essentially to be able to predict quantitatively the size distribution of these defects. Once again, we consider the growing evidence that D defects in CZ crystals are essentially voids and approach this task by assuming that D defects are the result of homogeneous nucleation of supersaturated vacancies. A major requirement for confirming the validity of this model will be its consistency with respect to the OSF-ring model. It will be assumed that native point defects are responsible for both defects; in the case of the OSF-ring we only require that point defects determine the *location* of the OSF-ring, and not necessarily its composition. In this context, consistency is imposed by requiring that the thermophysical data used in the OSF-ring model must also be used in the D defect model. This allows us to test both models against very different data.

2.5 Qualitative Models for Microdefect Formation

Before microdefect formation can be quantitatively predicted by simulation, a picture of the essential physics must be constructed. The first requirement is an adequate description of the microscopic processes involved, such as point defect transport, reaction, and aggregation (these processes are collectively referred to as *point defect dynamics*). The detailed quantitative characterization of point defect dynamics at the microscopic level is the subject of Chapters 3 and 4 and is not considered further here.

Other issues are the interaction of mobile point defect species with the crystal boundaries. The melt/solid interface is clearly the source of most species during growth; the difficulty lies in determining the incorporation mechanism which in turn determines the concentration within the crystal. Also, very little is currently understood about the role that the crystal cylindrical surface plays in shaping microdefect profiles. The observation of neutral zones at radial positions immediately adjacent to the surface in FZ crystals is indicative that some diffusional loss of point defects does occur here, but no quantitative information on surface point defect reactions is available.

Traditionally, microdefect formation models for FZ and CZ cases have been considered separately. The point of view taken in this thesis is that both systems are fundamentally similar. Although the basic constituents are the same in each case, the excess oxygen in CZ silicon must be given special attention because oxygen levels in a CZ crystal are above solubility limits for lower temperatures which is a driving force for oxygen precipitation. Most models for FZ microdefect generation have focused on the supersaturation of *intrinsic* point defects. Oxygen and carbon have been cast as nucleants, but the composition of the observed microdefects is mainly attributed to intrinsic point defects. On the other hand, CZ defect models are based heavily on the distribution of oxygen within the crystal. The following sections review two aspects of qualitative microdefect modeling; the first deals with the relative importance of self-

interstitials which applies equally to both FZ and CZ systems, while the second part is concerned with the nucleation mechanisms that form the experimentally observed microdefect structures. The latter considers oxygen-related defects separately.

2.5.1 Models for Point Defect Incorporation

The earliest efforts at explaining microdefect generation in FZ crystals were primarily concerned with determining the relative importance of self-interstitials and vacancies, and whether these were incorporated into the crystal by equilibrium or non-equilibrium mechanisms. A brief summary of some of the more common qualitative pictures based on the conclusions of Föll et al. (Föll *et al.*, 1977) is presented below. Note that these ideas are equally applicable to the case of CZ crystals. Models that are centered on the role of oxygen precipitation, as in the case of some CZ defects, are discussed separately in Section 2.5.2.

Equilibrium Models

Here, the term equilibrium refers to the incorporation of point defects into the crystal across the melt/solid interface. This idea was originally put forth by Seeger and coworkers (Seeger *et al.*, 1977; Föll *et al.*, 1977). These researchers assume that self-interstitials are incorporated into the crystal at equilibrium and exist in much higher concentrations than vacancies. They become supersaturated as the crystal cools and aggregate into B defects with the aid of inhomogeneous nucleants such as carbon. These three-dimensional aggregates then "collapse" into the observed dislocation loops that are known as A-defects once they reach a critical size, determined by their free energies. The model does not address D-defects or how vacancies affect the microdefect pattern.

Non-equilibrium Models

These models assume non-equilibrium point defect incorporation across the melt/solid interface and are based on the trapping theories of Webb (Webb, 1962) and Fletcher (Fletcher,

1976). These theories predict that a less perfect crystal is generated at higher pull rates due to the excess incorporation of point defects. Intuitively, these models are based on the concept that a rapidly advancing solidification front does not allow the system to fully achieve equilibrium because the random disorder in the liquid phase leads to a higher level of point defect incorporation. The degree of supersaturation is predicted to increase linearly with the *microscopic* velocity of the melt/solid interface. The microscopic growth rate is the actual velocity of the melt/solid interface which does not advance smoothly during crystal growth but rather in jumps, and, averaged in time, gives the observed *macroscopic* growth rate.

The increase of interstitial incorporation at higher pull rates cannot be used to explain the *vanishing* of swirls at higher pull rates, which implies decreasing self-interstitial presence. In order to do this, the concept of interface remelting must be employed. At lower macroscopic pull rates, the melt/solid interface can actually have negative velocity as the solid region immediately adjacent to it can temporarily melt. If this happens, the average positive microscopic growth rate must be higher than the macroscopic rate to make up for time lost in the remelt steps. Thus, since remelting decreases with increasing pull rate, this effect would lower the positive microscopic rate (which is responsible for defect incorporation) and decrease defect levels which gives the qualitatively observed trend for A-defect disappearance at higher rates. This model does not predict a *sudden* disappearance of A-defects at a critical growth rate and cannot be used to account for the reverse trend in D-defect distribution and has been discarded as a potential explanation.

The concept of remelting was used in a more plausible model for interstitial aggregation into A and B defects by Chikawa et al. (Chikawa and Shirai, 1977). Here, local remelting leads to the formation of microscopic liquid droplets near the melt/solid interface. Liquid silicon is denser than the solid phase, and point-defect pairs form at the surface of the locally melted region (Frenkel pair formation), injecting interstitials into the drop and ejecting vacancies into the surrounding bulk. This process allows the droplets to increase density as required by the silicon phase diagram and also leads to a supersaturation of vacancies at the melting temperature which can then

aggregate into D-defects. This model does not consider point defect recombination which is now known to proceed very rapidly at high temperatures (Antoniadis and Moskowitz, 1982).

Pure Vacancy Models

These models are based on the hypothesis that the vacancy is the only point species present in appreciable quantities (Van Vechten, 1975; Van Vechten, 1978). To explain the presence of interstitial-related dislocation clusters, Van Vechten has postulated several different mechanisms. The earliest ones (Matthews and Van Vechten, 1976) were based on impurity processes but these have been discounted because observations in ultrapure crystals still showed the presence of swirls. Another model was put forth to explain the formation of interstitial-type dislocation loops by the emission of vacancies (Matthews and Van Vechten, 1976). This is unlikely because already present vacancy supersaturation would require a large free energy source to allow this process to occur. Other explanations based on the vacancy-only picture proposed by Van Vechten require complicated and energetically unfavorable reconstructions to occur before experimental observations can be explained (Van Vechten, 1978).

These models show that it is very difficult to interpret A and B swirls in the context of a vacancy-only picture, and show that the presence of self-interstitials is crucial to explain these microdefects in a logical manner.

Vacancy-Interstitial Models

These models are based on the hypothesis that interstitials and vacancies coexist in thermal equilibrium. The processes that lead to this coexistence depend on the particular model. The model of Hu (Hu, 1977) allows for supersaturations of both defect to coexist by postulating that recombination is hindered by a large energy barrier, but irradiation experiments have shown that point defect recombination is actually very fast at high temperatures (Matthews and Ashby, 1973). Foll et al. (Föll *et al.*, 1981) therefore conclude that such a scenario is not likely and that vacancies and self-interstitials cannot exist in equally supersaturated states. This conclusion

is reasonable if one region of the crystal is considered but does not allow for the possibility that equal concentrations of self-interstitials and vacancies can lead to different microdefect profiles within the same crystal.

Perhaps the first model to recognize that precisely such a situation could arise is the model of Voronkov (Voronkov, 1982). This model was the first to be able to explain semi-quantitatively the transition between swirls and D-defects as a function of pull rate and thermal fields. It assumes that vacancies and interstitials are both equally important in shaping microdefect profiles and provides a very robust picture for the main defect regions found in FZ crystals. All recent efforts at numerically computing defect profiles during crystal growth are based on such a picture because of the overwhelming evidence that it is essentially correct.

The computational framework developed in Chapter 5 of this thesis, demonstrates that a vacancy-interstitial model is capable of explaining many of the experimental observations, not only in FZ crystals but also in CZ crystals, particularly the behavior of the OSF-ring, discussed in Section 2.2.6. Self-interstitials and vacancies are incorporated into the crystal in their equilibrium concentrations and then diffuse and recombine rapidly in the high temperature portion of the crystal. Depending on the growth conditions, one of these defect species survives in excess and leads to the formation of the corresponding microdefects, either swirls (self-interstitials) or D-defects (vacancies). This picture has been employed in several recent numerical studies, which are compared and contrasted with the model in this thesis.

2.5.2 Models for Nucleation and Aggregation of Microdefects

The above discussion does not consider mechanisms for the actual formation of A, B, and D defects by aggregation of supersaturated point defects. Once the source and concentration of intrinsic point defects has been determined, the next level of detail needs to address nucleation and aggregation. This issue is even more controversial as there is no experimental data on the structure and composition of very small point

defect clusters. The situation in CZ crystals is even more complicated due to the high oxygen levels which lead to structures such as the OSF-ring that are not observed in FZ crystals. This section briefly summarizes some of the nucleation models that have been proposed.

B-collapse Model

The essence of the B-collapse model is that self-interstitials aggregate into three-dimensional aggregates which become thermodynamically unstable as the crystal cools during growth. Then, depending on their size, they either dissolve, collapse to form the two-dimensional dislocation clusters that are known as A defects, or freeze into the crystal as B-defects. This qualitative picture was supported by several workers (Petroff and de Kock, 1975; Föll *et al.*, 1977). Evidence for the validity of this model is vague due to the invisibility of B-defects to TEM. However, this fact does indicate that B-defects exhibit relatively weak strain fields within the crystal.

The mechanism driving self-interstitial aggregation into B-defects is not clear. Some researchers have suggested that carbon plays a central role here, providing free volume to reduce an aggregate's free energy and thus stabilizes it. Evidence to support this idea comes from experiments that show carbon levels within a crystal have dramatic effects on the distribution of swirls and D-defects, as described in Section 2.3.

The Droplet Model

The concept of droplets was discussed briefly in Section 2.5.1 in the context of point defect sources during crystal growth. Here, the idea is used to make the B-collapse idea quantitatively useful. The B-collapse framework leaves open the question of whether nucleation of B-defects proceeds homogeneously or heterogeneously. The droplet model postulates that a "liquid-like" region forms as self-interstitials become highly supersaturated when the crystal begins to cool. Since the segregation coefficient of carbon is greater than one, carbon preferentially enters the liquid phase which stabilizes it as discussed above. If carbon is required for cluster growth, the B-defects

grow at a rate controlled by the diffusivity of carbon, i.e. with diffusion-limited kinetics (Waite, 1957).

Void Models

Recent experimental evidence for showing that D-defects in CZ crystals (and most likely in FZ crystals) are essentially voids has made obsolete the oxygen precipitation models employed for explaining these defects (Itsumi *et al.*, 1995). However, the same study detected the presence of small amounts of SiO_2 associated with these voids which could imply that a simple homogeneous vacancy nucleation model might not be accurate; this is currently an open question that is of vital technological importance in the area of microdefect engineering. The SiO_2 layer in these voids could provide the nucleation site for void formation, or it could form at the void boundary after homogeneous vacancy aggregation. Both these scenarios are possible and at this time neither can be ruled out.

The literature contains many previous attempts at computing the distribution of point defect aggregates in crystalline materials. Much of this work has been associated with the formation of interstitial-type dislocation loops (Hall, 1980) and voids (Katz and Wiedersich, 1971). These models are generally based on classical nucleation theory which predicts rapid aggregation for the high levels of supersaturation present during irradiation (Yacout and Stubbins, 1986) or thermal annealing (Johnson, 1968). In all cases, a quantitative description of the free energy for the distribution of cluster sizes is required. We pursue this in Chapter 4 using atomistic simulation and in Chapter 6 with a coarser phenomenological picture.

Models for Oxygen-Related Microdefects in CZ crystals

While oxygen precipitation does not appear to lead to observable defects in as-grown crystals, it is still important because of the evidence for the existence of latent nuclei which are invisible to current techniques. The kinetics of oxygen precipitation are determined almost exclusively by two factors: the transport of oxygen within the silicon lattice and the release of stress due to the volumetric mismatch between the

silicon and silicon dioxide phases. Models for oxygen precipitation generally consider one of two possible mechanisms for this; the SiO_2 phase can eject self-interstitials or absorb vacancies to gain the free volume it requires to grow. The self-interstitial mechanism is given by

$$(SiO_2)_i + (1+y)Si + 2O_I \longleftrightarrow (SiO_2)_{i+1} + yI + strain \tag{2.4}$$

The parameter, y, accounts for the fact that the interstitial strain-relief mechanism might not completely alleviate the compressive stress of the oxygen precipitate as it grows. The vacancy mechanism is similar and is given by,

$$(SiO_2)_i + Si + yV + 2O_I \longleftrightarrow (SiO_2)_{i+1} + strain \tag{2.5}$$

The OSF-ring structure discussed in Section 2.2 has been the focus of much attention. Models for the formation of this ring assume that oxygen microprecipitates form preferentially at the radius of the OSF-ring by the mechanisms described above and then act as nuclei for the stacking-faults that form during steam oxidation of wafers (Hasebe *et al.*, 1989). It is well known that the growth of interstitial-type stacking faults proceeds due to the interstitial injection into the crystal bulk across the Si/SiO_2 interface caused by steam oxidation. Less well understood are the defect dynamics that preferentially form an annular region of nuclei for these stacking-faults to grow from. The major emphasis of Chapter 5 is to provide a quantitative model that is capable of predicting the radial position of this region. The problem of describing point defect nucleation in this region is considered in Chapter 6, in an extension of this model.

2.6 Summary and a Quantitative Microscopic Model for Microdefect Formation

This chapter has focused on the multitude of microdefect observations and models in crystals grown by the CZ and FZ methods. There are many different possible explanations for each experimental result and the diversity of previously proposed models clearly shows this. Because these models are not always complementary, a single self-consistent picture for microdefect formation needs to be put together before any computational method can be implemented. This section describes the qualitative aspects of the model developed in this thesis in Chapters 5 and 6.

2.6.1 Intrinsic Point Defect Incorporation at the Melt/Solid Interface

Self-interstitials and vacancies are incorporated across the melt/solid interface in their thermal equilibrium quantities. These levels are not assumed to have particular levels and will be computed using the atomistic simulation techniques described in Chapter 3. Corrections to these estimates are discussed in Chapter 5 and are obtained by comparison to experimental measurements of the Oxidation-Induced Stacking-Fault Ring position as a function of crystal growth operating conditions. A crucial aspect of this approach is that no *a priori* assumptions are made concerning which point defect species is dominant.

While there is no direct evidence for equilibrium defect incorporation at the melt/solid interface, there is also no evidence to the contrary. Typical crystal growth rates are slow enough to give the system ample time to allow equilibrium to be attained. The trapping theories for non-equilibrium incorporation are more applicable in cases where rapid solidification exists, such as for very thin crystals. It will be shown quantitatively that the exact level of point defect incorporation across the interface does not critically affect the conclusions of this thesis.

2.6.2 Point Defect Transport and Reaction

Point defects are known to diffuse rapidly at high temperatures by Fickian diffusion. However, there has been significant controversy concerning quantitative estimates of these rates. In this thesis, atomistic simulation is used to compute first estimates for these rates which are then adjusted based on the same experimental data used to correct the equilibrium concentrations, as described in Chapter 5.

Other point defect transport mechanisms have been proposed. The most common of these is thermodiffusion due to the steep thermal gradients that appear in the crystal near the melt/solid interface during growth (Habu and Tomiura, 1996). There are currently no rigorous methods, either experimental or computational, to compute the magnitude of thermodiffusion of point defects in silicon. This thesis assumes that thermodiffusion is essentially negligible by computing an upper bound on it and then performing a sensitivity analysis on its effects on the predictions of the model in Chapter 5.

The rate at which point defect recombination proceeds is also partially computed with atomistic simulation. This mechanism for reducing supersaturation levels was not considered too carefully in earlier microdefect models but is now known to be very important. Its most important function is to provide a competitive link between the two native point defect species which is necessary to explain several experimental observations.

2.6.3 Microdefect Origins

The goal of this thesis is to provide a solid theoretical foundation for describing the basic mechanisms that lead to the "simplest" microdefects commonly found in *as-grown* crystals prior to processing. The term simple is used to refer to microdefects that we believe are primarily the consequence of homogeneous native point defect condensation. Thus, we assume that A and B defects are due to self-interstitial aggregation while D-defects are due to vacancy condensation. The OSF-ring and neutral zone that appear between these regions in CZ and FZ crystals, respectively,

are assumed to be regions in which intermediate conditions exist.

Oxygen precipitation is not considered to contribute to the formation kinetics of A, B, or D defects in both FZ and CZ crystals. Evidence for this picture comes from the lack of visible oxygen precipitates in as-grown crystals as well as the fact that oxygen precipitation is known to occur at temperatures well below the formation temperatures of the A, B, and D defects (Inoue *et al.*, 1982; Hu, 1981). In this view, the location of defect structures such as the OSF-ring would be determined exclusively by the dynamics of native point defects which is discussed in detail in Chapter 5.

The rates at which point defect clusters form are assumed to be functions of point defect supersaturation only; i.e. aggregation proceeds by homogeneous nucleation. The effects of impurity nucleants are neglected. The resulting defects are assumed to be spherical for all sizes. While these assumption is not likely to be completely accurate, it will lead to quantitative not qualitative errors. Microdefects are still not characterized well enough to justify added complexity in any nucleation model.

The detailed model of Voronkov (Voronkov, 1982) was the first to quantitatively account for the effects of the crystal boundary on the concentration of point defects. Voronkov assumed that point defect equilibrium was instantly achieved at the boundary and this accounted for the observed radial structure. However, other studies imply that point defect annihilation and generation rates at the crystal boundary are finite (Hu, 1994); these rates are treated as adjustable parameters in this thesis and their significance on the conclusions of this model will be discussed in Chapter 5.

None of the above discussions have considered the fact that point defects can exist in multiple charge states within the silicon lattice. The atomistic simulations described in Chapters 3 and 4 are based on an empirical interatomic potential that does not account for electrical charge and this thesis assumes that all defect species are neutral. This major assumption is justified because, as will be demonstrated later, most microdefect formation takes place at very high temperatures where charge states are "smeared out" across the lattice and not a property of an individual point defect. In this limit, it can be assumed that all point defects are neutral.

We now have a basis for developing a quantitative framework to model the forma-

tion of microdefects in a growing crystal. Clearly, it is necessary to make assumptions that cannot as of yet be rigorously justified. However, in most cases, the effect of these assumptions can be probed with an appropriate sensitivity analysis and their importance can be assessed.

Chapter 3

Point Defect Thermophysical Property Estimation

The equilibrium, transport, and kinetic properties of intrinsic point defects in silicon, namely their equilibrium concentrations, diffusivities, and recombination rate play an essential role in determining the quality of single-crystal silicon in many aspects of the fabrication of microelectronic devices, including the formation of microdefects during crystal growth from the melt (Abe, 1985; Brown *et al.*, 1994) and dopant implantation during device processing (Fahey *et al.*, 1989). The high temperature properties are especially important because of the role of thermally activated diffusion and reaction processes in the nucleation and aggregation of precipitates formed during thermal annealing of crystals. This chapter is primarily concerned with estimation of the thermophysical properties of intrinsic point defects in crystalline silicon; the self-interstitial and the vacancy.

There are two fundamental approaches to measuring point defect thermophysical properties; the first is with experiments designed to measure the effect of intrinsic point defects on various observable phenomena. The second approach is based on the use of atomistic simulation to directly compute microscopic properties related to defect dynamics and equilibrium properties. A discussion of the common current experimental techniques used to investigate point defect thermophysical properties is presented in Section 3.1. The computational framework employed in this thesis is

presented in Section 3.2 and is based on the methodology of Maroudas (Maroudas, 1992). The analysis is based on the use of an empirically-derived classical potential due to Stillinger and Weber (1985) .

Our choice of the SW interatomic potential is discussed in the context of other empirical potentials in Section 3.3. Previous analyses (Ungar *et al.*, 1993) have compared the relative strengths and weaknesses of various empirical potentials in detail and we use these studies to justify our choice. A discussion of first-principles or *ab initio* approaches is also presented in this section. These methods are far more detailed in their description of electronic structure but are also very computationally demanding.

Rigorous thermodynamic integration is performed for the computation of relative free energies, which are used to estimate equilibrium concentrations as a function of temperature. Absolute free energies are determined at low temperatures using the Quasi-Harmonic approximation (Lutsko *et al.*, 1988; Maroudas, 1992). Diffusion coefficients are calculated directly at high temperature using the Einstein relationship with a highly optimized molecular dynamics algorithm. The results of these calculations are presented in Section 3.6 and are compared with estimates previously computed by Maroudas (Maroudas, 1992). A model for point defect recombination is presented in Section 3.6.6 based on the results of these simulations.

3.1 Review of Experimental Techniques for Intrinsic Point Defect Thermophysical Property Estimation

No experimental technique currently exists that allows the direct measurement of self-interstitial and vacancy properties. To circumvent this problem, indirect effects known to be related to intrinsic point defects, such as the diffusion of dopant atoms, are measured. Diffusion-reaction models with the point defect properties as adjustable parameters are then used to fit the experimental data (Tan and Gösele, 1985). The

problems involved with such approaches are clear; the important physical and chemical processes governing point defect transport and reaction must be well understood and mathematically characterized, and the experimental data must be accurate and comprehensive enough to allow the simultaneous fitting of several quantities.

There are several commonly used experimental techniques for estimating intrinsic point defect transport and equilibrium properties. The metal gettering, or dopant diffusion, approach (Tan and Gösele, 1985; Zimmermann and Ryssel, 1992) diffuses a dopant species into a silicon wafer and then monitors the diffusion profile as a function of time and temperature. A common example is the diffusion of gold (Zimmermann and Ryssel, 1992; Stolwijk *et al.*, 1986) in silicon. Gold in silicon generally exists in the substitutional state (i.e. occupying silicon lattice sites) but is transported within the lattice in the interstitial state (Tan and Gösele, 1985). Thus, in order for a substitutional gold atom to move through the silicon lattice, it must first transform into the interstitial state in which it is mobile and then redissolve into the substitutional state. There are two mechanisms by which the interstitial to substitutional transformation can proceed. The first is the Frank-Turnbull mechanism (Frank and Turnbull, 1956) which involves vacancies and can be written as

$$Au_i + V \rightleftharpoons Au_s. \tag{3.1}$$

The Frank-Turnbull mechanism requires the consumption of vacancies as gold is diffused into the lattice, possibly creating an undersaturation. The other mechanism is known as the "kick-out" mechanism (Gösele *et al.*, 1980) and is expressed as

$$Au_i \rightleftharpoons Au_s + I. \tag{3.2}$$

In the kick-out mechanism, self-interstitials are created as the gold dissolves into the lattice leading to local supersaturations. In the dilute gold limit, the intrinsic point defect concentrations are independent of the gold concentration and both the kick-out and Frank-Turnbull mechanisms can be represented by the same set of reaction-

diffusion equations; which are written in one dimension as (Hu, 1994)

$$\frac{\partial C_i}{\partial t} = \frac{\partial}{\partial x}\left(D_i \frac{\partial C_i}{\partial x}\right) - k_{ft} C_V^{eq} C_i + k_{ko} C_I^{eq} C_s, \tag{3.3}$$

$$\frac{\partial C_s}{\partial t} = k_{ft} C_V^{eq} C_i - k_{ko} C_I^{eq} C_s, \tag{3.4}$$

where C_i and C_s are the interstitial and substitutional gold concentrations, $C_I^{eq}(T)$ and $C_V^{eq}(T)$ are the self-interstitial and vacancy equilibrium concentrations, which are functions of temperature only; k_{ko} and k_{ft} are the reaction rate constants for the kick-out and Frank-Turnbull reactions, respectively. In the case that the intrinsic point defect concentrations are functions of the gold concentrations, diffusion-reaction equations for the point defects must be considered simultaneously with eqs. (3.3) and (3.4).

The solution of eqs. (3.3) and (3.4) allow the estimation of either one of the intrinsic point defect self-diffusion contributions, $D_I C_I^{eq}(T)$ and $D_V C_V^{eq}(T)$, depending on the assumptions made on the relative importance of the Frank-Turnbull and kick-out mechanisms. An expression for the effective diffusivity for substitutional gold can be defined in terms of these products (Tan and Gösele, 1985) as

$$D_s^{eff} = \frac{D_I C_I^{eq} C_s^{eq}}{C_s^2} + \frac{D_V C_V^{eq}}{C_s^{eq}}, \tag{3.5}$$

which allows for relatively direct interpretation of experimental diffusion data in terms of $D_I C_I^{eq}(T)$ and $D_V C_V^{eq}(T)$. In order to estimate the individual properties, ($C_I^{eq}(T)$, $C_V^{eq}(T)$, $D_I(T)$, $D_V(T)$), further assumptions are needed (Tan and Gösele, 1985) making the estimation of these properties much more difficult and resulting in greater uncertainty. Other metals also are used in gettering experiments, such as zinc used by Bracht et al. (1995), and platinum by Zimmermann and Ryssel (1992).

Surface oxidation of silicon wafers by annealing in steam at high temperature also has been used extensively to estimate point defect properties using the same general methodology as described above. It has been found that surface oxidation leads to the *oxidation-enhanced diffusion* (OED) of (*B*, *Al*, *Ga*, *In*, *P*, *As*), and to the *oxidation-*

retarded diffusion (ORD) of *Sb* (Tan and Gösele, 1985; Frank *et al.*, 1984) within the bulk. At the same time, surface oxidation has been shown to lead to the generation and growth of interstitial-type oxidation-induced stacking-faults (OSF) (Hu, 1974). The morphology of these stacking faults has been well-characterized with transmission electron microscopy (Hirsch *et al.*, 1965). All these phenomena have been linked by the concept of interstitial injection from a Si/SiO_2 interface into the bulk. The SiO_2 layer which forms during steam oxidation requires about 120% more volume per silicon atom than pure silicon resulting in significant stress at the Si/SiO_2 interface. One way to release this stress is through the injection of interstitials from the SiO_2 phase into the silicon phase bulk, increasing free volume at the strained interface. The enhancement or retardation of diffusion of dopants is determined by whether they are transported through the lattice via a vacancy or self-interstitial assisted mechanism. Thus, by measuring the extent of diffusion enhancement or retardation, self-interstitial and vacancy thermophysical properties can be elucidated (Gösele and Tan, 1983). Additionally, the planar structure of the OSF allows for estimation of point defect properties based on the rate of OSF growth. Evidence for the validity of the interstitial injection picture was obtained with thermal nitridation of silicon wafer surfaces, which injects vacancies, rather than interstitials, into the silicon bulk (Mizuo and Higuchi, 1982). In these experiments, the diffusion of phosphorous was observed to be *retarded* while that of antimony was *enhanced*, in agreement with the above theory.

The common disadvantage in all the experimental methods described above is that they all rely on a fitting model to extract several parameters simultaneously. One recent development in experimental measurement of point defect properties has been the use of positron annihilation to measure vacancy formation enthalpies (Dannefaer *et al.*, 1986). Positron annihilation for charge neutral vacancies allows for a relatively direct method for measuring vacancy presence. Positrons injected into the lattice annihilate with lifetimes characteristic to a particular mechanism. Thus, in the case of the neutral vacancy, the formation energy can be measured quite accurately once the correct lifetime peak has been assigned. However, the quantitative interpretation

of positron annihilation data is quite involved and uncertainty in these results also exists.

3.2 Atomistic Simulation Methods

Experimental determination of point defect thermophysical properties in electronic materials is clearly difficult because of the indirect nature of most measurements. An alternative to such measurements is the estimation of the properties using atomistic simulation. The advantage of this approach is that simulation is essentially a direct measurement for a particular defect structure. It is possible to isolate a particular defect and measure directly its contribution to the energy of a system and thus infer its equilibrium concentration. Dynamical simulations track the evolution of atomic coordinates in real time, giving direct information about the mobility of a defect without the need to use an indirect property. Methods for linking the results of atomistic simulation to actual thermophysical properties such as diffusion coefficients and equilibrium concentrations are described in detail in Section 3.5.

Atomistic simulations are designed to give thermodynamic property information based solely on the physics input by a specified interatomic potential. The link between the input (microscopic) and the output (macroscopic) is obtained through statistical mechanics. A system of N spherical particles (or atoms in the case of elemental silicon) exists in $6N$ dimensional phase space, $\Gamma(\mathbf{r}^N, \mathbf{p}^N)$, where $\mathbf{r}$ and $\mathbf{p}$ are the coordinates and momenta of each particle, respectively.

The concept of a statistical ensemble is introduced to represent many microscopically distinct, but macroscopically equivalent states. Averaging over these states leads to values of macroscopic thermodynamic functions. Such averaging is performed by the Monte Carlo (MC) and Molecular Dynamics (MD) techniques; see Allen and Tildesley (1992) for a summary.

In many cases, the required thermodynamic property is the energy of a system in a particular configuration, under a specified set of thermodynamic constraints such as constant temperature, pressure, or volume. The total energy of a system of

N monoatomic particles is given by its Hamiltonian. The classical Hamiltonian is represented as a sum of kinetic and potential energies as

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \frac{1}{2m} \sum_{i=1}^{N} \mathbf{p_i} \cdot \mathbf{p_i} + \Phi(\mathbf{r}^N), \tag{3.6}$$

where $\{\mathbf{p_i}\}$ is the momentum of particle i and Φ is the configurational energy which is given directly by the interatomic potential. Given the Hamiltonian, $\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)$, the ensemble average of the system energy is calculated using the formula (McQuarrie, 1976; Hansen and McDonald, 1986) (McQuarrie, 1976; Hansen and Mcdonald, 1986)

$$< E >_{ens} \equiv \int d\Gamma \rho_{ens} \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N), \tag{3.7}$$

where the integration is carried out over all phase space and ρ_{ens} is the Boltzmann probability density distribution of states within the particular statistical ensemble. The evaluation of the above $6N$-dimensional integral is the reason for employing either the MC or MD methods.

3.2.1 The Molecular Dynamics Approach

Molecular Dynamics (MD) simulation is based upon the solution of Newton's equations motion for the N-particle system described by a specific interatomic potential, which determines the force that each particle experiences (Allen and Tildesley, 1992), and is the method employed in all the atomistic simulations described in this thesis. The location of the system state in phase space entails the solution of $3N$ nonlinear second-order differential equations as

$$\mathbf{f}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2}, \tag{3.8}$$

where

$$\mathbf{f}_i \equiv -\nabla_{\mathbf{r_i}} \Phi(\mathbf{r}^N), \tag{3.9}$$

is the force on the i^{th} particle created by the potential. Here, $\{m_i\}$ and $\{\mathbf{r_i}\}$ are the mass and cartesian coordinates of atom i, $\{\mathbf{f_i}\}$ is the force on atom i exerted by all the other atoms in the system and $\Phi(\mathbf{r}^N)$ is the interatomic potential. Equation (3.9) explicitly shows the nonlinearity of the system of equations represented by eq. (3.8) because $\Phi(\mathbf{r}^N)$ is nonlinear in terms of $\{\mathbf{r_i}\}$. This nonlinearity requires numerical integration for the solution of eqs. (3.8). In our particular case, the fifth-order order predictor-corrector developed by Gear is used (Gear, 1971). Here, the expressions for the *predicted* particle positions, velocities, accelerations, and higher time derivatives are expanded in Taylor series such that

$$\mathbf{r}_p(t+\Delta t) = \mathbf{r}(t) + \Delta t\mathbf{v}(t) + \frac{1}{2}\Delta t^2\mathbf{a}(t) + \cdots, \tag{3.10}$$

$$\mathbf{v}_p(t+\Delta t) = \mathbf{v}(t) + \Delta t\mathbf{a}(t) + \cdots. \tag{3.11}$$

Equations (3.10) and (3.11) give estimates for the position and velocity of each particle at time $t+\Delta t$ based on the position, velocity, acceleration, and higher time derivatives of each particle at time t. These estimates are improved in the *corrector* step where the difference between the accelerations at time t and $t+\Delta t$ is used to correct the functions

$$\mathbf{r}_c(t+\Delta t) = \mathbf{r}_p(t+\Delta t) + c_0\Delta\mathbf{a}(t+\Delta t), \tag{3.12}$$

and

$$\mathbf{v}_c(t+\Delta t) = \mathbf{v}_p(t+\Delta t) + c_1\Delta\mathbf{a}(t+\Delta t), \tag{3.13}$$

where the $\{c_i\}$ are coefficients in the integration algorithm (Gear, 1971).

If the energy at each time step is included in a running average, the time averaged energy is generated as

$$< E >_{ens} = \lim_{t\to\infty} \frac{1}{t}\int_0^t dt'\mathcal{H}(t'), \tag{3.14}$$

which, if the ergodic hypothesis is satisfied, is the same energy given by MC sampling (Allen and Tildesley, 1992).

3.3 Interatomic Potentials for Atomistic Simulation of Crystalline Silicon

The appropriate choice of interatomic potential for the atomistic simulation of crystalline silicon is vital if meaningful results are to be obtained. Interatomic potentials for silicon range from relatively simple classical potentials to sophisticated schemes that solve, to varying degrees of accuracy, for the electronic structure of the system. The choice of potential is based on several considerations; while fully *ab initio* calculations are more accurate, they are computationally demanding, limiting the system size and time scale for any simulation. These limitations motivate the use of empirical potentials especially for the simulation of larger defect clusters which require many atoms and longer simulations times. These difference are discussed below.

3.3.1 Empirical Potentials for Silicon

Empirical potentials are generally constructed by postulating an analytic form for the potential energy as a function of atomic coordinates. Such an approach requires the use of several adjustable parameters which are fitted to known physical properties, such as melting points, lattice parameters, cohesive energies, and diamond structure bulk modulus. According to the classification of Carlsson (1990) , empirical potentials can be grouped into two main entities; cluster potentials and cluster functionals. The former model many-body interactions using separate terms for each bond-order term,

$$U_N = \sum_{i<j}^{N} u_2(\mathbf{r_i}, \mathbf{r_j}) + \sum_{i<j<k}^{N} u_3(\mathbf{r_i}, \mathbf{r_j}, \mathbf{r_k}) + \cdots, \tag{3.15}$$

where $u_2(\mathbf{r}_i, \mathbf{r}_j)$ is the potential energy contribution due to pairwise interactions, $u_3(\mathbf{r_i}, \mathbf{r_j}, \mathbf{r_k})$ is the contribution due to three-way interactions and so on. The interaction terms, $\{u_i\}$, $i = 1, 2, 3, \cdots$ ideally will quickly converge to zero as i increases. Most potentials truncate the series in eq. (3.15) after the the third term because of the computational expense associated with calculating many-body interactions. Examples of three-body potentials for silicon are the Stillinger-Weber potential (Stillinger

and Weber, 1985), and the Biswas-Hamann potential (Biswas and Hamann, 1985).

Cluster functional potentials differ from the above potentials in that they do not explicitly consider a three-body term for the angular forces in tetrahedral silicon, but incorporate these effects into the pair-wise term (Balamane *et al.*, 1992). The most common formalism for doing this for silicon is the Tersoff method (Tersoff, 1988a). In this approach, the pairwise interactions are written as

$$U_N = \sum_{i<j}^{N} f_c(r_{ij})[A_1\phi_1(r_{ij}) - A_2\phi_2(r_{ij})p(\zeta_{ij}], \tag{3.16}$$

where p is a measure of the bond order and is a function of the effective coordination number, ζ_{ij} which is given by

$$\zeta_{ij} = \sum_{k\neq i,j} \psi((r_{ij}, r_{ik})g(\theta_i). \tag{3.17}$$

ψ and g are effectively the three-body terms which favor the tetrahedral geometry of the diamond lattice. Examples of cluster functional potentials are the three Tersoff potentials (Tersoff, 1988a; Tersoff, 1988b) and the Dodson potential (Dodson, 1987) which is closely related to the first Tersoff potential.

Our choice for the study of point defect thermodynamics is the Stillinger-Weber (SW) potential. The motivation for this choice is somewhat arbitrary in that no empirical potential has been proven to be significantly more accurate than the others. However, the SW potential has been shown to predict fairly well the thermodynamic and structural properties of intrinsic point defects (Balamane *et al.*, 1992). This fact coupled with the popularity of the potential which provides ample data for checking our results makes the SW potential attractive. Section 3.3.2 provides a summary of previous calculations performed with various empirical potentials. The results of the SW calculations are compared directly to others and the relative strengths and weaknesses of each potential is discussed.

The Stillinger-Weber Interatomic Potential

The simulations performed in this thesis are based on the empirically-derived Stillinger-Weber (SW) interatomic potential for silicon (Stillinger and Weber, 1985). Extensive work has proven its ability to accurately predict many properties of bulk silicon (Broughton and Li, 1987) . The success of the SW potential is largely due to its anticipation of the low energy tetrahedral geometry of the perfect silicon lattice, accounted for through its three-body component. This term serves to stabilize the diamond structure of bulk silicon, which is not possible by utilizing a two-body interatomic potential. The resulting two- and three-body components take the following form:

$$u_2(r_{ij}) = \epsilon f_2(r_{ij}/\sigma), \tag{3.18}$$

$$u_3(r_{ij}, r_{ik}, r_{jk}) = \epsilon f_3(r_{ij}/\sigma, r_{ik}/\sigma, r_{jk}/\sigma) \tag{3.19}$$

where

$$f_2(r) = \begin{cases} A(Br^{-p} - r^{-q})\exp[(r-a)^{-1}], & r < a, \\ 0, & r \geq a. \end{cases} \tag{3.20}$$

and

$$f_3 = h(r_{ij}, r_{ik}, \theta_{jik}) + h(r_{ji}, r_{jk}, \theta_{ijk}) + h(r_{ki}, r_{kj}, \theta_{ikj}). \tag{3.21}$$

θ_{ijk} represents the angle between the position vectors $\mathbf{r_{ij}}$ and $\mathbf{r_{ik}}$. The constants σ and ϵ specify the length and energy scales, respectively (Stillinger and Weber, 1985). The function h is given by the expression

$$h(r_{ij}, r_{ik}, \theta_{jik}) = \lambda \exp[\gamma(r_{ij} - a)^{-1} + \gamma(r_{ik} - a)^{-1}] \times (\cos\theta_{jik} + 1/3)^2, \tag{3.22}$$

where a is a prescribed cutoff radius. The function h vanishes identically if either $|\mathbf{r_{ij}}| > a$ or $|\mathbf{r_{ik}}| > a$. The various adjustable parameters in the model (see Table 3.1) have been fit to various physical properties of silicon, including the stability of the diamond phase, the melting temperature, and the liquid phase structure factor (Stillinger and Weber, 1985). Lattice sums were performed to determine the relative stability of different crystalline arrangements at low pressure. Parameter sets that predicted

$A = 7.049556$	$B = 0.602225$	$a = 1.80$
$p = 4$	$q = 0$	$\lambda = 21.0$
$\gamma = 1.20$	$\epsilon = 25150\text{K}$	$\sigma = 2.0951\text{Å}$

Table 3.1: SW interatomic potential parameters for silicon

the diamond structure to be most stable were separated from the rest. Next, these parameter sets were used to compute predictions of the thermodynamic melting point and liquid structure data. The liquid phase structure factor is related to the radial distribution factor by

$$S(k) = 1 + \rho \int \exp(i\mathbf{k} \cdot \mathbf{r})[g(r) - 1]d\mathbf{r}. \tag{3.23}$$

The parameter set described in Table 3.1 leads to good qualitative and reasonable quantitative agreement with experimental diffraction data. It is important to note that the SW potential is not fit to any point defect energy or structure data.

Other Classical Potentials for Tetrahedral Silicon

One of the shortcomings of the empirical potentials for silicon described above is their inability to model small clusters of silicon atoms. These structures are useful for evaluating a potentials accuracy with comparison to *ab initio* results. They also test the limits of a potentials ability to model atomic arrangements far away from the regular tetrahedral structure of the diamond lattice. The unique feature of small clusters is their effective low coordination number; a large percentage of available bonds are unfilled. Similar situations arise at surfaces and in the local region near a bulk vacancy. The potential of Chelikowsky et al. (1989) introduced a "dangling bond" vector in which bond strength is transferred from unbonded orbitals to back bonds, creating more compact structures as predicted by electronic structure calculations.

This potential, however, has not been shown to provide additional accuracy in the modeling of point defect thermophysical properties.

The potential of Bolding and Andersen (1990) is one of the most complicated empirical potentials available. It attempts to account for σ and π bonding independently and is a generalization of the Tersoff formulation. The potential considers five-body interactions and requires the fitting of over thirty parameters. The extensive data base of properties to which it is fit includes phase diagram information, static properties of the diamond phase, and cluster energies and geometries. While it is quite accurate in its reproduction of cluster energy maps and surface energies, it has not been shown to do particularly well in modeling bulk defects. This is an example of the local validity of empirical potentials; few are capable of predicting well properties not considered in their parametrization. Furthermore, this particular potential suffers greatly from its complexity; the computational cost of dynamical simulations such as the ones performed in this thesis are a foremost consideration when choosing a potential. The next section discusses quantitatively the performance of several empirical potentials for silicon for a variety of systems, ranging from point defects to surfaces.

3.3.2 Quantitative Comparison of Empirical Interatomic Potentials for Silicon

The most direct method of determining a potential's suitability for a given calculation is to compare its predictions to results from *ab initio* and other empirical potentials; this is a consequence of the lack of accurate experimental results for most crystal properties. Balamane et al. (1992) have studied and compared extensively several empirical, many-body potentials for silicon. A general feature of empirical potentials is that they are parameterized by fitting to a small number of physical properties; the particular choice of properties will determine ultimately the suitability of the potential for a given calculation.

For our purposes, a potential's ability is evaluated mainly by its ability to predict correctly point defect structures and energies. Table 3.2 summarizes some of the

	DFT	PTHT	BH	SW	DOD	T3	C&P
Vacancy	3-4	0.77	2.12	2.82	2.57	3.70	4.70
Split Vacancy	4.19	2.83	2.30	3.36	4.17	3.50	N/A
I_{Tetra}	5-6	0.63	1.56	5.25	3.03	3.45	7.6
I_{Hexag}	4-6	0.84	2.89	6.95	2.61	4.61	9.0
I_{Bond}	4-5	1.92	2.54	5.99	4.39	5.86	N/A

Table 3.2: Comparison of point defect energy predictions for several empirical and LDA calculations at 0 K; DFT = Density Functional theory, Car et al. (1984), Baraff and Schluter (1984) , Bar-Yam and Joannopoulos (1984), Kelly et al. (1986); PTHT = Pearson et al. (1984) ; BH = Biswas and Hamann (1985); SW = Stillinger and Weber (1985); DOD = Dodson (1987); T3 = Tersoff (1988); C & P = Chelikowsky and Phillips (1990) . All values in eV.

results of Balamane et al. (1992) and includes other potentials such as the one developed by Chelikowsky and Phillips (1990). More detailed comparisons of the SW predictions to electronic structure calculations are presented later; here, our goal is to show that the SW potential compares favorably to other empirical alternatives. Table 3.2 clearly shows that the SW and T3 potentials are the best choices for point defect calculations. There is no clear advantage in using either of these two potentials. Balamane et al. (1990) also computed elastic constants and phonon frequencies using these potentials and compared them to experimental data; Table 3.3 is a summary of these results. Again, SW and T3 both do a good job, but the agreement in the case of the SW potential is more significant because T3 was fit to the elastic constants while SW was not. This wide range of applicability of the SW potential

	Experiment	PTHT	BH	SW	DOD	T3
C_{11}	1.67	2.97	2.04	1.62	1.21	1.43
C_{12}	0.65	2.70	1.52	0.82	0.72	0.75
C_{44}	0.81	0.45	0.45	0.60	0.66	0.69
$\nu_{TA}(X)$	4.4	4.5	5.6	6.7		9.0
$\nu_{TO}(X)$	13.9	19.3	14.5	15.9		16.0
$\nu_{LOA}(X)$	12.3	13.8	12.2	13.1		12.0
$\nu_{TO}(\Gamma)$	15.3	18.3	16.0	18.1		16.0

Table 3.3: Comparison with experiment of point defect energy predictions for several empirical potentials; PTHT = Pearson et al. (1984); BH = Biswas and Hamann (1985); SW = Stillinger and Weber (1985); DOD = Dodson (1987); T3 = Tersoff (1988). Elastic constants in Mbar, phonon frequencies in THz.

has made it very popular for atomistic studies of silicon. The shortcomings of the SW potential lie mainly in its ability to model isolated atomic clusters and to some extent surfaces. These structures all exhibit low coordination and a high number of unbonded molecular orbitals (Chelikowsky and Phillips, 1990). Potentials such as the one developed by Chelikowsky and Phillips (1990) are specifically designed to mimic the electronic response in such cases and are thus better than the SW potential here. However, this does not imply that they are also better for bulk point defect studies as demonstrated in Table 3.2.

3.3.3 Electronic Structure Calculations

While empirically derived interatomic potentials are useful because of their relative simplicity, they do not consider the electronic picture explicitly, instead relying on *ad hoc* analytical expressions with adjustable parameters to capture the essential features of electronic bonding. The development of computational frameworks to calculate the electronic band structure around crystal defects originated with the pseudopotential method where the assumption that the entire electronic potential can be replaced by the valence electron potential. The electronic configuration of a silicon atom is $1s^2 2s^2 sp^6 3s^2 3p^2$. In the crystalline phase, these electronic states can be subdivided into those belonging to a core (the filled $1s^2 2s^2 2p^6$ levels) and valence electrons which are involved in bonding. The valence electrons are effectively shielded from the ion and can be assumed to be exist as nearly free electrons. The wave functions of these electrons can be approximated by two component functions; a rapidly oscillating function near the core and a smooth, *pseudo-wave function.* This approximation of the true electronic potential is called a *pseudopotential.*

Before the pseudopotential approximation can be applied in a practical computation, it is necessary to simplify the general expression for the total energy of an electronic/ionic system. The total energy of such a system is written as

$$E_{tot} = E_{e-e} + E_{e-ion} + E_{ion-ion} \tag{3.24}$$

where the first term represents the energy due to electronic self-interaction, the second is the energy due to ion-electron interactions and the third is the energy due to ionic interactions. Equation (3.24) takes into account the Born-Oppenhiemer approximation which separates the ionic and electronic degrees-of-freedom because of their very different masses. Equation (3.24) can be written as

$$E_{tot} = F[n(\mathbf{r})] + \int d^3r V_{ion}(\mathbf{r}) n(\mathbf{r}) + E_{ion-ion} \tag{3.25}$$

where $F[n(\mathbf{r})]$ is a functional of the ground state energy configuration. Kohn and Sham (1965) have shown that it possible to reduce the multiparticle functional, F, to a set of self-consistent single-electron equations so that

$$E_{e-e} = 2\sum_i \int d^3r \psi_i^*(\mathbf{r})(-\frac{1}{2})\nabla^2\psi_i(\mathbf{r}) + \frac{1}{2}\int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{XC}n(\mathbf{r}), \tag{3.26}$$

where ψ is an electronic state and the electronic density is given by

$$n(\mathbf{r}) = 2\sum_i |\psi(\mathbf{r})|^2 \tag{3.27}$$

The first term represents the kinetic energy of non-interacting electrons, the second term is the electrostatic potential and the third is the exchange-correlation energy which is related to electronic spin states. If eq. (3.26) is minimized with respect to the wavefunctions, the Kohn-Sham equations are obtained:

$$\left[-\frac{1}{2}\nabla^2 + V_{ion}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r})\right]\psi_i(\mathbf{r}) = E_i\psi_i(\mathbf{r}) \tag{3.28}$$

In eq. (3.28), V_H is the Hartee potential (the electrostatic potential), and V_{XC} is the exchange-correlation potential. The *ab intio* method consists of solving the Kohn-Sham equations self-consistently so that the resulting wavefunctions give the same electronic density that was used to construct the potential. In this approach, iteration towards self-consistency leaves no free parameters to be fitted to experimental data. The *local density approximation* (LDA) is used to simplify the form of the exchange-

correlation potential. The LDA assumes that the exchange-correlation energy for an electron at position ($\mathbf{r}$) is that of a homogeneous electron gas at the same position.

Another method of solving eq. (3.28) is the so-called *empirical* pseudopotential method (EPM). The (EPM) obtains $V(\mathbf{r})$ by fitting predictions to certain experimental data such as optical reflectivity and band structure data. The fitting is performed by computing the Fourier components of the pseudopotential to generate *pseudopotential form factors* (Yu and Cardona, 1996)

Car and Parinello (1985) further extended the usefulness of electronic density functional theory by incorporating it into fully dynamic, finite temperature simulations based on the molecular dynamics method. Current algorithms for diagonalizing the Kohn-Sham matrices allow for LDA schemes to simulate $O(10^2)$ atoms for upto $O(10^{-12})$ seconds. These limits are just enough to allow for the relaxation of single point defects in materials such as silicon but until significantly faster computers or diagonalization approaches are found, fully *ab intio* methods will remain very computationally expensive.

The Tight-Binding Method

A much cheaper approach to computing electronic band structures is by using the theory of Linear Combination of Molecular Orbitals (LCAO). This method is also known as the Tight-Binding (TB) approximation. The salient features of molecular orbital theory is that whenever two atoms bond, the interaction produces two kinds of new orbital, a *bonding* one and an *antibonding* one. The former is lower in energy than the unbonded orbitals while the latter possess higher energy. These energy differences are determined by the interaction Hamiltonian of the system.

For silicon, there are two kinds of orbitals which participate in bonding; s and p. s orbitals are spherical while p orbitals are directional and have three perpendicular orientations, p_x, p_y, and p_z. Because of orientation, p orbitals on neighboring atoms can interact in two different ways, either forming σ bonds (or antibonds) or π bonds (or antibonds).

The elements of the interaction Hamiltonian correspond to the different kinds

of overlap possible, e.g. an $s-p$ overlap. Each of these elements is known as an *overlap parameter.* For a system such as tetrahedral silicon, which contains only s and p valence electrons, it can be shown that there are only four non-zero overlap parameters (Yu and Cardona, 1996), which are given below:

$$V_{ss\sigma} \equiv \langle s|H|s\rangle \quad V_{sp\sigma} \equiv \langle s|H|p_z\rangle \quad V_{pp\sigma} \equiv \langle p_z|H|p_z\rangle \quad V_{pp\pi} \equiv \langle p_x|H|p_x\rangle \tag{3.29}$$

The computational framework for carrying out a tight-binding calculation begins with the local wave equation for the electron bands on each atom:

$$h_l\phi_{ml}(\mathbf{r}-\mathbf{r}_{jl}) = E_{ml}\phi_{ml}(\mathbf{r}-\mathbf{r}_{jl}), \tag{3.30}$$

where E_{ml} and ϕ_{ml} are the eigenvalues and eigenfunctions of the state m. The index l is the index of atom l in primitive cell j (Yu and Cardona, 1996). The atomic orbitals, ϕ_{ml} are constructed so that wavefunctions centered at different atoms are orthogonal to each other. The total Hamiltonian of the system is approximated as the sum of the non-interacting atomic hamiltonians and an interaction term, H_{int}, so that

$$H = H_0 + H_{int}, \tag{3.31}$$

where

$$H_0 = \sum_{j,l} h_l(\mathbf{r}-\mathbf{r}_{jl}). \tag{3.32}$$

The hamiltonian, H, is a matrix which, when diagonalized, gives the electronic states. The elements of H are combinations of the overlap parameters in eq. (3.30) and once these are determined, the problem can be solved. These parameters are fit to band structure data from *ab initio* calculations.

The results of various tight-binding and LDA calculations are compared to our SW estimates in Section 3.6. These calculations provide a solid base against which the results from less physically realistic potentials can be judged. However, it is important to note, that while in theory, LDA calculations are very accurate, their

practical implementation requires the use of fairly small simulation cells and coarse lattice sampling which can lead to structures that are not fully relaxed and hence artificially large energies. In the case of tight-binding calculations, parameter fitting is still required which always introduces a degree of uncertainty. Ultimately, comparison to available experimental data provides the best test of a potentials predictive capabilities. Unfortunately, in the case of point defect thermophysical properties, direct experimental measurements are not possible and uncertainties are very large, as discussed in Section 3.6.5.

3.3.4 Previous Applications of the SW potential

As mentioned above, the SW potential has been employed in many investigations involving silicon. This section presents a brief summary of these applications in order to establish the surprisingly broad range of this potential's validity. These investigations can be classified into three main categories. The first is the study of defect structures and energies. Defects investigated with the SW potential are certainly not limited to point defects or even intrinsic defects.

Defects in Silicon

Numerous researchers have performed atomistic simulations with the SW potential to compute the energies and structures of defects in silicon. These studies have not been limited to point defects. Bulatov et al. (1997) have computed migration energies for line-dislocation motion by jog and kink mechanisms. The {113} planar interstitial defect has been observed in silicon after electron irradiation and is thought to be an important structure in self-interstitial aggregation in silicon and germanium. The SW potential has been used to investigate the energetics and stability of this defect (Kohyama and Takeda, 1992), and has demonstrated that such structures are very stable with respect to isolated self-interstitials.

The subject of point defect aggregation in silicon is of tremendous importance as discussed in Chapter 2. The SW potential offers a compromise between physical

accuracy and computational efficacy for investigation of aggregation on an atomistic level. A detailed discussion of aggregation is presented in Chapter 4 but previous investigations have already attempted these calculations; Zachariah et al. (1996) , for example, used the SW potential to study the aggregation of silicon nanoclusters. Vacancy aggregation (void formation) is another very important process in silicon processing and previous work in this area with the SW potential includes that of Li and Lowther (1996) who investigated local strain effects around very small vacancy clusters. Such investigations test the limits of the applicability of the SW potential which is somewhat less accurate in systems where atomic coordination numbers deviate significantly from four.

Phase Transitions in Silicon

The amorphization of silicon by irradiation (Diaz de la Rubia *et al.*, 1996) and rapid melt cooling (Park *et al.*, 1996) have been studied extensively with the SW potential. The motivation for understanding the properties of amorphous silicon stems from its usefulness in solar cells; its energy absorption properties in the visible light range are significantly better that those of crystalline silicon (Yu and Cardona, 1996), The SW potential is known to give a reasonable picture of the liquid phase and predicts the melting temperature well by virtue of its parametrization. These facts make it appealing in the study of phase transitions such as rapid solidification and amorphization.

Surfaces and Silicon Growth

Atomistic surface processes are extremely important in growth and processing of silicon crystals. Even though the reliability of the SW potential is questionable in surface geometries because of the large number of dangling bonds, its computational efficiency has led many researchers to use it anyway. The SW potential, however, is capable of ordering surface energies correctly; i.e. $E_{<100>} > E_{<111>}$ (Balamane *et al.*, 1992) and has been shown to be at least qualitatively accurate in its description of the $< 100 >$ surface as compared to *ab initio* calculations. A particularly well studied

process is the epitaxial growth of silicon or other substances on silicon substrates which allows for the manufacture of specialty devices such as lasers and ultrafast photodetectors (Ethier and Lewis, 1992).

Reparameterizations of the SW potential

The growth of heterostructures on silicon (Ethier and Lewis, 1992) requires modifications to the SW potential. Ethier and Lewis (1992) reparameterized the SW potential for Germanium and then applied simple Lorentz-Berthelot mixing rules to determine the $Si-Ge$ interactions. Allen et al. (1997) have recently investigated the diffusion of geramanium atoms on the silicon $< 111 >$ surface and found good agreement with experiment. The SW potential has also been employed in the investigation of impurity defects in silicon crystals which is of obvious importance in growth and processing. Rasband et al. (1996) used a modified SW potential to study boron-intrinsic point defect pairs. In this case, the SW parameters were adjusted by fitting to empirical tight-binding results for boron-silicon interactions.

The above discussion serves to highlight the tremendous range of previous applications of the SW potential, most with at least qualitative success. Some of these applications are far more daring than the study of intrinsic point defects in bulk, crystalline silicon. In general, the further away from the diamond crystalline state a particular system is, the less likely that quantitative data will be obtained with the SW potential; this is true of all empirical potentials that have been fitted to the diamond phase of silicon. Potentials designed for specific purposes such as small clusters, or surfaces, will generally not perform as well in other areas. In conclusion, there is ample evidence that out of the presently available potentials, SW provides a near-optimal compromise between accuracy and computational cost.

3.4 Optimization of the Molecular Dynamics Method

The extremely high computational demands of molecular dynamics simulation have led to the development of several methods for reducing the number of operations

required to compute energies and forces at each time step. The most of common of these are the Verlet neighbor list (Verlet, 1967) and the Cell method (Allen and Tildesley, 1992) which will be briefly discussed in the following sections. These methods are generally useful for simulations of larger systems containing over 200 particles because of the increased overhead due to additional data structures. Section 3.4.2 discusses a new data structure introduced specifically for three-body potentials that further increases computational efficiency. All the above methods are useful only in cases where the interatomic potential is short-ranged; this will be explained in further detail in the following sections. The effectiveness of these schemes will be discussed in terms of how the computational load scales with the number of particles in a simulation.

3.4.1 The Verlet Neighbor List and the Cell Method

The Verlet neighbor list is designed to keep track of every particle's nearest neighbors. In this way, only particles within a certain distance of a particular atom are included in the energy and force calculation. Consider the computational expense required for calculating energies and forces with a three-body potential. In the absence of the Verlet neighbor list, the load scales as N^3, where N is the number of atoms in the simulation. If a neighbor list is created, this load is reduced to NN_n^2, where N_n is the average number of atoms in each particle's neighbor list. However, the actual construction of the neighbor list has to now be included in the total computational expense, which scales as N^2. A key point is that it is not necessary to update the neighbor list for every force calculation, since it is usually constructed to include a "buffer" shell slightly thicker than the interaction range of the potential. The update frequency is therefore a function of the time step and the thickness of this shell, and it is usually about once every ten-twenty time steps for condensed systems.

For a short ranged potential such as SW, the cell method is also very useful. Here, the simulation domain is spatially discretized into smaller units possessing the same geometry as the original system. If the size of these cells is chosen to be slightly larger than the potential interaction distance, then only particles in adjacent cells

can interact. Therefore, during the construction of the neighbor list, only particles in adjacent cells need to be considered. The optimal size for these cells is determined by minimizing the number of particles in each cell. Clearly, for a given density, the finer the cellular subdivision, the fewer particles there are in each cell. However, once the cell size becomes smaller than the pair interaction range, it is possible for two interacting atoms to reside on *non-adjacent* cells. These arguments suggest that the optimal cell size is equal to the interaction range for atomic pairs.

In order to fully take advantage of the cell method an efficient data structure known as the linked list is used (Allen and Tildesley, 1992). Here, two data arrays are created. The first one, IHEAD, has one element for each cell which contains an atom label in that cell. The second array, ILIST, is structured so that the element corresponding to that atom label contains the number of the next atom in that cell, which also serves as the index for the next atom and so forth. The end of this sequence is signified by a zero in the array ILIST. An example of the use IHEAD and ILIST is shown schematically in Figure 3-1 In this example cell 1 contains the atoms numbered

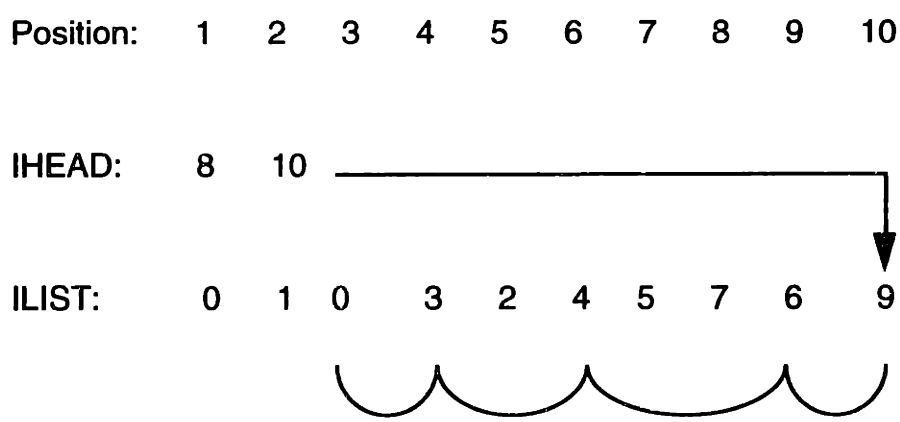

Figure 3-1: Extraction of Cell content using the linked list data structure.

1, 2, 5, 7, 8 and cell 2 contains the atoms numbered 3, 4, 6, 9, 10.

The number of operations required to construct the Verlet neighbor list was previously shown scale as N^2; in three-dimensional space, the cell method reduces this

to $27N_c$ where N_c is the number of particles in each cell. We can now compare the computational requirements of a brute force calculation as opposed to one employing the Verlet neighbor list and the cell method by considering the ratio

$$\frac{\tau_{opt}}{\tau_b} = \frac{\frac{27N_c}{F} + N_n + N_n^2}{N^2}, \tag{3.33}$$

where τ_{opt} and τ_b are the operation counts for the optimized and brute force calculation, respectively, and F is the number of time steps between each neighbor list update. The first term in the numerator of eq. (3.33) is associated with neighbor list construction, the second and third terms correspond to the potential and force calculations for the two and three body contributions, respectively. For a typical molecular dynamics simulation of crystalline silicon at zero pressure with the SW potential, $F \sim 20$, $N_n \sim 20$, and $N_c \sim 10$, eq. (3.33) becomes

$$\frac{\tau_{opt}}{\tau_b} \sim \frac{400}{N^2}. \tag{3.34}$$

Thus for systems containing more than 20 atoms, the overhead is justified and the simulation cost decreases as N^2 when compared to the brute force approach.

3.4.2 The Triplet List

Equation (3.33) shows that the bulk of the cost associated with MD simulation with the SW potential is incurred in the force and potential calculation due to the three-body terms. A new data structure has been created to address this problem. In addition to the neighbor list, a *triplet* list is created and updated every F time steps. The algorithm uses the previously created Verlet list to reduce the search space.

The cost of creating this list scales as NN_n^2 but reduces the triplet force calculation from NN_n^2 to NN_t, where N_t is the number of triplets associated with each particle. For SW silicon at typical simulation conditions, N_t is about $N_n^2/4$, which implies a 400 % speedup in code execution. A fully optimized implementation of the molecular dynamics method employing the SW potential executes about 2700 time steps per

second per particle. For systems containing $O(10^2)$ atoms, $O(10^6)$ time steps, or several nanoseconds of real time are feasible. This has allowed the direct simulation of solid-state point defect diffusion processes, as is discussed in Section 3.5.3.

3.5 Statistical Mechanics of Property Estimation

Depending on the statistical ensemble used in a simulation, different thermodynamic functions can be obtained, and, in theory, while it is possible to obtain free energies directly from a molecular dynamics simulation, this is not a practical approach. In order to accurately compute the entropy of an atomic system by direct MD simulation, it is necessary to sample the system in all its possible states, which is practically impossible as some of these states are extremely unlikely. Thus, the computation of free energies for equilibrium property investigation is performed indirectly with other approaches. These will be described in the next sections. The present methodology is based on work of Maroudas (Maroudas, 1992). Figure 3-2 summarizes the main features of this framework.

3.5.1 Formation Properties and Equilibrium Concentrations

Statistical thermodynamics dictates that a finite concentration of intrinsic point defects must exist in a lattice at finite temperature (Girifalco, 1973). The temperature dependent expression for the equilibrium concentration is derived by minimizing the lattice free energy with respect to the number of point defects present. The free energy of a lattice containing n point defects in N lattice sites at finite temperature is written as

$$G = G_0 + n(G^f) - kTln\frac{N!}{(N-n)!n!}, \tag{3.35}$$

where H^f and S^f are the energy and entropy of formation of point defects, G_0 is the perfect crystal free energy, N the number of lattice sites, and the last term represents the configurational entropy. The formation properties of point defects are defined as the difference in the energies (entropies) of the relaxed defected lattices and the

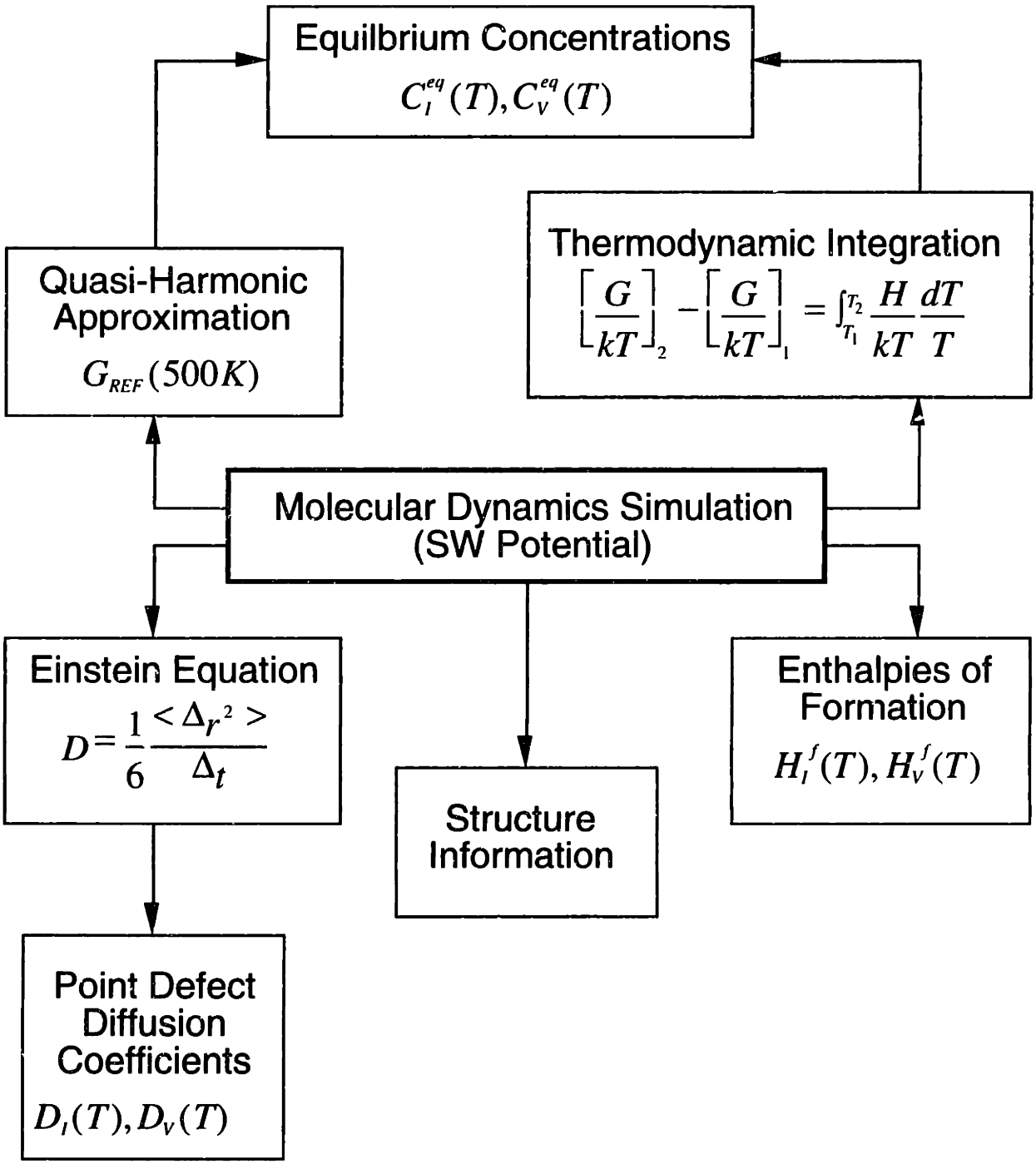

Figure 3-2: Computational Framework for investigating intrinsic point defect thermophysical properties.

energy of the relaxed perfect lattice scaled to the same number of atoms (Benedek *et al.*, 1992). For a given property, $(X = H, G, F)$, the formation quantity is given by

$$X^f = X(N-1) - (N-1)[X(N)/N], \tag{3.36}$$

for a vacancy, and

$$X^f = X(N+1) - (N-1)[X(N)/N], \tag{3.37}$$

for a self-interstitial.

Differentiation of eq. (3.35) with respect to n gives (McQuarrie, 1976)

$$n^{eq} = N \exp\left(-\frac{G^f}{kT}\right). \tag{3.38}$$

3.5.2 Free Energy Calculations

Equation (3.38) shows that in order to compute equilibrium concentrations for the point defects, it is necessary to estimate absolute free energies of formation for these species. Free energies are generally not computed directly from MD or MC simulations because entropy-related quantities are directly related to the partition function of a system which represents *all* states. Many of these states are rarely observed in a simulation and so a direct averaging method will lead to very poor estimates for the free energy.

The usual approach for computing free energies is to compute a single absolute free energy at some reference temperature and then extrapolate this value to higher temperatures by computing free energy *differences* (Maroudas, 1992; Ungar *et al.*, 1993; Sinno *et al.*, 1995). The reason for this approach is that absolute free energy computations are usually very computationally expensive and cannot be justified for every temperature of interest. Furthermore, some of these methods (Lutsko *et al.*, 1988) are based on idealized models which are quite accurate at low temperatures but become increasingly uncertain at higher temperatures. For example, Maroudas (1992) computed absolute free energies of formation using a synthetic method termed

the Quasi-Harmonic Approximation (QHA), which is based on harmonic crystal theory (Girifalco, 1973; Ashcroft and Mermin, 1979) with non-harmonic corrections provided by atomistic simulation (Lutsko *et al.*, 1988). This is the approach used in this thesis and a theoretical description is given below.

The Quasi-Harmonic Approximation

If the lattice potential energy function is expanded in a Taylor series about an equilibrium state, only quadratic terms need to be retained if atomic motion is not too large, i.e. the temperature is sufficiently low. In fact, the linear term vanishes identically because the forces in the system cancel when the particles are in their equilibrium positions. Therefore

$$\Phi_{total} = \Phi_0 + \Phi_{harmonic}, \tag{3.39}$$

where Φ_0 is the potential energy of the system at equilibrium and

$$\Phi_{harmonic} \equiv \frac{1}{2}\mathbf{u}^T \cdot \mathbf{D} \cdot \mathbf{u}. \tag{3.40}$$

In eq. (3.40), $\mathbf{u}$ is the $3N$-dimensional displacement vector, and $\mathbf{D}$ is the *Dynamical Matrix* (Ashcroft and Mermin, 1979), which is defined as

$$D_{i,j,\alpha,\beta} \equiv \left(\frac{\partial^2 \Phi}{\partial r_{i,\alpha} \partial r_{j,\beta}}\right)_{r_0^N}, \tag{3.41}$$

where (i,j) are particle indices and (α,β) are coordinate indices. Thus, within the harmonic approximation, the free energy is approximated as

$$F_{harmonic} = \Phi_0 + \frac{kT}{2} \ln\left(\frac{\Lambda^2}{2\pi kT} \times \det(\mathbf{D})\right), \tag{3.42}$$

where Λ is the thermal wavelength defined as (McQuarrie, 1976)

$$\Lambda \equiv \left(\frac{h^2}{2\pi mkT}\right)^{1/2}, \tag{3.43}$$

and h is Planck's constant. Anharmonic effects are accounted for by the physics incorporated in the interatomic potential which is used to compute the density at which the dynamical matrix elements are evaluated.

Clearly, the harmonic approximation is not a good assumption at higher temperatures, where atomic vibrations are large enough to be significantly anharmonic. It has been shown that the QHA is acceptable for temperatures upto 75% of the melting temperature in perfect crystal systems (Lutsko *et al.*, 1988). Ungar et al. (1994) have criticized the use of the QHA for computing absolute free energies of defected lattices even at low temperature because the crystal regions near defects might exhibit locally significant anharmonicities. An alternative to the QHA which is more accurate but also more computationally demanding is an integration scheme, originally proposed by Frenkel and Ladd (1984). Consider the isobaric-isothermal partition function

$$Q_{NPT}(\lambda) = \sum_{\Gamma} \sum_{V} \exp\left(-\frac{U(\lambda) + PV}{kT}\right), \tag{3.44}$$

which is related to the Gibbs free energy of a system by

$$G(\lambda) = -kT \ln Q_{NPT}(\lambda), \tag{3.45}$$

where λ is a variational parameter and will be defined below. If eq. (3.44) is differentiated with respect to λ, we get a quantity which can be obtained directly from simulation:

$$\frac{\partial G}{\partial \lambda} = -kT \frac{\partial \ln Q_{NPT}}{\partial \lambda} = \left\langle \frac{\partial U}{\partial \lambda} \right\rangle. \tag{3.46}$$

Integration of eq. (3.46) gives

$$G(\lambda = 1) - G(\lambda = 0) = \int_0^1 \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{\lambda'} d\lambda'. \tag{3.47}$$

The parameter λ is used to concoct a varying potential that changes smoothly from the desired one (e.g. the SW potential) to a simple one for which the free energy can be computed analytically, such as a harmonic approximation. This potential can be

written as

$$U(\lambda) = (1 - \lambda)U_{harm} + \lambda U_{full}. \tag{3.48}$$

The derivative of the free energy with respect to λ is given by

$$\left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{\lambda'} = \langle U_{full} - U_{harm} \rangle . \tag{3.49}$$

The integral in eq. (3.47) is evaluated numerically by performing simulations at closely spaced increments in λ, and since $G(\lambda = 0)$ is known analytically, an absolute free energy can be computed.

The so-called λ-integration method described above is computationally very expensive in comparison to the QHA, especially if different defect configurations need to be investigated. However, it is not possible to quantitatively estimate the error incurred in using the QHA without direct comparison to the thermodynamic integration method. We justify the use of the QHA at low temperatures for point defect investigations by noting that atomic motion in the vicinity of point defects is the same as motion in perfect parts of the crystal. In other words, thermal vibrations of atoms are unaffected by the presence of a defect and no long range motion of the defect (or neighboring atoms) is observed except for diffusion which occurs extremely rarely at lower temperatures. Figure 3-3 shows the atomic trajectories of two atoms which lie in defected and perfect regions, respectively. The trajectory of the atom in the perfect crystal is distributed isotropically about its equilibrium position. In the case of the atoms near a vacancy, we see that although an anisotropy is introduced, the trajectories are still essentially centered closely about a well-defined equilibrium position.

Relative Free Energy Calculations at High Temperature

Once a reference free energy is computed at low temperature, it is sufficient to calculate *relative* free energies for all other temperatures. Maroudas (1992) used a cumulant expansion (Maroudas, 1992) to extrapolate the free energy estimates to temperatures upto the melting point referenced by an estimate for the absolute free energy at 500 K

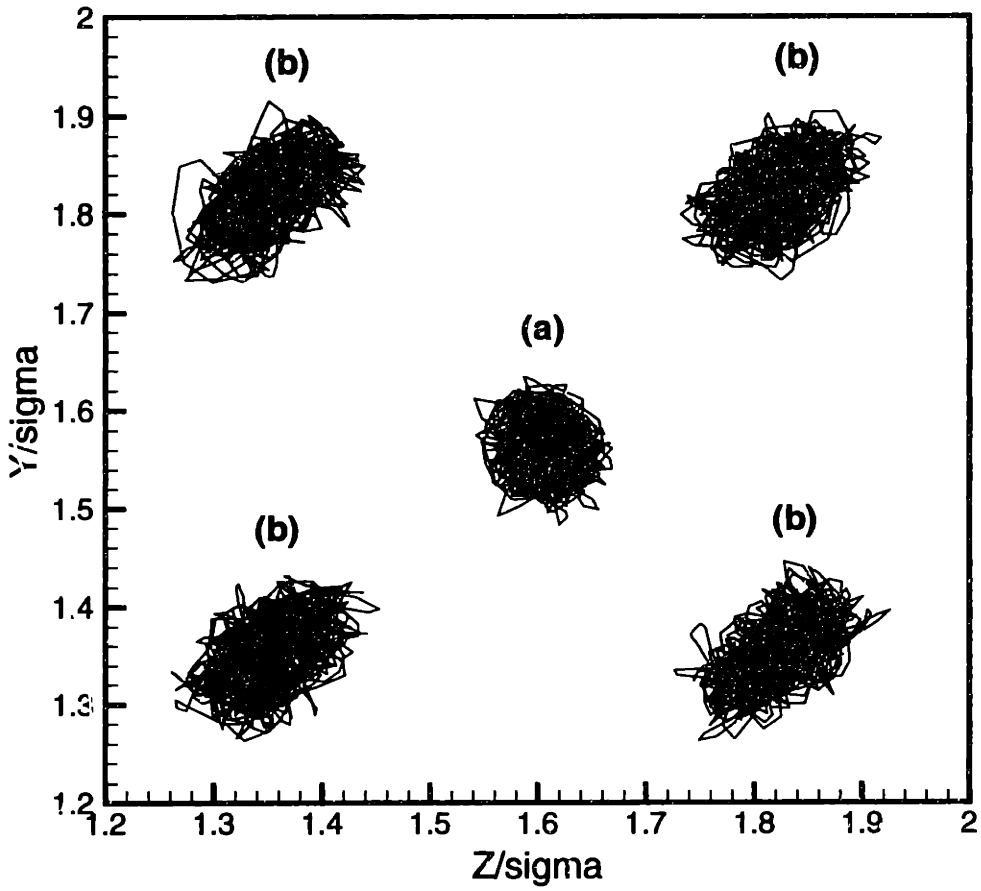

Figure 3-3: Atomic trajectories projected onto the yz plane; (a) atom in a perfect crystal, (b) nearest-neighbor atoms of an equilibrated vacancy. $T = 500K$.

using the QHA. The difference in the quantity G/kT at two temperatures is given by

$$\Delta(\beta G) \equiv \beta G - \beta_{ref} G_{ref} = -\ln\langle \exp(\beta_{ref} - \beta) H \rangle_{\beta_{ref}}, \tag{3.50}$$

where $\beta \equiv 1/kT$ and the subscript ref represents some reference state. The right hand side of eq. (3.50) can be expressed by a series of cumulants (Maroudas, 1992):

$$\langle \exp(\beta H) \rangle \equiv \exp\left(\sum_{n=1}^{\infty} \frac{\beta^n}{n!} C_n(N, p, \beta) \right), \tag{3.51}$$

where C_n is the n^{th} cumulant of the enthalpy distribution function. If eq. (3.51) is substituted into eq. (3.50) we obtain

$$\Delta(\beta G) = \sum_{n=1}^{\infty} \frac{(\beta - \beta_{ref})^n}{n!} (-1)^{n+1} C_n, \tag{3.52}$$

where

$$C_1 \equiv \langle H \rangle, \tag{3.53}$$

$$C_n \equiv \langle (H - \langle H \rangle)^n \rangle; \qquad n > 1. \tag{3.54}$$

The major advantage of this method for computing relative free energies is that a single MD simulation is sufficient to estimate the temperature dependence of any thermodynamic property (Maroudas, 1992; Phillpot and Rickman, 1991). Here, however, we use rigorous thermodynamic integration to compute relative free energies at the higher temperatures (Allen and Tildesley, 1992). In the isothermal-isobaric statistical ensemble, the Gibbs free energy difference for a system at two temperatures, T_1 and T_2 is given by (Allen and Tildesley, 1992)

$$\left(\frac{G}{NkT}\right)_2 - \left(\frac{G}{NkT}\right)_1 = \int_{T1}^{T2} \left(\frac{H}{NkT}\right) \frac{dT}{T}. \tag{3.55}$$

Thus, for a given reference Gibbs free energy, absolute values for the free energy are computed at any temperature. Clearly, this method is computationally expensive because in order to obtain a reasonable numerical estimate for the integral in eq. (3.55)

it is necessary to perform a full MD simulation every 50-100 K from the reference temperature (500 K in our case) to near the melting point (1600 K). The results presented in Section 3.6.4 will show that the added computational expense is justified at higher temperatures due to delocalization effects.

3.5.3 Diffusion Coefficients

Point defect diffusion is modeled as a thermally activated jump process, and can be expressed by rate theory in terms of a jump rate which is given by (Vineyard, 1957)

$$\omega = \nu \exp\left(-\frac{H^m}{kT}\right), \tag{3.56}$$

where ν is an attempt frequency and H^m is an activation energy barrier for migration which represents the energy difference of the system between the equilibrium and saddle-point configurations (Maroudas, 1992). Applying the harmonic crystal approximation to the system, the attempt frequency can be written as

$$\nu \equiv \frac{\prod_{i=1}^{3N} \nu_i}{\prod_{i=1}^{3N-1} \nu_i^*}, \tag{3.57}$$

where the $\{\nu_i\}$ are the normal mode frequencies and $(*)$ denotes the saddle-point configuration. The right-hand-side of eq. (3.57) is evaluated as the determinant of the dynamical matrix of the system in equilibrium and at the saddle-point configuration (see eq. (3.41)). The isotropic diffusion coefficient in a cubic crystal is related to the jump frequency by

$$D = \frac{1}{6} d_0^2 \omega, \tag{3.58}$$

where d_0 is the jump distance. Combining eqs. (3.58) and (3.56) gives (Maroudas, 1992)

$$D = \frac{1}{6} d_0^2 \nu \exp\left(-\frac{H^m}{kT}\right). \tag{3.59}$$

Maroudas (1992) computed temperature dependent point defect diffusion coefficients with eq. (3.59) and constrained path atomistic simulation. In this approach,

a diffusion path for the point defect is postulated and a series of simulations are performed, constraining the position of the defect at various points along this path. For each simulation, the energy is computed and the activation barrier is constructed incrementally. In addition, at the saddle-point (highest energy configuration), the dynamical matrix is evaluated so that the jump frequency can be estimated. This approach is effective if the diffusion mechanism is precisely known in terms of atomic displacements, otherwise many different paths have to be tested before the lowest barrier can be found. The results presented below in Section 3.6 show that in the case of the vacancy, the diffusion path is very simple and the above methodology provides a good estimate for the diffusion coefficient. Maroudas (1992) computed the energy barrier for vacancy diffusion by considering a path that takes a neighboring lattice atom directly into the vacant site. The saddle-point corresponds to the point where the moving atom is centered between two lattice sites and two "half-vacancies" exist on either side; this configuration is known as the split-vacancy. Results for the self-interstitial are more difficult to interpret in terms of a mechanism.

A more direct approach, and the method used in the present work, is the use of the Einstein equation which relates the mean square displacement, $< \Delta \mathbf{r}^2 >$ of the migrating particle to the diffusion coefficient

$$D = \frac{1}{6} \frac{< \Delta \mathbf{r}^2 >}{\Delta t}. \tag{3.60}$$

The limiting factor in this approach is the computational expense which becomes prohibitive at lower temperatures where point defect diffusion is very slow. The use of an Arrhenius expression such as the one in eq. (3.56) allows for the extrapolation of high temperature results to lower temperatures. Estimates for point defect diffusion coefficients computed with eqs. (3.60) and (3.59) are compared in Section 3.6.

3.6 Results

3.6.1 Simulation Parameters and Conditions

Results are presented for the application of the SW molecular dynamics simulations to the calculation of equilibrium concentrations and diffusion coefficients of self-interstitials and vacancies. All the simulations are based on supercells initially containing 216 silicon atoms arranged in a perfect tetrahedral lattice. This has been shown to be sufficiently large to make finite size effects unimportant (Maroudas, 1992). Vacancies are created by initially removing a single atom from the lattice, while self-interstitials are created by placing an additional atom in a tetrahedral interstitial position (Maroudas, 1992). Three-dimensional periodic boundary conditions were applied in all calculations.

The free energy calculations are performed in the isothermal-isobaric statistical ensemble (or constant temperature, pressure, and number of particles) at zero pressure. Constant pressure is attained using the piston method of Andersen (Andersen, 1980). The thermodynamic integration is carried out by performing multiple simulations at each temperature to minimize the statistical uncertainty in the enthalpies of formation. A minimum of six runs are performed at each temperature, ranging from 500 K to 1600 K in increments of 100 K. Each run is allowed to equilibrate for about 20,000 time steps and then energy averaging is performed for an additional 30,000 time steps. The uncertainty in the computed enthalpies of formation is given by

$$\sigma_{tot} = \sqrt{\frac{(\sigma_p^2 + \sigma_d^2)}{N}}, \tag{3.61}$$

where σ_p and σ_d are the uncertainties in the perfect and defected crystal enthalpies respectively. For each temperature, enough runs are performed to ensure that $\sigma_{tot} \leq kT/2$.

The diffusion calculations were performed in the canonical statistical ensemble (constant temperature, volume, and number of particles). The simulations were allowed to run for two million time steps to ensure a sufficient number of atomic jumps.

Simulations were performed at (1200 K, 1300 K, 1400 K, 1600 K) for both the self-interstitial and vacancy. The four data points in each case were fitted to Arrhenius expressions to determine constant preexponential factors and activation energy barriers.

3.6.2 Estimates for Point Defect Diffusivities

The results for the point defect thermophysical properties obtained with calculations with the SW potential have been fitted to Arrhenius-type expressions. In the case of the diffusion coefficients, both preexponential factors and activation energies are assumed to be constants during the fitting procedure. The total mean-square displacement as a function of time for the self-interstitial and vacancy runs are shown in Figure 3-4. In each case, the mean-square displacement is fitted well by a straight line, although the self-interstitial diffusivity at 1200 K is already low enough to make the individual jumps visible. The SW potential clearly predicts a more mobile vacancy at all the temperatures for which simulations were performed; this is also discernible by the relative smoothness of the vacancy mean-square displacements. The fitting of the diffusion coefficients to Arrhenius-type expressions is demonstrated in Figures 3-5 and 3-6. The resulting expressions for point defect diffusion coefficients as a function of temperature are given below as

$$D_I(T) = 1.76 \times 10^{-2} \exp\left(-\frac{0.94}{kT}\right) \quad cm^2/s \tag{3.62}$$

$$D_V(T) = 1.70 \times 10^{-3} \exp\left(-\frac{0.46}{kT}\right) \quad cm^2/s \tag{3.63}$$

The vacancy activation energy for migration, (0.46 eV) is extremely close to the value computed by Maroudas (1992) using constrained-path energy minimization (0.43 eV). This result confirms the assumption that vacancies diffuse by passing through the split-vacancy transition state. A plot of atomic trajectories projected onto the yz-plane near a vacancy is shown in Figure 3-7 and demonstrates the direct path of vacancy diffusion. The bulk of the atomic trajectories are simply vibrational dis-

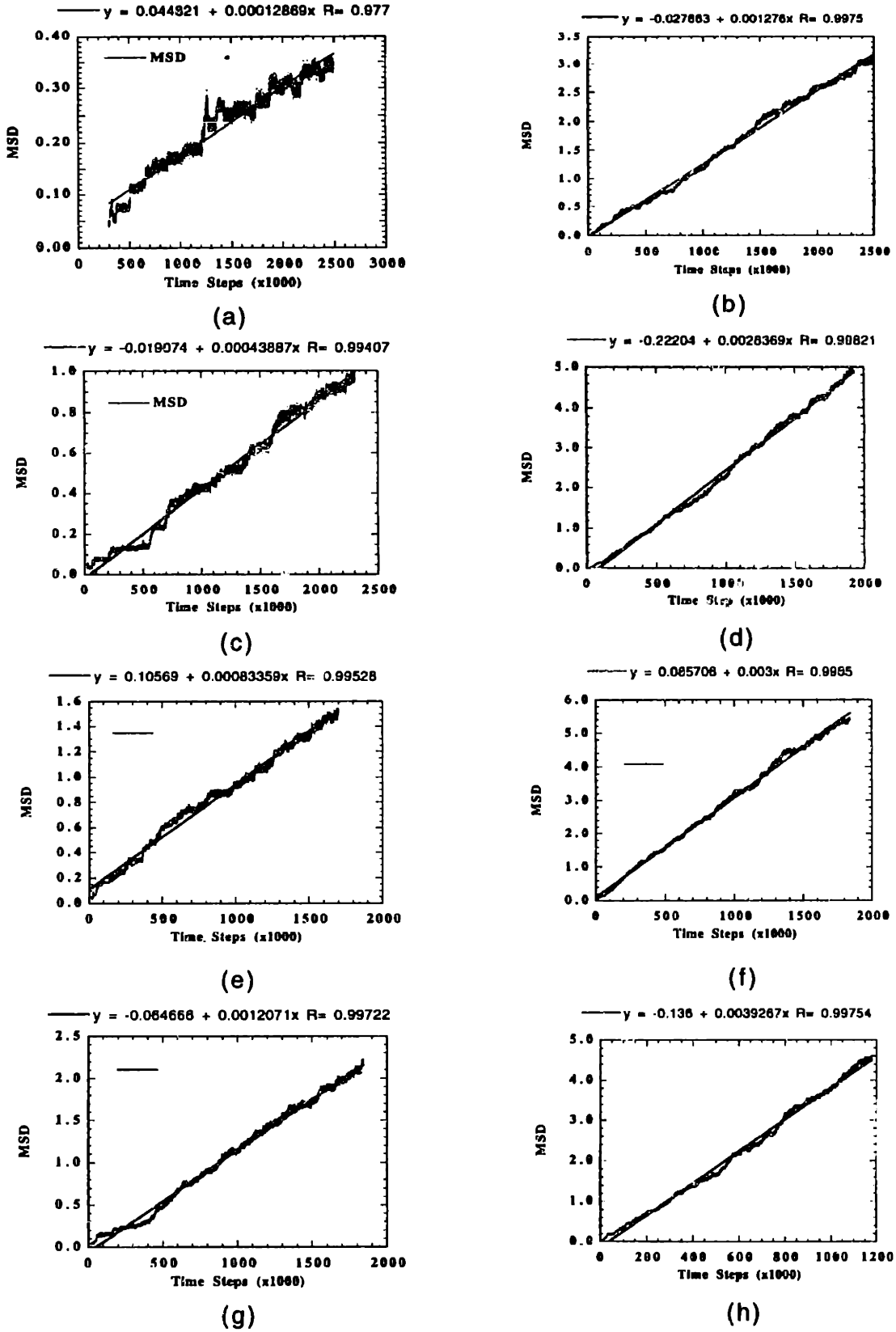

Figure 3-4: Mean-square displacement as a function of time for; (a) I at 1200 K, (b) V at 1200 K, (c) I at 1400 K, (d) V at 1400 K, (e) I at 1500 K, (f) V at 1500 K, (g) I at 1600 K, (h) V at 1600 K

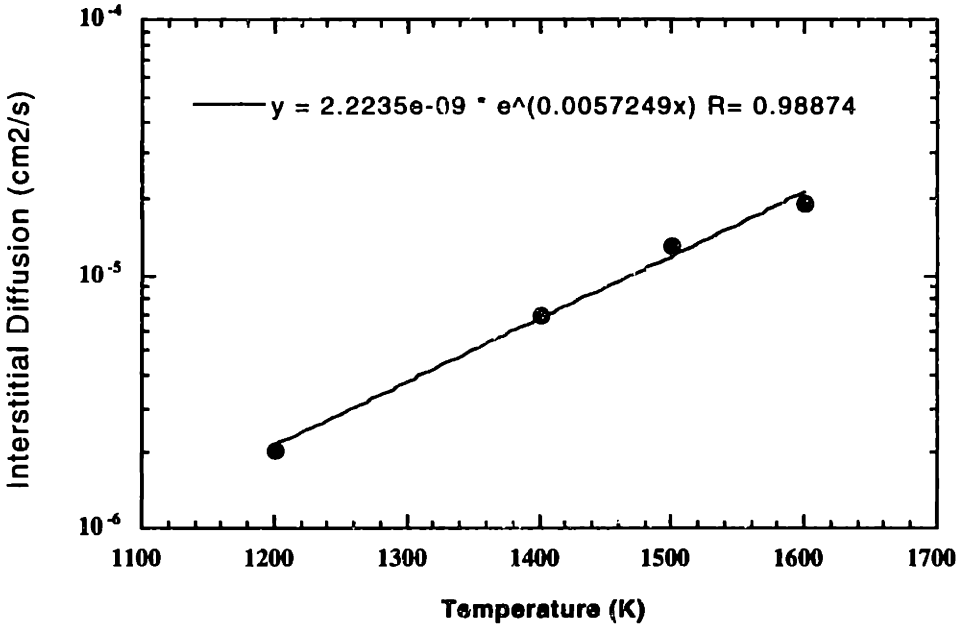

Figure 3-5: Interstitial diffusion coefficient as a function of temperature.

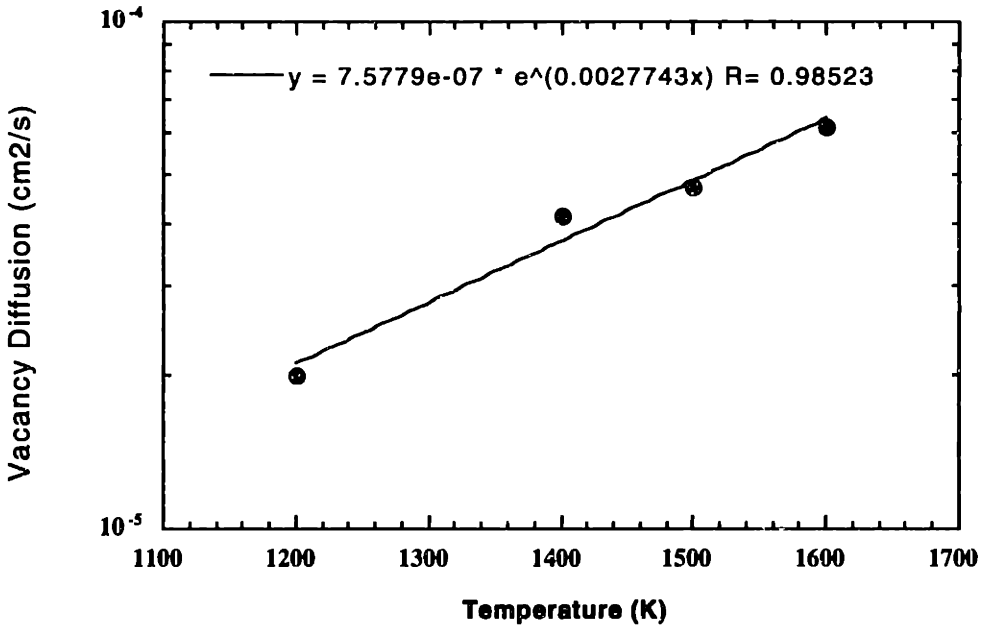

Figure 3-6: Vacancy Diffusion Coefficient as a function of temperature

placements about an equilibrium lattice position. Once an atom is located suitably far from this equilibrium position and possesses the necessary kinetic energy to overcome the activation barrier, it moves almost linearly from one lattice position to the next and resumes its oscillatory motion about its new equilibrium position.

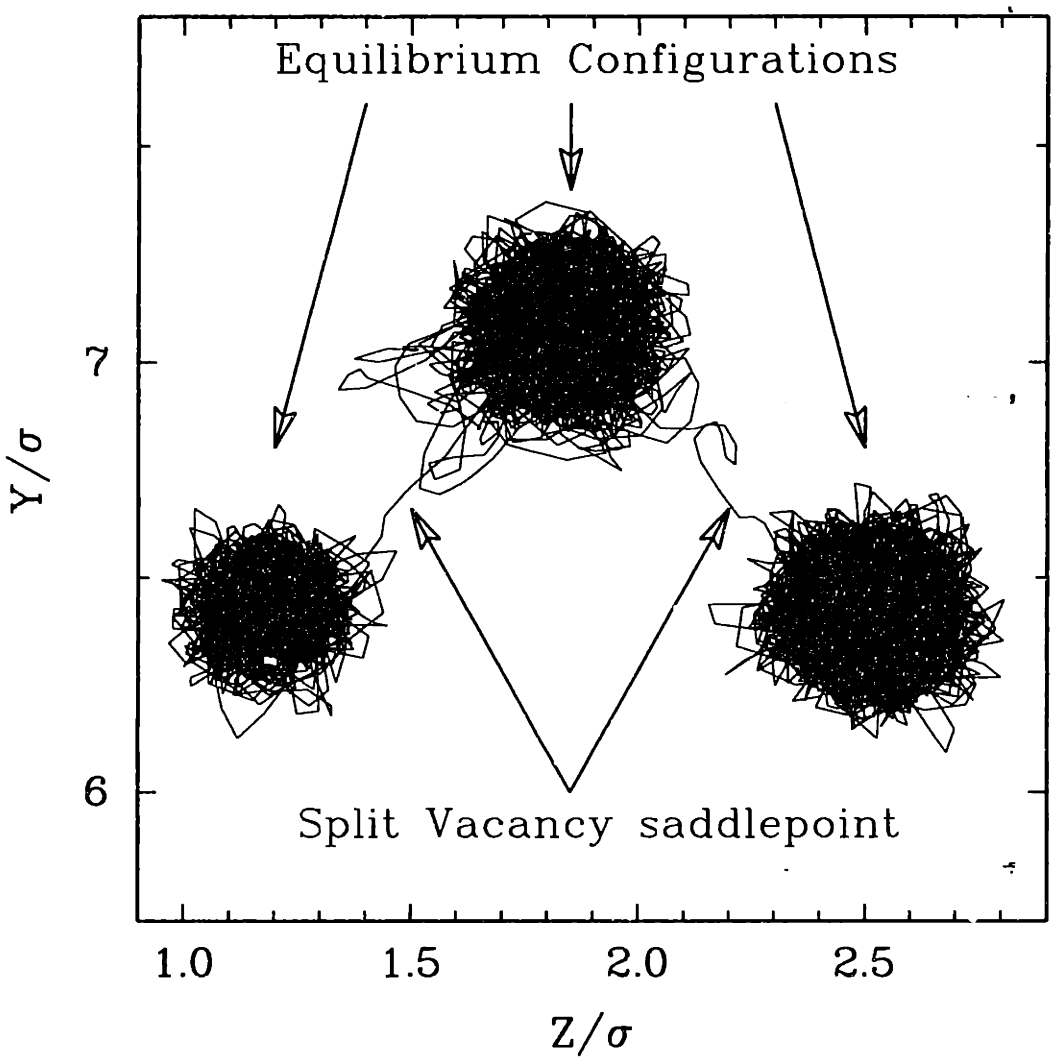

Figure 3-7: Atomic trajectories in the vicinity of a vacancy projected onto the yz plane during a molecular dynamics simulation; σ is the SW length scale.

Maroudas (1992) computed the diffusion energy barrier for the tetrahedral-hexagonal self-interstitial transformation as 1.74 eV. Since the tetrahedral interstitial is not stable with respect to the < 110 >-dumbbell, we cannot use the Einstein relationship to compute this barrier. The constrained-path minimization method is not practical for computing the diffusion barrier for the < 110 > dumbbell because the atomic re-

arrangements involved are complex and it is difficult map out a concerted path along which we can compute energies. To overcome these difficulties, Maroudas (1992) assumed that the transition state for the diffusion path of the $< 110 >$ dumbbell would be the tetrahedral interstitial, and that the resulting barrier would simply be the difference between their respective formation energies, 1.2 eV. The present computations estimate the migration barrier to be 0.94 eV which shows that the above argument is reasonable. In fact, other molecular dynamics studies using the SW potential have shown that the interstitial migrates preferentially along any of the six equivalent $< 110 >$ directions (Gilmer *et al.*, 1995).

Estimates in the literature for the diffusivities of point defects in silicon span several orders-of-magnitude at any given temperature. In general, diffusivities computed by atomistic simulation tend to be much higher than those obtained by fitting to experimental data. The SW estimates are compared to experimental values in Figures 3-8 and 3-9 to experimental values at high temperatures ranging from 1000 K to the melting point. For interstitials, estimates range from $1 \times 10^{-3}\ cm^2/s$ (Bracht *et al.*, 1995) to $1 \times 10^{-6}\ cm^2/s$ (Wijaranakula, 1990) at the melting temperature. The SW estimate of $1 \times 10^{-4}\ cm^2/s$ lies comfortably within this range and is in good agreement with the result of Zimmermann and Ryssel (1992).

In the case of the vacancy diffusion coefficient, the SW calculations predicts $1 \times 10^{-4}\ cm^2/s$ which is much higher than most experimental estimates; other estimates are at least one to two orders-of-magnitude smaller at T_m. Interestingly, the LDA (Blöchl *et al.*, 1993) predicts values closer to the SW result or about $1 \times 10^{-4}\ cm^2/s$ for $T \sim 1500\ K$.

3.6.3 Point Defect Equilibrium Structures and Formation Energies

The equilibrium structure computed for the self-interstitial is shown in Figure 3-10 as the $< 110 >$-oriented split-dumbbell. This structure has consistently been found to be the lowest energy state for the self-interstitial predicted with the SW poten-

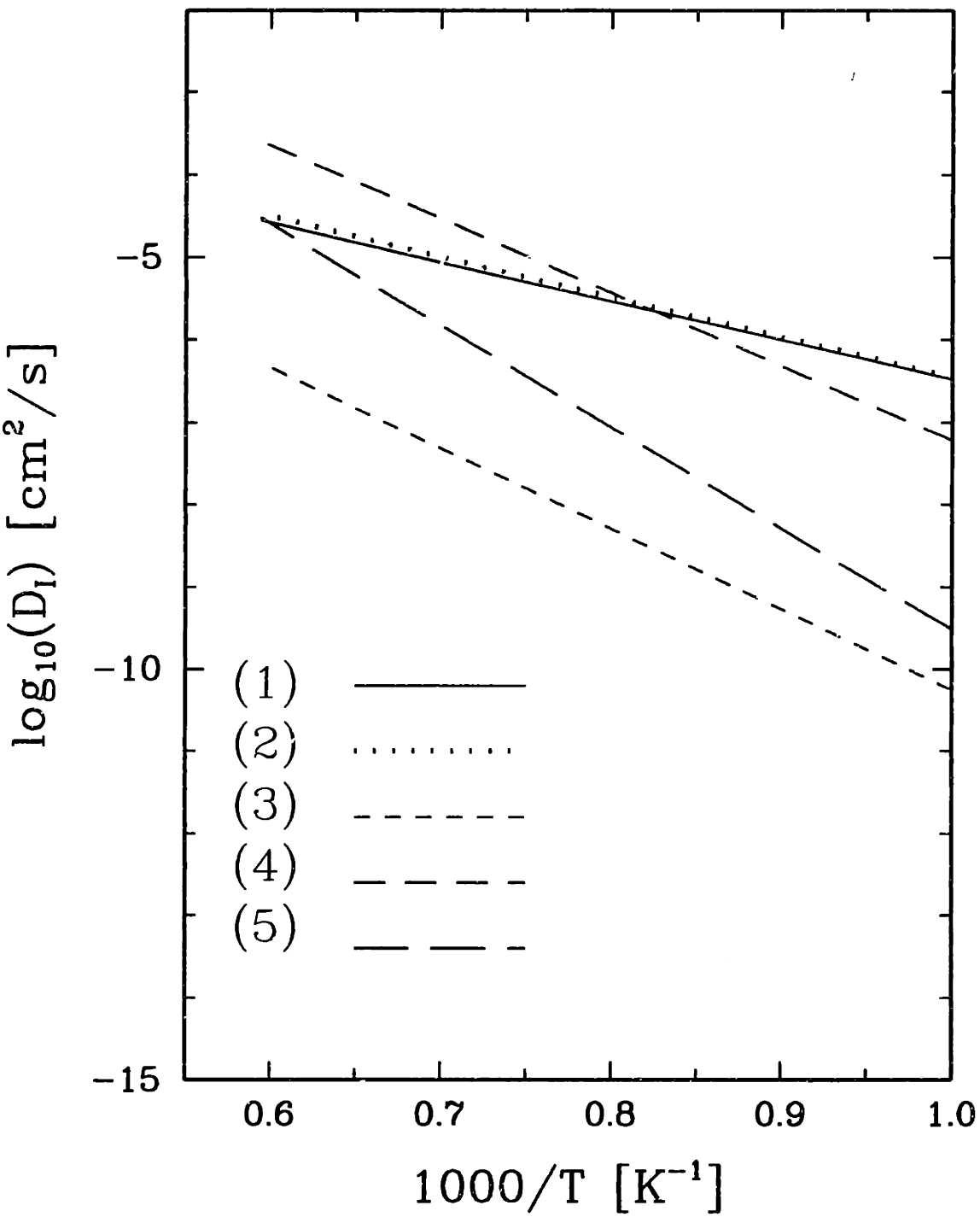

Figure 3-8: Self-interstitial diffusivities as a function of temperature: (1) This work; (2) Gold Diffusion (Zimmermann and Ryssel, 1992); (3) OED/ORD (Wijaranakula, 1990); (4) Zinc Diffusion (Bracht *et al.*, 1995); (5) OED/Boron Diffusion (Park and Law, 1992)

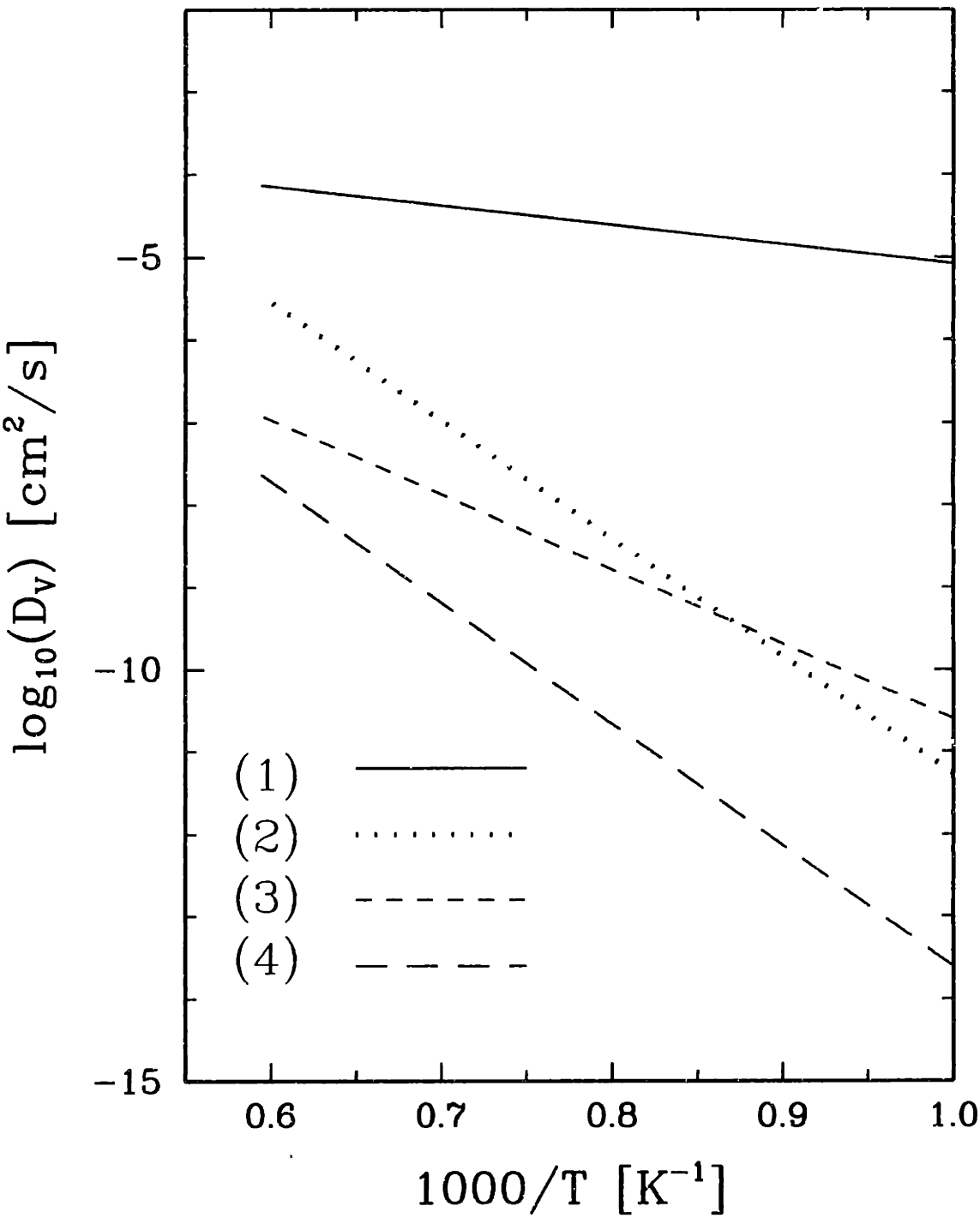

Figure 3-9: Vacancy diffusivities as a function of temperature: (1) This work; (2) Gold Diffusion (Zimmermann and Ryssel, 1992); (3) Zinc Diffusion (Bracht *et al.*, 1995); (4) OED/Boron Diffusion (Park and Law, 1992)

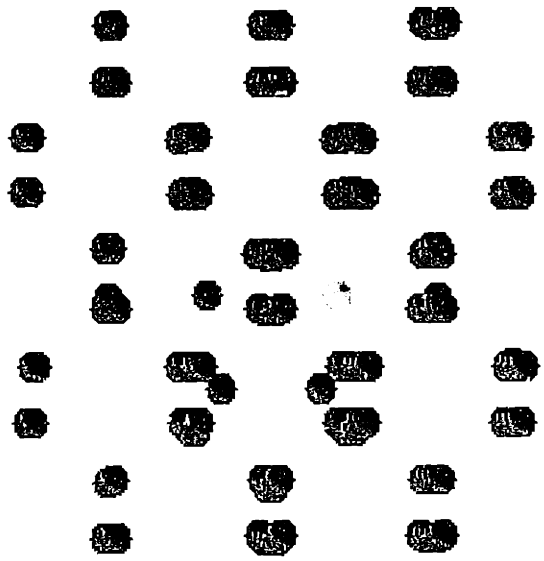

Figure 3-10: Equilibrium structure of the self-interstitial at zero temperature: The $< 110 >$-oriented split-dumbbell.

tial (Gilmer *et al.*, 1995). The structure was generated in MD simulations by allowing the system to relax without constraint until no further changes were observed. The ground state (zero-temperature) enthalpy for the $< 110 >$ dumbbell was calculated to be 3.66 eV, as previously determined by Monte Carlo simulations using the SW potential (Maroudas, 1992). The zero temperature energy was determined using the steepest descent algorithm to quench the atomic configuration (Maroudas, 1992).

Other atomistic simulation studies employing both empirical potentials and LDA total energy methods have predicted very similar structures for the lowest energy configuration of the self-interstitial. The *ab initio* results of Blöchl et al. (1993) give the ground state formation energy of the $< 110 >$ dumbbell to be 3.3 eV. This was also found to be the lowest energy structure for the self-interstitial. Other studies that have investigated the formation enthalpy of the $< 110 >$ dumbbell have led to similar estimates. A summary of these results is given in Table 3.4.

Maroudas (1992) demonstrated that the SW potential predicts an inwardly relaxed configuration around a neutral vacancy. This atomic structure is shown in

Calculation Method	< 110 >-dumbbell Formation Energy
SW Potential	2.66 eV
LDA (1)	4.1 eV
LDA (2)	3.65 eV
Tight-Binding (1)	3.97 eV
Tight-Binding (2)	3.96 eV

Table 3.4: Zero Temperature Energies of Formation for the < 110 >-dumbbell Configuration of the Silicon Self-Interstitial; LDA (1) (Blöchl *et al.*, 1993); LDA (2) (Zhu *et al.*, 1995); Tight-Binding (1) (Tang *et al.*, 1997); Tight-Binding (2) (Goodwin *et al.*, 1989)

Figure 3-11. The inward relaxation has been shown to be consistent with recent LDA

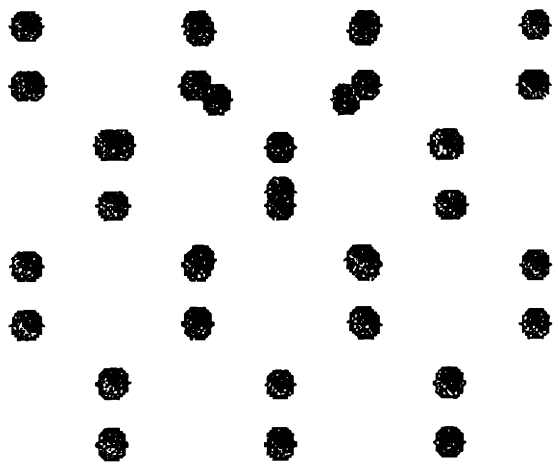

Figure 3-11: Equilibrium structure of the inwardly relaxed vacancy at zero temperature.

calculations (Zhu *et al.*, 1995). The zero temperature formation enthalpy for the vacancy is predicted to be 2.66 eV using the SW potential. The same result was obtained with the Monte Carlo simulations of Maroudas (1992). There is considerably greater uncertainty in the reported simulation estimates for the vacancy formation energy. Table 3.5 summarizes these results. In general, the classical potential estimates are lower then the *ab intio* results which give estimates as high as 4.1 eV. Electronic structure calculations have shown that the formation of a neutral vacancy is accompanied by a Jahn-Teller distortion which pairs the four dangling bonds around the vacancy (Blöchl *et al.*, 1993). Since the Jahn-Teller distortion is an electronic effect, it is not captured with a classical potential. This fact probably explains the relatively lower formation energies predicted with the classical potentials as compared to LDA estimates. However, it cannot be said at this point whether any of these studies is accurate because large disagreements exist even within the *ab initio* calculations. The tight-binding (TB) results predict vacancy formation energies that lie in between the classical and LDA results. The Jahn-Teller distortions are captured by the TB

Calculation Method	Vacancy Formation Energy	Calculation Method	Vacancy Formation Energy
SW Potential	2.66 eV	Classical (1)	3.2 eV
LDA (1)	2.8 eV	LDA (2)	3.3-3.6 eV
LDA (3)	4.1 eV	LDA (4)	3.5 eV
Tight-Binding (1)	3.7 eV	Tight-Binding (2)	3.8 eV
Tight-Binding (3)	3.15 eV	Tight-Binding (4)	3.68 eV

Table 3.5: Zero Temperature Energies of Formation for the Vacancy; Classical (1), (Baskes *et al.*, 1989); LDA (1), (Virkkunen *et al.*, 1993); LDA (2), (Kelly and Car, 1992); LDA (3), (Blöchl *et al.*, 1993); LDA (4), (Nichols *et al.*, 1989); Tight-Binding (1), (Wang *et al.*, 1991); Tight-Binding (2), (Mercer and Chou, 1993); Tight-Binding (3), (Rasband *et al.*, 1996a); Tight-Binding (4), (Song *et al.*, 1993)

method (Song *et al.*, 1993) which accounts for the relatively close agreement between the TB and LDA estimates for the vacancy formation energies.

3.6.4 Free Energies of Formation and Equilibrium Concentrations

The free energies of formation computed here have been separated into temperature dependent expressions for the enthalpies and entropies of formations in the linear form

$$X^f = X_0^f + TX_1^f, \tag{3.64}$$

where X^f is a formation property for either self-interstitials or vacancies. The temperature dependent enthalpies and entropies of formation for interstitials and vacancies computed with the thermodynamic integration scheme described in Section 3.5.1 are plotted in Figure 3-12 and compared with the cumulant analysis results of Maroudas (1992). The results for the enthalpies of formation from both methods are quite close and very weakly temperature dependent. However, there exists a large discrepancy in the case of the entropies of formation. The cumulant expansion appears to predict a *decreasing* vacancy entropy of formation as temperature increases, approaching zero entropy at the melting temperature. The interstitial value peaks at about 6 k for $T = 800\ K$, and decreases to 4 k near the melting temperature. The thermodynamic integration results show that both entropies of formation are strongly temperature dependent and increase linearly to 8 k for self-interstitials and 5 k for vacancies.

These large increases are suggestive of the increased delocalization of each point defect with increasing T. Seeger and Chik (1968) first postulated the delocalization of point defects with increasing temperature and Seeger et al. (1977) used this idea to explain the silicon self-diffusion data of Mayer et al. (1977). The parameter fitting used in this analysis yielded S_I^f from 1.00 k at $T = 570\ K$ to 6.11 k at $T = 1658\ K$, while the enthalpy of formation varied only between 2.59 eV and 3.04 eV across the same temperature range. The change in formation entropy predicted by thermodynamic integration is the same as found by Seeger et al. (1977) although the

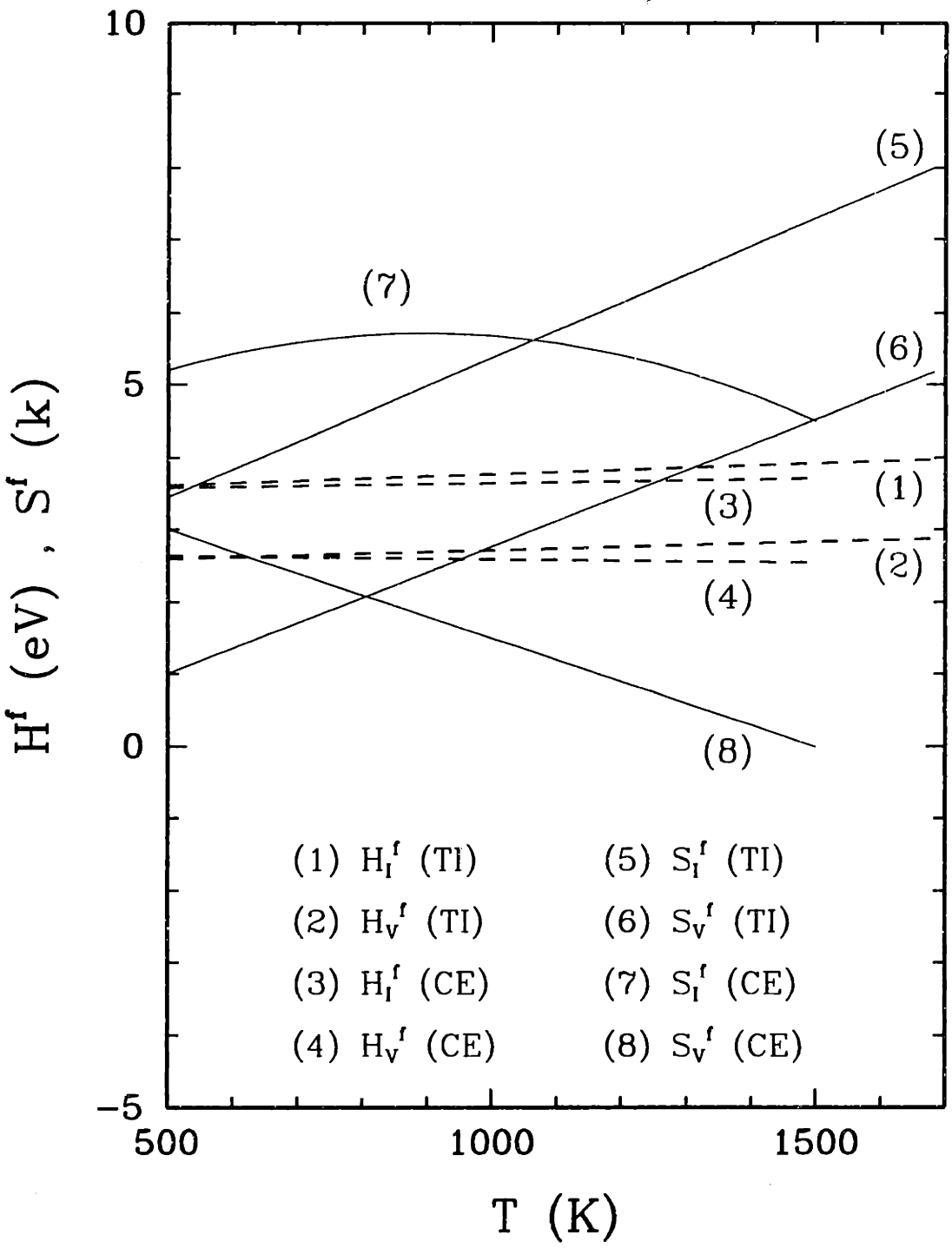

Figure 3-12: Enthalpies and entropies of formation for self-interstitials and vacancies; (TI) = thermodynamic integration, (CE) = cumulant expansion.

absolute values of our estimates for S_I^f are 3 k higher.

The *ab initio* results of Blöchl et al. (1993) estimated S_I^f to be constant at 6 k. This result was computed using the Local Harmonic Approximation (LHA) method of Najafabadi et al. (1989) which decouples the vibrational motion of each atom and leads to N 3×3 local dynamical matrices for a system of N atoms. The result is in qualitative agreement with our thermodynamic integration results. There are serious limitations on the use of the LHA (Ungar *et al.*, 1993) and it is likely that the accuracy of the LDA result of Blochl et al (1993) is limited. Tang et al. (1997) used TB to compute diffusion and formation energies and fitted the entropy of formation to experimental data for the interstitial self-diffusion rate, $C_I^{eq} D_I$. Their analysis gives $S_I^f(1523\ K) = 11.2\ k$. While these numbers are not in quantitative agreement, they all indicate that the high temperature self-interstitial is quite delocalized and possesses a large entropy of formation. Characterization of point defect delocalization is important for modeling point defect recombination kinetics, as is discussed in Section 3.6.6.

The general expression for the equilibrium concentrations of point defects is given by

$$C^{eq} \equiv A \exp\left(S_0^f + T S_1^f\right) \exp\left(-\frac{H_0^f + T H_1^f}{kT}\right). \tag{3.65}$$

The estimates for the equilibrium concentrations are summarized below in eqs. (3.66)-(3.67).

$$C_I^{eq}(T) = 2.97 \times 10^{23} \exp\left(\frac{1.53 + 3.85 \times 10^{-3} T}{k}\right) \exp\left(-\frac{3.46 + 3.08 \times 10^{-4} T}{kT}\right) \ cm^{-3} \tag{3.66}$$

$$C_V^{eq}(T) = 4.97 \times 10^{22} \exp\left(\frac{0.76 + 3.53 \times 10^{-3} T}{k}\right) \exp\left(-\frac{2.48 + 2.33 \times 10^{-4} T}{kT}\right) \ cm^{-3} \tag{3.67}$$

The SW estimates for the point defect equilibrium concentrations are compared in Figures 3-13 and 3-14 with other experimental estimates for the temperature range $1000 \leq T \leq 1683\ K$. Most estimates for the self-interstitial equilibrium concentrations near the melting point are in agreement to within an order-of-magnitude, and range between $10^{15} - 10^{16}\ cm^{-3}$. The OED study of Park and Law (1992) is the

highest estimate at lower temperatures and deviates from the other values. It is noticeable that all the metal diffusion experiments agree very well with each other. As in the case of the point defect diffusivities, the vacancy equilibrium concentrations exhibit much larger uncertainties. The OED/ORD estimate of Park and Law (1992) is again the highest value by about an order-of-magnitude at T_m. Even though the SW potential gave a relatively low estimate for the formation energy compared to other computational estimates (see Table 3.5, it is still higher than all the experimental formation energies shown in Figure 3-14 (steeper Arrhenius curve).

Most experimental estimates predict that the equilibrium concentration of vacancies is higher than the interstitial concentrations, especially at lower temperatures. The SW calculations agree with this picture, but *ab initio* calculations strongly favor self-interstitials at high temperatures (Blöchl *et al.*, 1993). This contradiction is largely due to the high formation energy estimates predicted by these calculations.

3.6.5 Self-Diffusion Coefficients

The OED/ORD and metal diffusion estimates for point defect properties are subject to significant uncertainty. The lack of agreement on the values of point defect properties between different experiments has yet to be completely understood. While OED/ORD experiments predict $D_I(1373K) \sim (10^{-9})\ cm^2/s$, metal in-diffusion results generally give $D_I(1373\ K) \sim O(10^{-6})\ cm^2/s$. In the case of OED/ORD experiments, the transport model is applied near the Si/SiO_2 interface, while for metal in-diffusion experiments the model is applied in the bulk of the wafer. Agrawal and Dunham (1995) have provided an explanation for the discrepancy by including surface segregation effects in the modeling of OED/ORD experiments, giving diffusion coefficients much closer to the metal in-diffusion results. In general, at temperatures near the melting point, the OED/ORD experiments give equilibrium concentrations that are two orders-of-magnitude higher and diffusivities that are two orders-of-magnitude lower than values from the metal in-diffusion experiments.

The discussion in Section 3.1, however, made the point that these methods allow for relatively direct estimation of the *self-diffusion coefficients* which are defined as

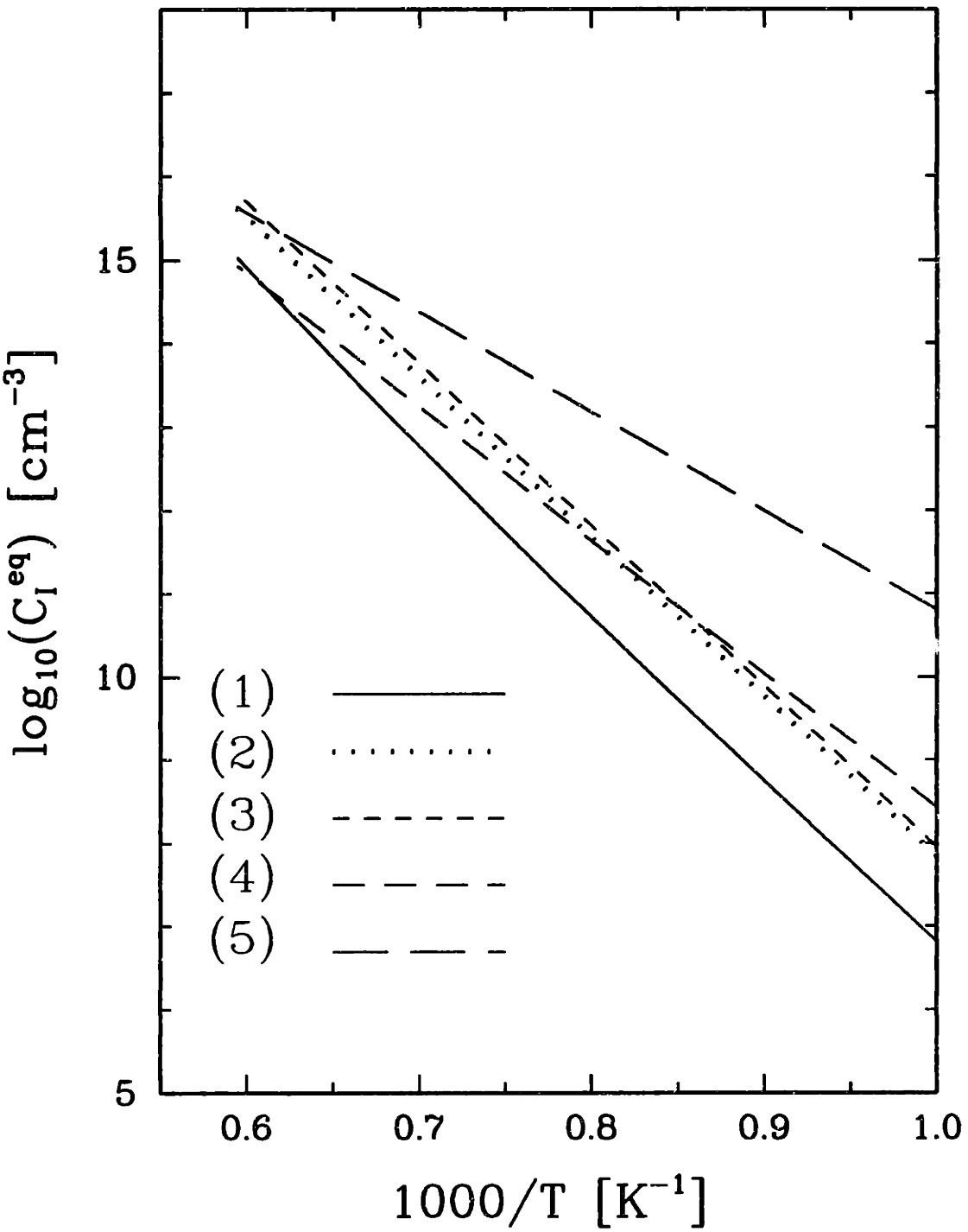

Figure 3-13: Self-interstitial equilibrium concentration as a function of temperature: (1) This work; (2) Diffusion (Morehead, 1988); (3) Gold Diffusion (Zimmermann and Ryssel, 1992); (4) Zinc Diffusion (Bracht *et al.*, 1995); (5) OED/Boron Diffusion (Park and Law, 1992)

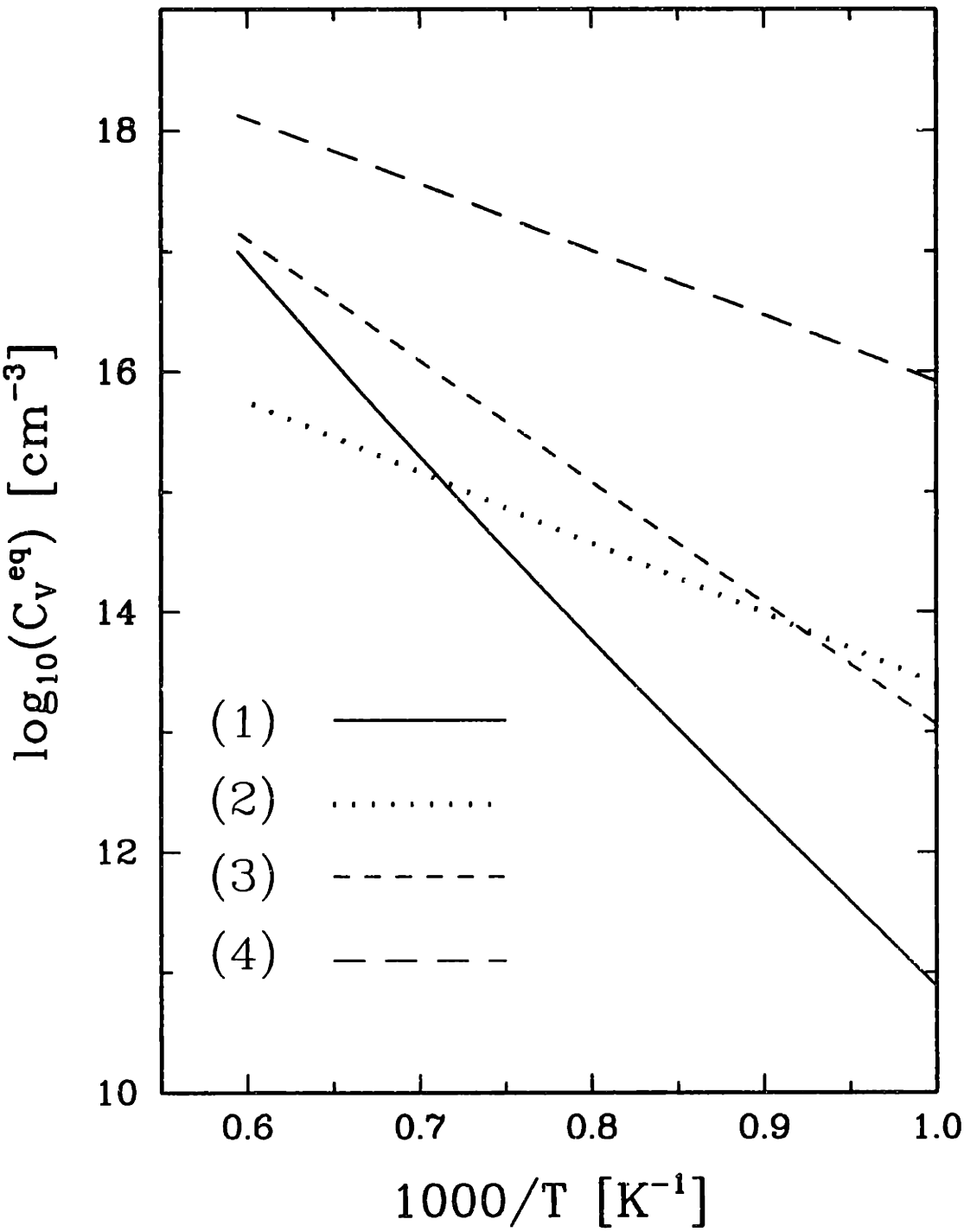

Figure 3-14: Vacancy diffusivities as a function of temperature: (1) This work; (2) Gold Diffusion (Zimmermann and Ryssel, 1992); (3) Zinc Diffusion (Bracht *et al.*, 1995); (4) OED/Boron Diffusion (Park and Law, 1992)

the product of point defect diffusivities and equilibrium concentrations. These are generally written in Arrhenius form as

$$D_X C_X^{eq} = A \exp\left(-\frac{H_X^m + H_X^f}{kT}\right), \tag{3.68}$$

where $X = I, V$ is the point defect species. The SW estimate for these properties for self-interstitials and vacancies are given below as

$$D_I C_I^{eq} = 5.23 \times 10^{21} \exp\left(\frac{1.53 + 3.85 \times 10^{-3} T}{k}\right) \exp\left(-\frac{4.40 + 3.08 \times 10^{-4} T}{kT}\right) \ cm^{-3}, \tag{3.69}$$

$$D_V C_V^{eq} = 8.45 \times 10^{19} \exp\left(\frac{0.76 + 3.53 \times 10^{-3} T}{k}\right) \exp\left(-\frac{2.94 + 2.33 \times 10^{-4} T}{kT}\right) \ cm^{-3}. \tag{3.70}$$

The SW estimates for the self-diffusion coefficients are compared in Figures 3-15 and 3-16 to those obtained from metal gettering and OED/ORD experiments. For both the vacancy and self-interstitial cases, there is very good agreement between all the experimental results both for activation energies and preexponential factors, resulting in scatter over less than a factor of three between all these estimates. It is generally agreed that the correlation of Stolwijk et al. (1986)

$$D_I C_I^{eq} = 3.2 \times 10^{25} \exp\left(-\frac{4.8\ eV}{kT}\right) \ cm^{-1} s^{-1}, \tag{3.71}$$

represents a good estimate. The SW approximation for the self-interstitial contribution to self-diffusion is just under one order-of-magnitude lower than these estimates, with good agreement for $H_I^m + H_I^f$. The SW vacancy estimate is not in agreement with the experimental data shown with the SW estimate almost three orders-of-magnitude higher across the temperature range plotted; alternatively, $H_V^m + H_V^f$, appears to be in reasonable accord. The disagreement is due mainly to the very high vacancy diffusion coefficient estimate. While the actual values for the self-diffusion coefficient predicted with the SW potential are not entirely in agreement with experimental measurements, they are at least qualitatively accurate. The activation energies of migration and formation are well predicted for both point defects. The main source

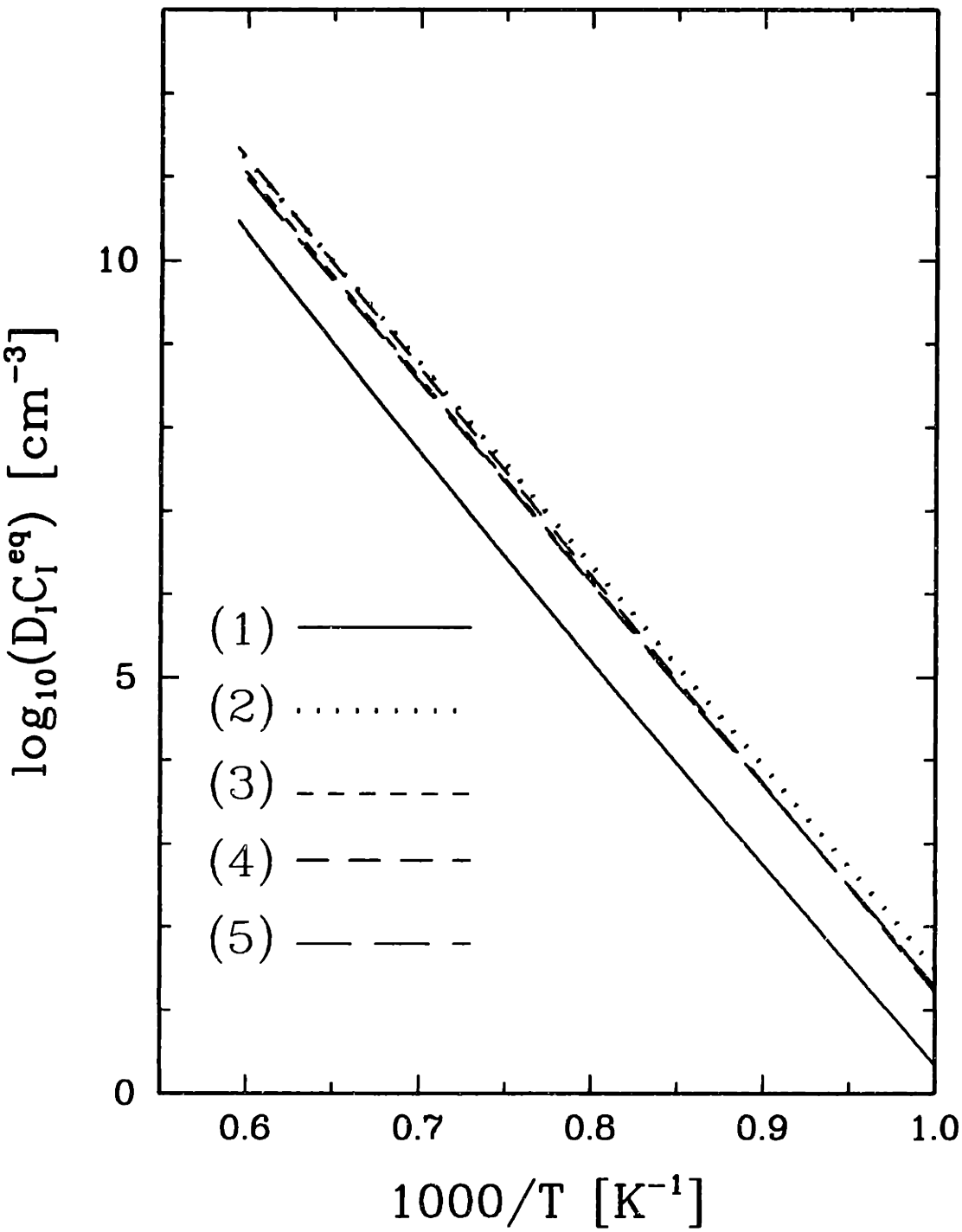

Figure 3-15: Estimates for $D_I C_I^{eq}$ as a function of temperature: (1) This work; (2) Gold Diffusion (Zimmermann and Ryssel, 1992); (3) Zinc Diffusion (Bracht *et al.*, 1995); (4) OED/Boron Diffusion (Park and Law, 1992)

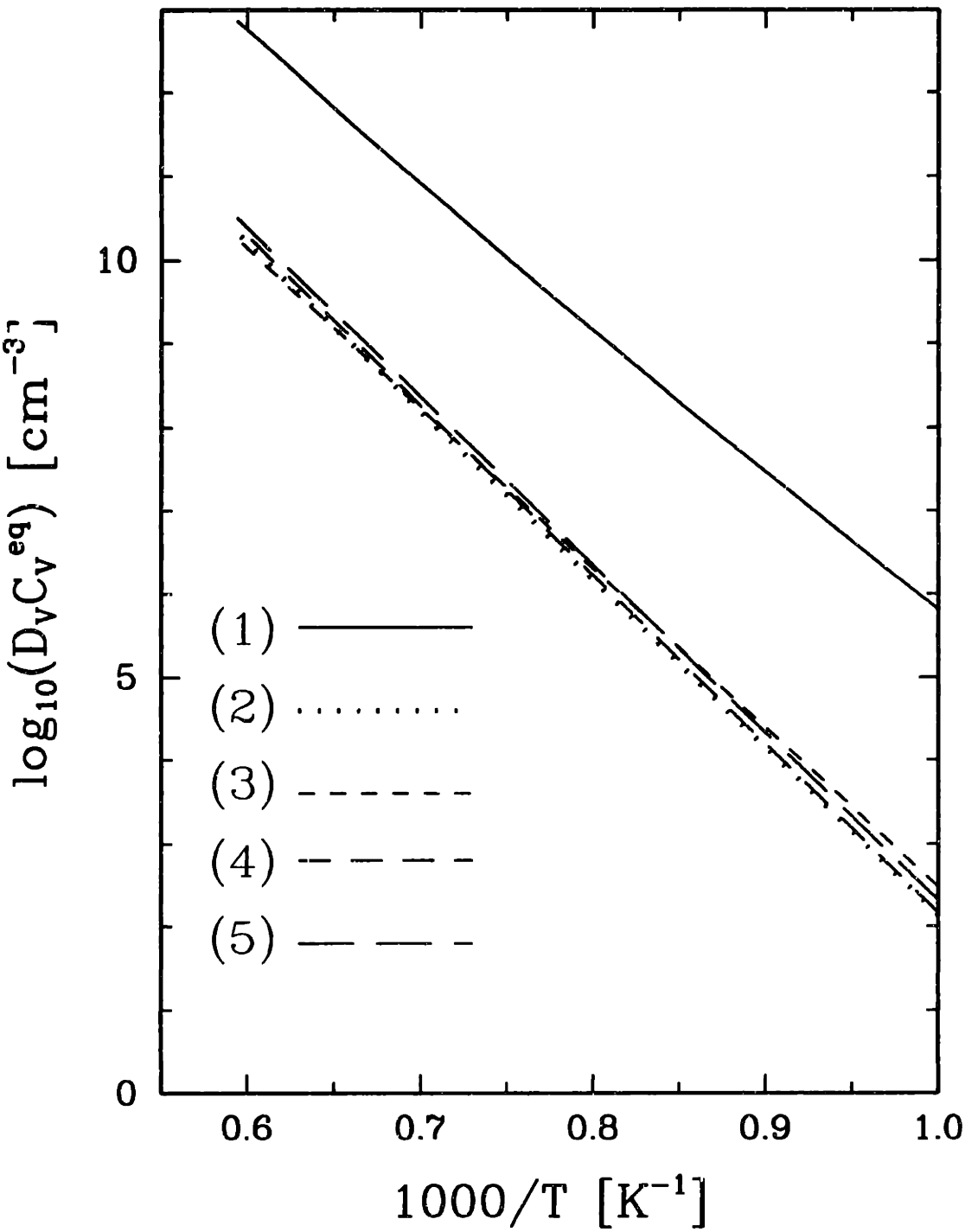

Figure 3-16: Estimates for $D_V C_V^{eq}$ as a function of temperature: (1) This work; (2) Gold Diffusion (Zimmermann and Ryssel, 1992); (3) Zinc Diffusion (Bracht *et al.*, 1995); (4) Gold Diffusion (Gösele and Tan, 1983); (5) OED/Boron Diffusion (Park and Law, 1992)

of error appears to stem from the computation of the entropic contribution, which is large especially at high temperature. It is also fairly evident that the SW potential models the thermophysical properties of the self-interstitial more accurately than those of the vacancy. This is possibly a consequence of the important role of electronic structure in the vacancy geometry. It is well known that the SW potential is not well suited to computing the energies of "open" structures such as surfaces and small clusters (Ungar *et al.*, 1993; Chelikowsky and Phillips, 1990). These structures are similar to the vacancy in the sense that they all contain a large number of unbonded molecular orbitals. Empirical potentials with special attention to this effect have been developed (Chelikowsky and Phillips, 1990). The potential of Chelikowsky and Phillips (1989) (CP) in fact predicts a very large enthalpy of formation for the vacancy, 4.7 eV. This estimate is higher than most accepted values. The CP potential exhibits a very sharp gradient in the strain field near the vacancy, which effectively means that the lattice is inelastic and imposes a large energy penalty on the presence of a vacancy (Glassford *et al.*, 1991). The high estimate of Glassford et al. (1991) for the vacancy formation energy shows that it is probably necessary to include the electronic problem to obtain a good value for the vacancy formation energy.

3.6.6 Recombination Kinetics

One of the most important factors in determining the resulting intrinsic point defect concentrations for a given set of thermal conditions during crystal growth and wafer annealing is the recombination rate between self-interstitials and vacancies. The reversible reaction can be expressed as

$$I + V \rightleftharpoons 0, \tag{3.72}$$

where 0 refers to a perfect lattice atom. The effect of this reaction is to produce an interdependency between interstitial and vacancy distributions. Despite the importance of this reaction, its kinetics are poorly understood (Fahey, Griffin, and Plummer, 1989). The reaction rate is dependent on an activation energy barrier which is a

function of diffusion and kinetic limitations. These limitations are incorporated into a rate expression (Fahey, Griffin, and Plummer, 1989)

$$k_{IV}(T) = \frac{4\pi a_r}{\Omega C_s}[D_I(T) + D_V(T)] \exp(-\frac{\Delta G_{IV}(T)}{kT}), \tag{3.73}$$

where a_r is an effective capture radius, C_s is the atomic concentration of lattice sites in silicon, Ω is the atomic volume of a silicon atom, D_x is the diffusivity of point defect species (x), and ΔG_{IV} is the free energy barrier against recombination. The diffusion rates have been discussed in the previous sections. However, the nature of the kinetic barrier is poorly understood. The activation barrier consists of enthalpic and entropic components (Seeger and Chik, 1968) written here as

$$\Delta G_{IV}(T) \equiv \Delta H(T) - T\Delta S(T). \tag{3.74}$$

There is a general consensus that the overall barrier is of order 1.5 eV (Antoniadis and Moskowitz, 1982; Fahey *et al.*, 1989). Antoniadis and Moskowitz interpreted their defect lifetime measurements at 1373 K in terms of a purely enthalpic barrier with zero entropic contribution. Their estimate was derived by comparing the experimentally observed lifetimes to a theoretical value based on diffusion limited reaction theory (Waite, 1957). However, Gösele et al. (1982), interpreted the data in terms of a purely entropic barrier of 11.5 k. The theoretical basis for this kind of barrier comes from the behavior of point defects at higher temperatures, where they lose their point-like characteristics and become more diffuse or extended. This feature causes the entropy of formation of the point defect to increase, making it more stable and less amenable to the recombination reaction, which has been confirmed with the thermodynamic integration results presented in Section 3.6.3. Figure 3-17 shows the extent of distortion near an interstitial defect as the temperature is increased. In addition to the increased vibrational energy, which increases the baseline level, it can be seen that the distortion persists across a longer distance away from the defect.

Recently, tight-binding molecular dynamics simulations have been applied to the study of the interstitial/vacancy recombination energy barrier (Tang *et al.*, 1997).

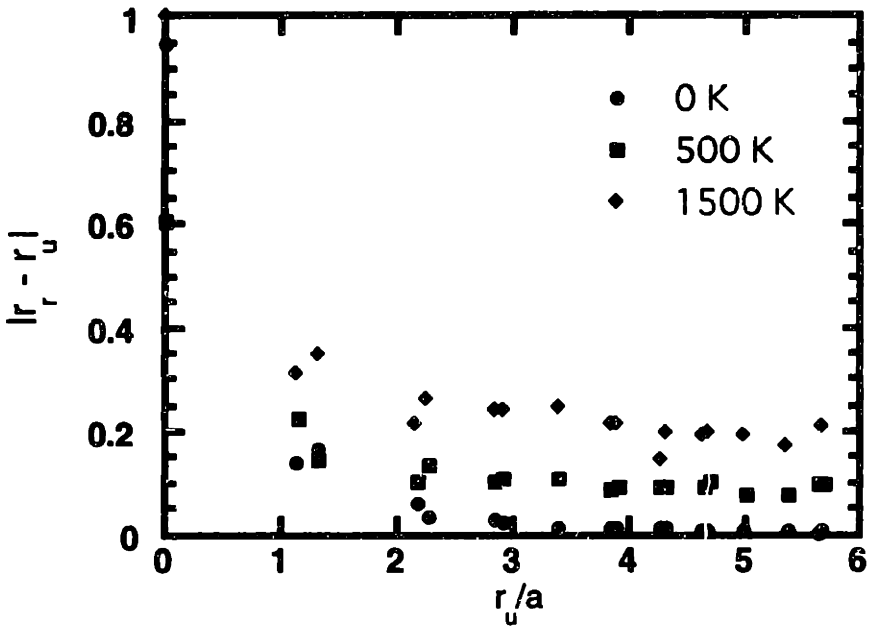

Figure 3-17: Extent of Lattice distortion around a self-interstitial as a function of temperature; (a) 500 K, (b) 1000 K, and (c) 1500 K.

These simulations have determined an *enthalpic barrier* of about 1.1 eV for the annihilation of a metastable $I - V$ complex. This value is in reasonable agreement with the experimental measurement of Antoniadis and Moskowitz (1982).

The aim of this Section is to present a model for determining the functional form for the free energy barrier to interstitial/vacancy recombination based on the single-temperature data point of Antoniadis and Moskowitz (1982) and the results of the SW simulations presented above. The basis for this model is a hypothetical reaction pathway, which, combined with free energy of formation data, is used to give a quantitative expression for the free energy barrier. As shown in Figure 3-18, the recombination process is postulated to take place in several sequential stages. First, the two extended defects diffuse to within the interaction radius. Then, before recombination can take place, the point defects have to revert to their zero-temperature, or localized, states. The recombination of the interstitial and the vacancy in their localized form is the final step. In additional to the delocalization entropy that must be lost, a structural reorientation must place which would possibly require that the

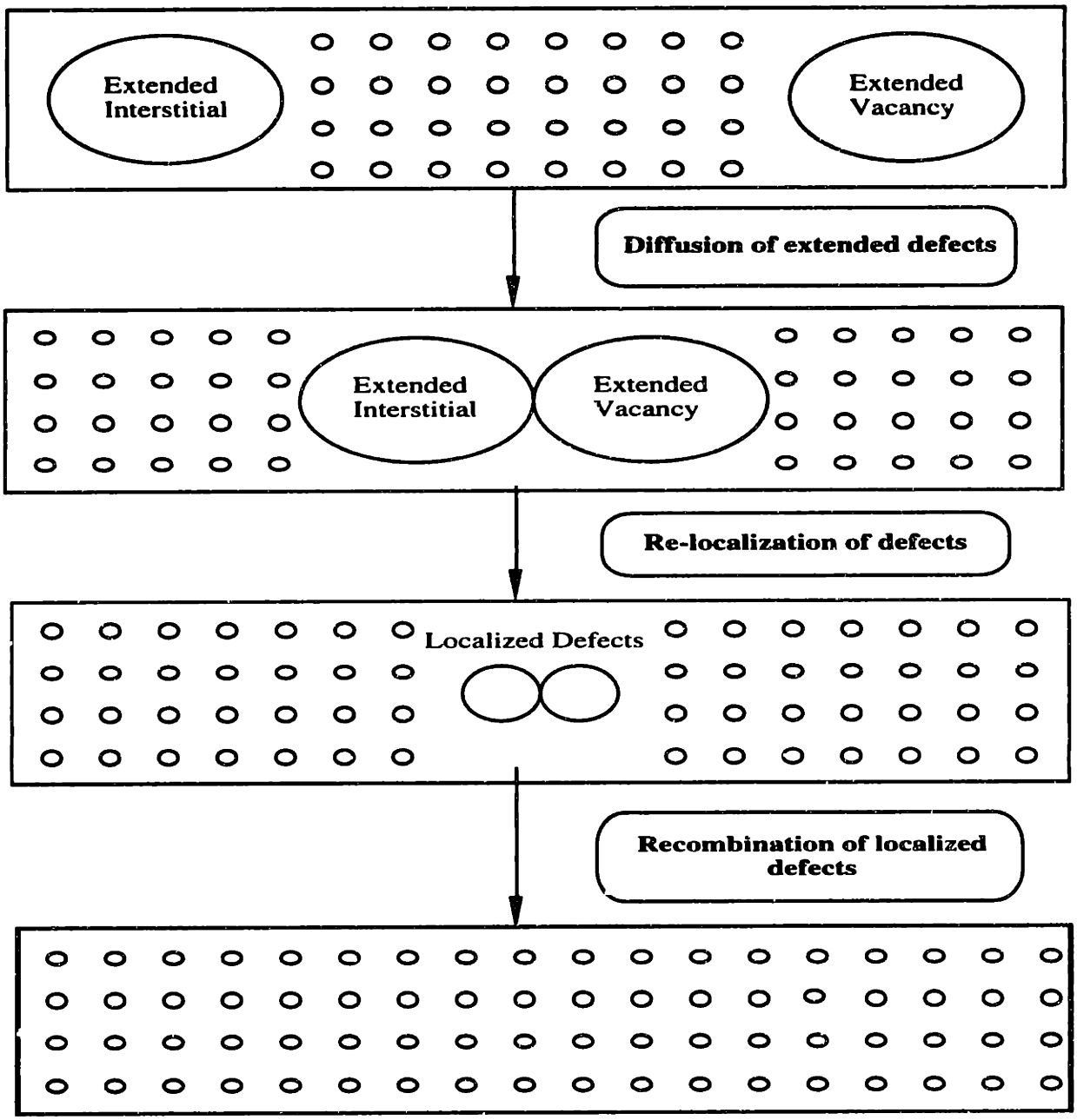

Figure 3-18: Diffusion-limited reaction model for the recombination of self-interstitials and vacancies with enthalpic and entropic contributions to the free energy activation barrier.

local atomic coordinates pass through a high-energy transition state leading to an enthalpic component to the overall barrier. The enthalpic contribution is estimated as the remainder of the Antoniadis and Moskowitz barrier once the entropies of formation for the two point defects have been subtracted. The overall barrier to the recombination process therefore is represented as

$$\Delta G_{IV} = [S_i^f(T) - S_i^f(0)] + [S_v^f(T) - S_v^f(0)] + \Delta H_{IV} \approx 1.5eV, \tag{3.75}$$

where ΔH_{IV} is the temperature independent enthalpic portion of the energy barrier. This model gives the temperature dependent recombination rate as

$$k_{IV}(T) = 1.2 \times 10^{-6} \exp\left(-\frac{0.58 - T(2.29 + 7.38 \times 10^{-3}T)}{kT}\right) \tag{3.76}$$

The enthalpy and entropy terms are given in units of eV and Boltzmann constants, respectively. Other distributions are possible with varying amounts of enthalpic and entropic contributions. Eqs. (3.77) and (3.78) show the free energy barrier for enthalpy-only and entropy-only models, respectively.

$$k_{IV}(T) = 1.2 \times 10^{-6} \exp\left(-\frac{1.5}{kT}\right). \tag{3.77}$$

$$k_{IV}(T) = 1.2 \times 10^{-6} \exp\left(-\frac{2.29 + 7.38 \times 10^{-3}}{k}\right). \tag{3.78}$$

Figure 3-19 shows a plot of the recombination rate as a function of temperature according to our model as compared to purely enthalpic or entropic models, which represent the two limits for describing the barrier. The purely entropic barrier actually leads to an increase of the reaction rate with decreasing temperature, but finally starts to decrease after 600 K as point defect diffusion becomes very small. The enthalpic model gives higher recombination rates near the melting temperature but decreases very rapidly as the crystal cools.

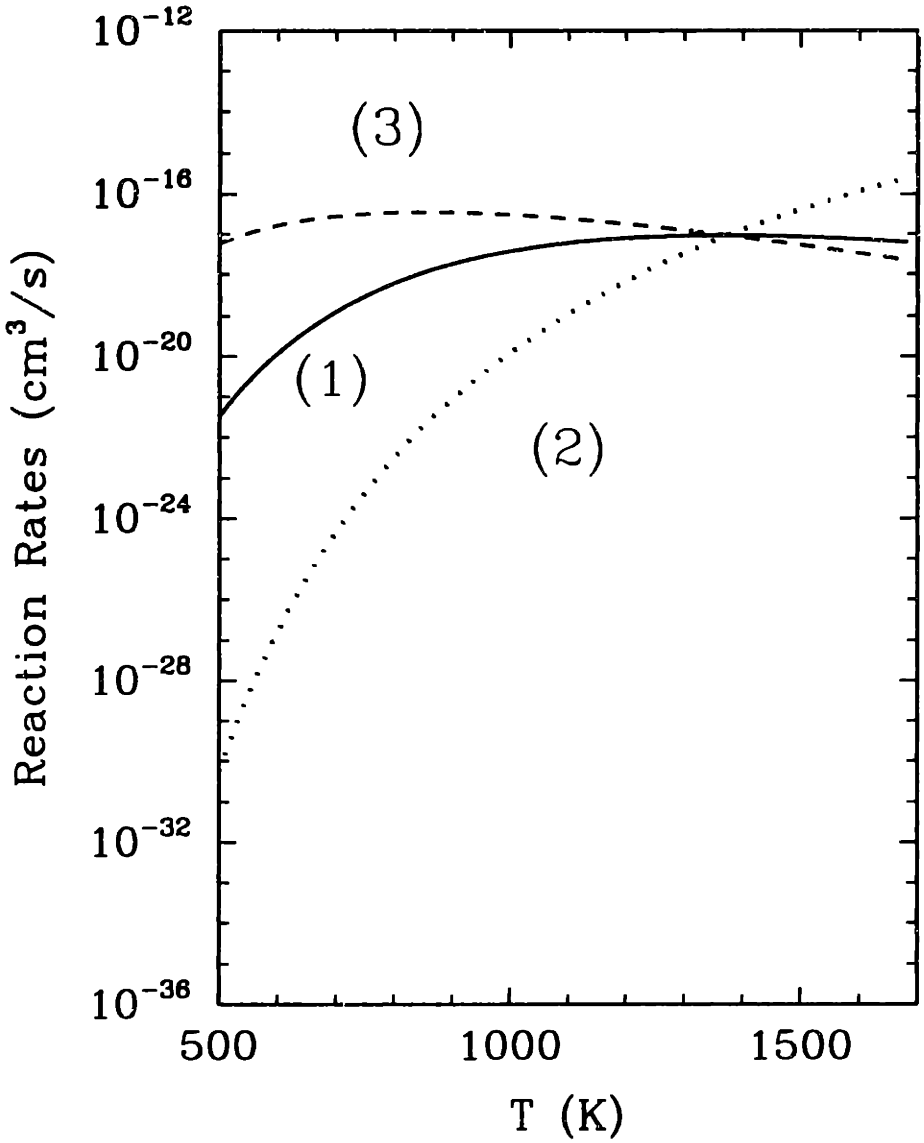

Figure 3-19: Point defect recombination rates computed with (1) Entropic/Enthalpic model, eq. (3.76), (2) Enthalpic only model, eq. (3.77) (3) Entropic only model, eq. (3.78).

3.7 Summary and Discussion

The Stillinger-Weber interatomic potential has been used in a comprehensive computational framework to calculate the thermophysical properties of self-interstitials and vacancies, namely the equilibrium concentrations, diffusivities, and the recombination rate. Optimization of the molecular dynamics method has allowed for the use of computationally intensive thermodynamic integration to be used in the computation of temperature dependent free energies of formation as well as direct evaluation of the diffusion coefficients.

Our results for point defect dynamics are in qualitative agreement with experimental data, especially in the case of the self-interstitial. The vacancy properties, in particular the diffusion coefficient, are susceptible to more uncertainty. A possible explanation for this is the importance of electronic structure in the case of the neutral vacancy which the Stillinger-Weber potential effectively ignores. However, most estimates for vacancy diffusion coefficients and equilibrium properties in the literature are still uncertain and atomistic simulation is clearly a very viable option given the expense of experimental measurements.

The advent of faster computing is certain to make atomistic simulation even more appealing as the use of *ab initio* schemes becomes more feasible. Furthermore, methods such as the tight-binding approximation is allowing for some level of electronic structure detail to be already incorporated into large, dynamic simulations of point defect diffusion and reaction (Rasband *et al.*, 1996a; Tang *et al.*, 1997).

Chapter 4

Atomistic Simulation of Point Defect Aggregation

4.1 Introduction

The previous Chapter demonstrated that atomistic simulation is extremely useful in the investigation of point defect structures and thermodynamics. The ultimate goal in the atomistic study of crystalline defects however, is the detailed characterization of point defect aggregation processes which lead to microdefect formation. In principle, the extension of the methodology described in Chapter 3 to the investigation of point defect clusters is very straightforward; multiple point defects must be considered within the same simulation cell during a particular simulation and they must be allowed to interact. In order to allow such a system to fully relax, the total number of atoms within the simulation cell must be increased dramatically to accommodate multiple point defects. The increase is required to keep the defect concentration low enough to minimize finite-cell effects.

The increase in computational expense as a function of simulation size is dependent on the form of the interatomic potential; in general, the more complicated the potential the worse the scaling. In the case of the three-body Stillinger-Weber interatomic potential, simulations rapidly become prohibitively expensive on single-processor workstations as the total number of atoms approaches $O(10^4)$. For point

defect concentrations equal to those considered low enough to eliminate finite size effects (about one point defect in two hundred lattice atoms as in the simulations performed in Chapter 3), this would allow a maximum of 30-40 point defects to be considered simultaneously.

A solution to this limitation can be provided by the use of parallel computer architectures to divide the simulation onto several processors allowing for much larger systems to be considered. Although there have been many examples of parallel implementation of molecular dynamics simulations, almost all of these implementations have focused on application to interatomic potentials with only pairwise interactions (Nakano *et al.*, 1994). Only Esselink et al. (Esselink and Hilbers, 1993) have considered parallel algorithms for N-body potentials ($N \geq 2$), but this research is not directly applicable to the SW potential because of its description of the three-body interactions; this will be discussed in greater detail later on.

The present approach is to achieve parallel computation by dividing physical space among $M = P^3$, $P = 1, 2, \cdots$, processors, and by letting atoms or particles to move according to Newton's laws of motion through this space as in any regular molecular dynamics simulation. The parallel efficiency of the algorithm depends on the amount and speed of the communications needed between the processors to share the coordinate information so that forces can be computed for each atom. The algorithm was implemented using the SW potential on an IBM SP2 computer having 8 processors ($N = 8$) with inter-processor communication performed in the MPL message passing library. Each processor is responsible for a varying number of particles which migrate freely across the processor boundaries. The algorithms developed here optimize the search for unique 2- and 3-body interactions across all the processors by equally distributing the work load across all processors.

4.2 Computational Method

4.2.1 General Parallelization Strategy

Geometric parallelization is used to divide equally the physical simulation space among M processors of the SP2 computer. The simulation domain is always cubic and this discretization results in M smaller cubes. The parallelization strategy is simplified by requiring that the simulation space is equally divided in each dimension. This implies that $M = P^3$, where P is any integer. The IBM SP2 computer used in this study has a total of 12 processors, thus a maximum of 8 processors ($P = 2$) can be employed in these simulations. For $P = 3, 4, 5$, would require the use of 27, 64, and 125 processors respectively. Most current parallel architectures are relatively coarse grained, employing a few, powerful processors; the SP2 is an example. Therefore, for most practical purposes, we can expect that $P = 4$ (64 nodes) will be the largest simulation we will be able to perform. This limitation has some negative consequences for the efficiency of a parallel code which will be discussed later. The IBM SP2 architecture is based on a high speed switch that effectively allows for direct communication between every pair of processors. In effect, all processors are nearest neighbors which eliminates the need to account for hardware structure when designing the communication scheme (Mel'čuk *et al.*, 1991; Fincham, 1987).

The IBM SP2 is prototypical of a type of SIMD parallel computer with relatively high speed processors, each with large processor memory. The machine used here has processors with 128 megabytes of memory associated with each CPU and delivers 130 MFLOPS on a LINPACK benchmark (Dongarra, 1992) running on a single node.

There is, however, a problem inherent in geometric parallelization using a relatively small number of processors; the three-dimensional representation in Figure 4-1 of the connectivity between eight processors for a simulation space subject to periodic boundary conditions (Allen and Tildesley, 1992; Haile, 1992) highlights the problem. For a system of eight processors, all are immediate neighbors of each other *in simulation space.* In this case, every processor will receive information about *all* the other particles in the system whenever particle information needs to be exchanged.

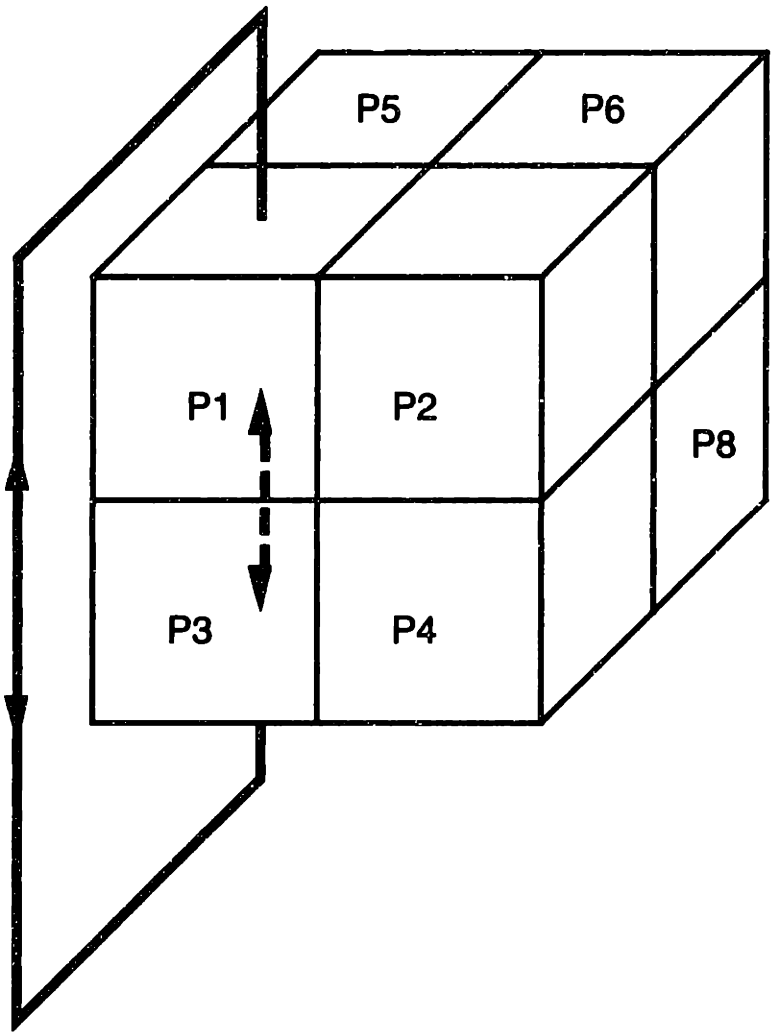

Figure 4-1: 3-dimensional representation of processor connectivity in simulation space.

Furthermore, Figure 4-1 shows an example of the communication redundancy in the case of eight processors; processors $P1$ and $P3$ are neighbors via *two* paths. The eight processor machine, P^3 with $P = 2$, is clearly the worst case but the situation is not much improved for larger, but still moderately sized machines; for $P = 4$ or 64 processors, each processor will still have to communicate with about half of total number (27 nearest neighbors in three dimensions out of 64).

4.2.2 The MPL Message Passing Protocol

Communication between processors is performed using the MPL message passing library. MPL is the IBM specific version of the more commonly used MPI library. Both libraries are based on very similar constructs for exchanging information between individual processors. The core commands used to transfer information between processors are the *send* and *receive* functions which operate either between individual processors or to all processors simultaneously. One of the main concerns in the implementation of message passing algorithms is that not all processors perform their individual tasks at the same speed. This can be due either to unbalanced loads or to external influences such as other jobs running concurrently on one or more of the processors. Thus, it is possible to have situations where a processor that has not yet received information required in a particular calculation to proceed without "waiting". Such possibilities can lead to catastrophic failure of an MD simulation and care must be taken to avoid them.

Processor synchronicity can be affected by several mechanisms. The most commonly employed are the *blocking* send and receive operations. These commands require that a processor wait for completion of any message passing before proceeding. The disadvantage in the use of these functions is the potential for program *deadlock*. This can result when two or more processors are waiting for each other to send or receive messages before either can continue. If this happens, the program simply waits indefinitely and must be terminated manually. Deadlock can be avoided by limiting the use of blocking commands to places where it is crucial that information be received before further computation. The other potential disadvantage of excessive use

of blocking commands is the detrimental effect on program speed. Every time that the processors are synchronized, the program is forced to execute at the speed of the slowest processor; the more often this is done the slower the overall speed.

4.2.3 The Cell Method in a Parallel Environment

For a short ranged potential such as SW, only particles near the spatial boundaries of each processor are capable of interacting across processors. In order to separate out these particles from the rest, and thus avoid wasted communication, each processor's domain is further subdivided into cells (Allen and Tildesley, 1992) as discussed in Chapter 3. If the size of these cells is chosen carefully, only those particles in the cells adjacent to the processor's boundary (the shaded regions) need be exchanged with other processors, as shown in Figure 4-2. The effects of imposing periodic boundary conditions on the structure of the shaded region are also noted there. In the case of parallel molecular dynamics, the optimal size for these cells is determined by minimizing the number of particles that have to be exchanged across processors during message passing. This number is the product of the number of cells that must be exchanged and the number of particles in each cell. Clearly, for a given density, the finer the cellular subdivision on each processor, the fewer particles there are in each cell. However, once the cell size becomes smaller than the pair interaction range, as given by *rlist* of the SW potential, it is possible for two interacting atoms to reside on *non-adjacent* cells. If this occurs, two or more layers of cells must now be exchanged across processors to ensure that all interactions are taken into account. These arguments suggest that the optimal cell size is equal to the interaction range for atomic pairs, *rlist*.

4.2.4 Structure of the SW interatomic potential

The form of the Stillinger-Weber interatomic potential is the most important consideration in determining how to divide the computation across the SP2 nodes shown in Figure 4-1. The complicating factor arises from the structure of the three-body

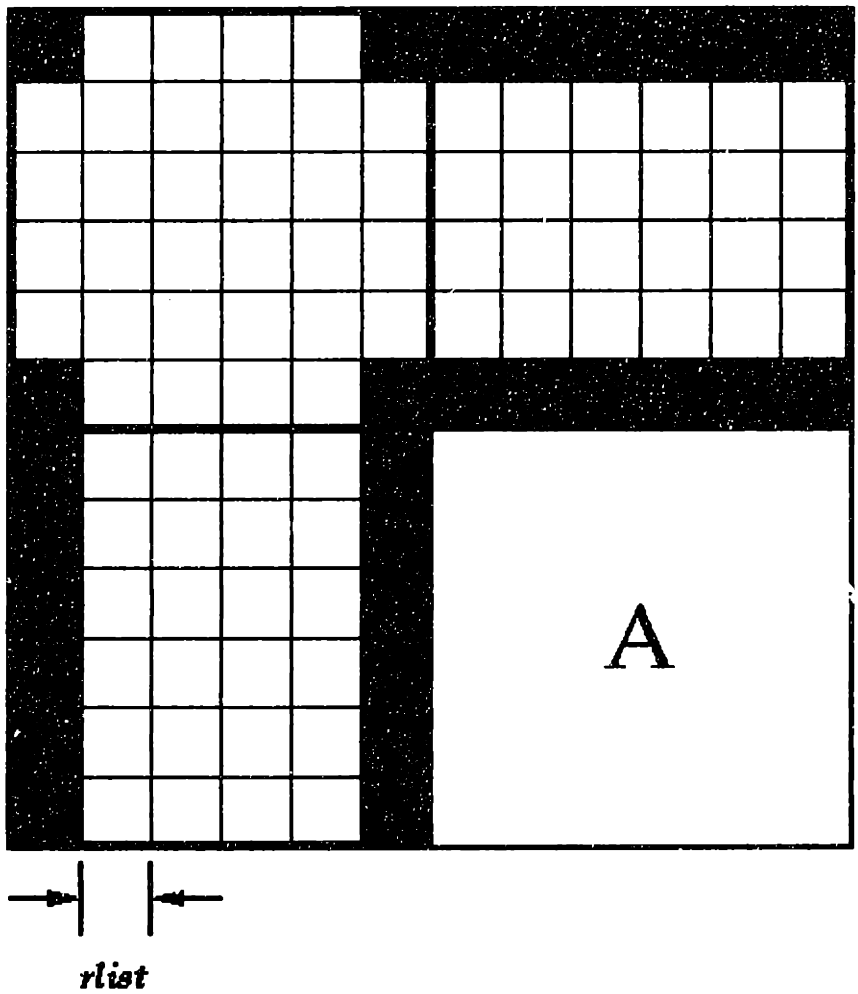

Figure 4-2: Node subdivision into cells reduces communication environment of node A to shaded regions. Note the effect of imposing periodic boundary conditions.

interaction term which is rewritten below as

$$f_3 = h(r_{ij}, r_{ik}, \theta_{jik}) + h(r_{ji}, r_{jk}, \theta_{ijk}) + h(r_{ki}, r_{kj}, \theta_{ikj}). \tag{4.1}$$

θ_{ijk} represents the angle between the position vectors $\mathbf{r_{ij}}$ and $\mathbf{r_{ik}}$. The function h is given by the expression

$$h(r_{ij}, r_{ik}, \theta_{jik}) = \lambda \exp[\gamma(r_{ij} - a)^{-1} + \gamma(r_{ik} - a)^{-1}] \times (\cos\theta_{jik} + 1/3)^2, \tag{4.2}$$

if $|\mathbf{r_{ij}}|$ and $|\mathbf{r_{ik}}|$ are within the cutoff distance, a, and 0 otherwise. Note that the SW potential does not constrain $|\mathbf{r_{jk}}|$ to be smaller than the cutoff distance, a. This results in two kinds of triplets that contribute to the overall energy of the system. Triplets of type A are defined as those where all three atom pairs, $[(i,j),(i,k),(j,k)]$, are within the cutoff distance of each other, while those of type B only have 2 atom pairs, $[(i,j),(i,k)]$, within the cutoff radius. Clearly, type A triplets are more compact and may span a maximum of 4 cells while those of type B may span a maximum of 9 cells in a 2-dimensional system. These cases are shown in Figure 4-3.

Distinguishing between these two types of triplets makes the use of generalized search algorithms, such as the one developed by Esselink *et al.* (Esselink and Hilbers, 1993; Esselink *et al.*, 1993) inefficient because the type B triplets require large cells and hence coarser discretizations, limiting the usefulness of the cell method. In the algorithm of Esselink et al., a many-body interaction is subject to the constraint that all atoms are within the cutoff radius of each other. Thus, for arbitrarily many interactions, the effective interaction range is equal to the pair interaction range.

4.3 Program Structure

The overall structure of the parallel molecular dynamics simulation is essentially unchanged with respect to the serial version. Once again, Newton's equations of motion for the set of N atoms are solved using Gear's 5^{th}-order predictor-corrector algorithm (Allen and Tildesley, 1992; Haile, 1992). The essential features of the program

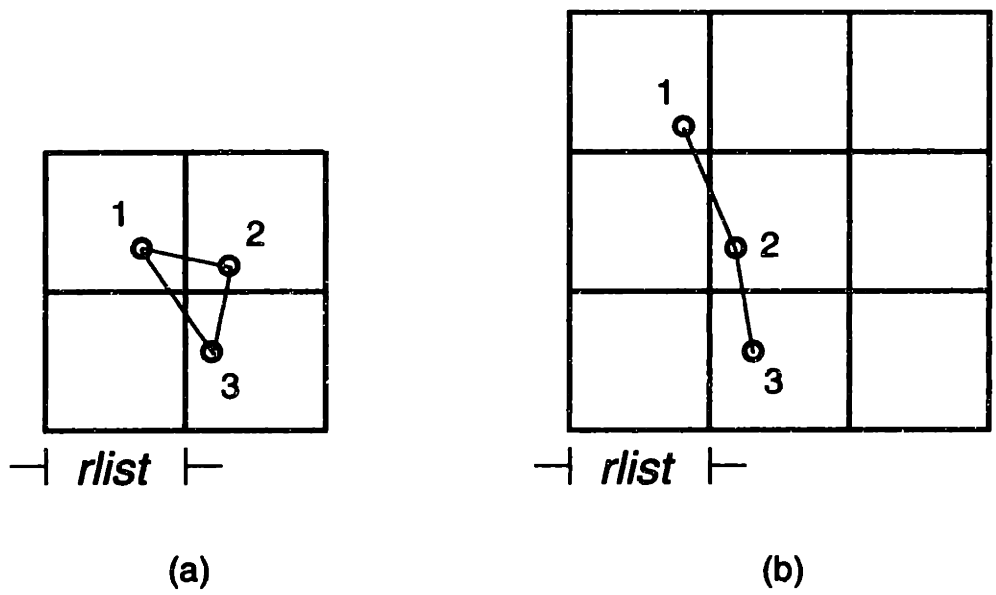

Figure 4-3: SW potential contains contributions from 2 different kinds of triplets (a) all 3 pair interactions within the cutoff distance (type A), and (b) only 2 pair interactions within the cutoff distance (type B).

structure are shown schematically in Figure 4-4. At each time step, estimates for the particle positions are calculated in the routine PREDCT. The potential and forces experienced by each particle based on the SW potential are computed in EVAL and the positions of the particles are corrected in the routine CORR using the forces calculated in EVAL. This sequence of steps is repeated at each time step. Periodically the neighbor lists (Allen and Tildesley, 1992) are updated using NLIST. The procedures PREDCT and CORR remain completely unchanged in the parallel implementation because these only require information about particles that reside locally on each processor. In contrast, EVAL and NLIST now require the positions of all particles that lie within the interaction range of all particles that reside locally on a node. Since such particles may lie on different nodes, particle information needs to be exchanged between processors; this is performed using the new subroutine COMMUN. The routine ALLOCATE is created to ensure that each particle is always assigned to the correct processor and is similar to the routine described by Esselink et al. (Esselink *et al.*, 1993). These subroutines are discussed in greater detail in the subsequent sections.

4.3.1 ALLOCATE

All particles are distributed among the N processors based on their instantaneous spatial location, so that each processor is responsible for particles in a certain *fixed* region of space. This kind of partitioning (geometric partitioning) minimizes communication overhead because particles that are spatially close will usually be assigned to the same node. However, since the particles are free to move, the list of particles which reside on each node must be updated periodically, typically every 10 - 20 time steps. This is accomplished in the subroutine ALLOCATE.

Within each processor, the routine ALLOCATE searches over the current list of particles, and determines whether each particle's coordinates are still within the node's boundaries. If not, the particle label and all its information (position, velocity, acceleration, and higher order derivatives) are transferred to the correct node. It is obviously not possible to determine *a priori* whether any particles will need to be

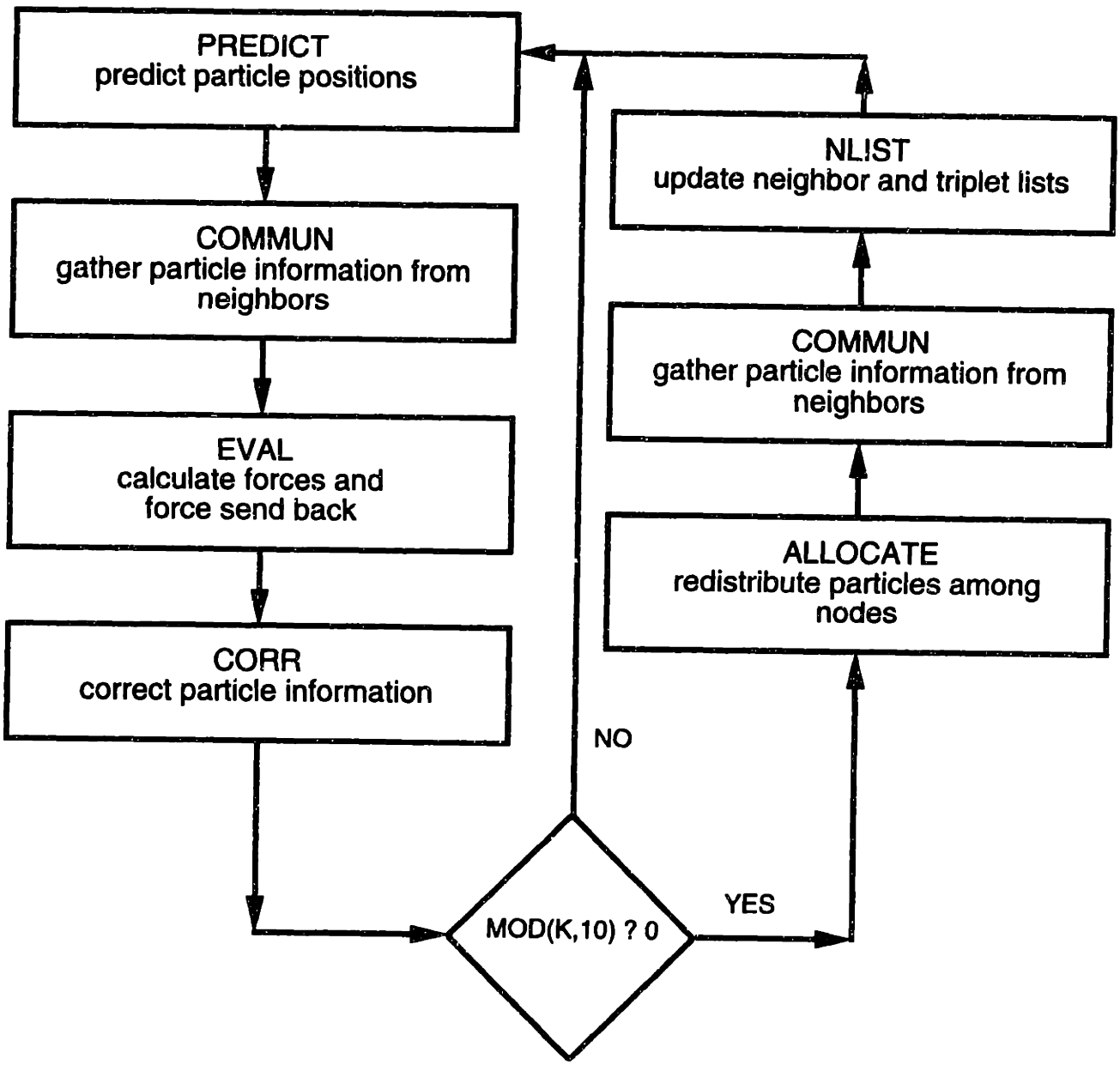

Figure 4-4: Schematic of program structure.

removed from a nodes memory, and typically there will be some instances that a particular processor will not require ALLOCATE to do anything. However, in order to avoid synchronization problems, ALLOCATE sends messages in a predetermined sequence to all neighboring processors, even if the message contains no information. The increased use of unnecessary message passing is justified here because of synchronicity requirements. In other words even if only two processors required a particle transfer, the rest would still have to wait anyway. Each node communicates with the other nodes in the same order to minimize waiting time; e.g., all nodes send to the right and receive from the left, etc.

4.3.2 NLIST

The subroutine NLIST generates two lists containing the neighbors of each particle, and a list of all triplets as in the serial version of the simulation code. As discussed in Section 3.4.1 the calculation of pair and triplet interactions are order N^2 and N^3 operations respectively, where N is the number of particles. The Verlet neighbor list (Allen and Tildesley, 1992) is used once again to generate the pair interaction lists. Traditionally, the size of the neighbor list is reduced by half by invoking Newton's second law of motion (NLM); the neighborhood list contains the atom pair (i, j) but not (j, i) because $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$. In a parallel MD simulation based on only pair interactions, NLM is implemented by only communicating with half the neighboring cells thus reducing the search space. However, for the SW potential, no efficient method was found to generate the triplet list if NLM was implemented. This limitation is caused by triplets of type B, as described in Section 2.2. All triplets, (i, j, k), are found by selecting a particle i, looping over its neighbors to get particle j, then looping over the neighbors of j to obtain particle k. Overcounted triplets are eliminated in a subsequent step. The computational overhead for generating the triplet list scales as $N_p N_n^2$, where N_p is the average number of particles on a single node, and N_n is the number of particles within a typical particles neighbor list.

In order to account for all possible neighbors of particles on a node, the particles in the first layer of cells on neighboring nodes (shaded region in Figure 4-2) must

be considered. To generate the triplet list, the neighbors of these particles also are needed, so that the neighbor list must also be computed for particles in the shaded region. The latter calculation requires information from the second cell layer adjacent to the shaded region. Thus, in order to find all triplets, information from *two* layers of surrounding cells must be transferred to each processor. This requirement imposes a heavy penalty on the communication overhead. For systems with up to 125 cells per processor, almost all the particles must be exchanged. Only in cases where each processor has more than 1000 cells does this fraction become less than 50 %.

4.3.3 COMMUN

COMMUN is the subroutine responsible for transferring particle information from the two cell layers closest to the surface of each processor to adjacent processors. COMMUN is structured so that each processor performs the same *send* and *receive*

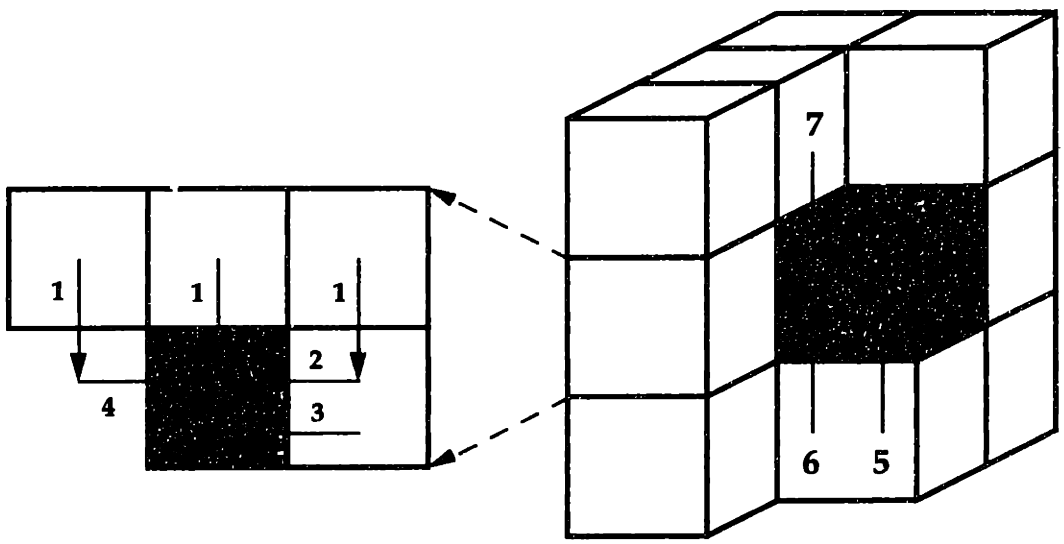

Figure 4-5: Communication strategy involves a sequence of seven steps. Each processor performs identical operations to minimize waiting time.

operations simultaneously so that no one processor has to wait for the others, as in ALLOCATE. The sequential steps performed in COMMUN are shown schematically

in Figure 4-5 and correspond to the communication strategy presented by Esselink et al. (Esselink *et al.*, 1993). For example, all nodes send information to the nodes below in step 1. At the same time all processors receive information from the nodes above. In step 2, all nodes send the left plane of cells to the node on their left and so on. In seven steps each processor has gathered particle information from its 13 neighbors. Since NLM is not implemented in this code, this process must be repeated for the remaining 13 neighbors. Once the layer of cells closest to the surface have been exchanged, the entire process is repeated for the second layer. The order in which the communication steps are performed is used to create pointer arrays which allow for rapid extraction of the particle information.

4.3.4 EVAL

The subroutine EVAL is responsible for the calculation of the potential and force. It consists of two sections; the energy and force computations, as in the serial implementation, and the communication of the forces, which is new. The pair potentials and forces are computed only if $i < j$, where i and j are the particle labels. This restriction eliminates double counting any pair interactions, and implies that only the force of particle i on particle j is computed explicitly. Suppose now that particle i exists on processor A, while particle j is on processor B. If $i < j$, then only A will calculate the interaction, and B will not have access to the new value of the force exerted on j, unless A passes this information to B. Each particle that is involved in a force evaluation is tagged and then sorted into message bins based on its coordinates, which indicate its destination. Once the sorting procedure is complete, the messages are sent out sequentially to the appropriate processors which then update their force data.

4.4 Results and Discussion

4.4.1 Algorithm Performance

The performance of the parallel MD code was evaluated by direct comparison with an identical serial version of the code running on a single processor of the IBM SP2. The parallel efficiency is defined as :

$$\eta(\%) = \frac{T_{serial}}{N_{nodes} \times T_{parallel}} \times 100 \tag{4.3}$$

where T_{serial} and $T_{parallel}$ are the CPU times required for the execution of the serial and parallel codes respectively. The efficiency computed as a function of number of particles is shown in Figure 4-6. The optimal number of cells was used for each configuration to ensure maximum efficiency, as described in Section 3.4.1. Eight processors (N=8) were used for all the parallel calculations. The performance efficiency computed only for the subroutine EVAL also is shown in Figure 4-6. The overall efficiency increases as the number of particles is increased and reflects an increase in the work load of each processor, relative to the communication load and associated synchronization time. For a system of 21,952 particles, the overall efficiency reaches 70 %. The force calculations in the serial and parallel version of EVAL are identical. However, as discussed in Section 4.3.4, the parallel version of the code requires that some of the forces be sent back to neighboring processors. This communication phase is responsible for the less than perfect efficiency of EVAL. The efficiency of EVAL increases from 65 % to about 80 % with increasing the number of particles because of the decreasing surface area to volume ratio with increasing number of particles per processor. The fraction of particles on a node that are involved in inter-processor communications thus becomes smaller to increase the efficiency of EVAL.

The contribution of each of the major subroutines to the overall simulation time are reported in Figure 4-7. The routine NLIST contributes only about 10 % to the total simulation time. The amount of time spent evaluating the forces scales linearly with the *total* number of particles. In contrast, COMMUN and the force communica-

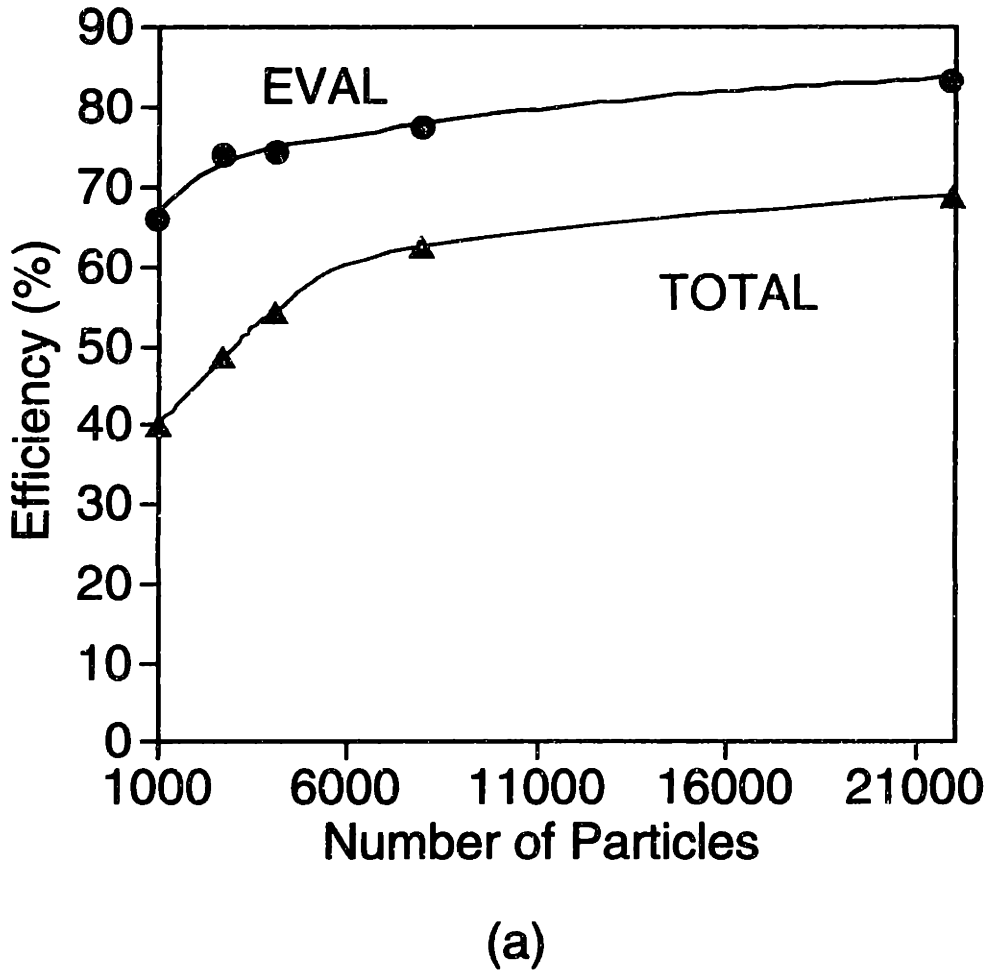

Figure 4-6: Efficiency determined via comparison to a serial version of the code.

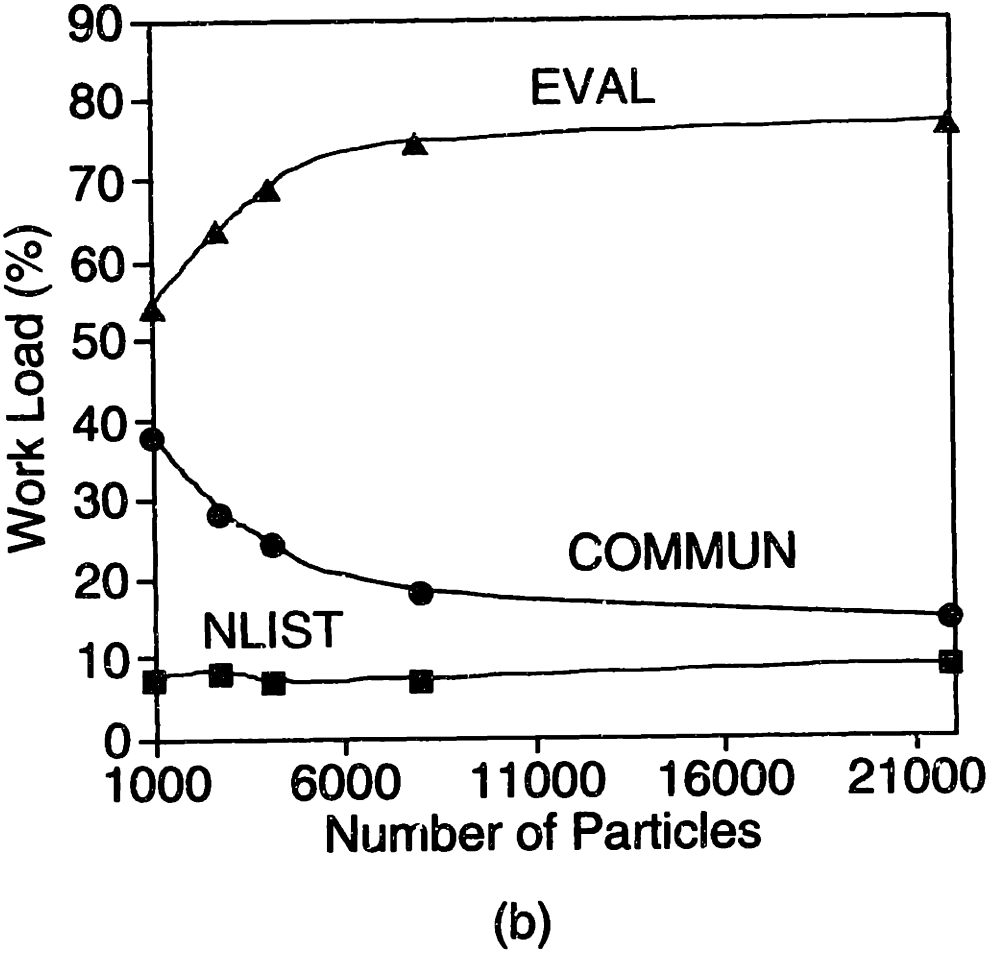

Figure 4-7: Work load distribution among various subroutines in the parallel code.

tions scale linearly with the number of particles *in the surface cells* on each node. As the number of particles per node increases, the communication overhead, as reflected by the time spent in COMMUN, becomes a smaller fraction of the overall load. Correspondingly, the contribution of the routine EVAL to the total computational load increases, as is evident in Figure 4-7.

The code was written for P^3 processors. Eight processors represents a special case. Due to periodic boundary conditions, the processor to the left and the right (or above and below) each node correspond to the same node. This results in duplication of information during communication. This duplication was not eliminated for the sake of program generality.

4.4.2 Thermodynamics of Point Defect Clusters

MD simulations were performed employing the constant pressure-constant temperature statistical ensemble (NPT). The temperature was fixed using velocity rescaling and the pressure was kept constant by using the piston method (Andersen, 1980). A simulation cell containing 2744 silicon atoms initially arranged in a perfect diamond lattice was used as the host lattice for all calculations described in this section.

There is no reliable experimental data suggesting a particular structure for even very small point defect clusters in silicon. A series of simulations was designed to allow the configuration of each cluster to equilibrate at a given temperature. To do this, the size of the defect increased (or decreased) by adding (or subtracting) one particle at a time. For example, in the case of interstitial clusters, the first simulation employed the perfect diamond lattice described above with one extra silicon atom placed in a tetrahedral position (Maroudas and Brown, 1993). The configuration of a single self-interstitial was equilibrated by a long-time simulation. Using this as a first configuration, another silicon atom was placed in the vicinity of the first defect, and the simulation was continued until equilibrium is reached. Equilibrium was defined as the state in which the energy no longer changed with time, and after the defects had clearly interacted with each other, as determined by observing a sudden large reduction in the energy of the system. The sequence of adding an interstitial and then

relaxing the extended system was continued until seven interstitials had been added to the original perfect lattice. The same sequence of simulations was performed for calculations with vacancies; here an atom is removed at each equilibration step.

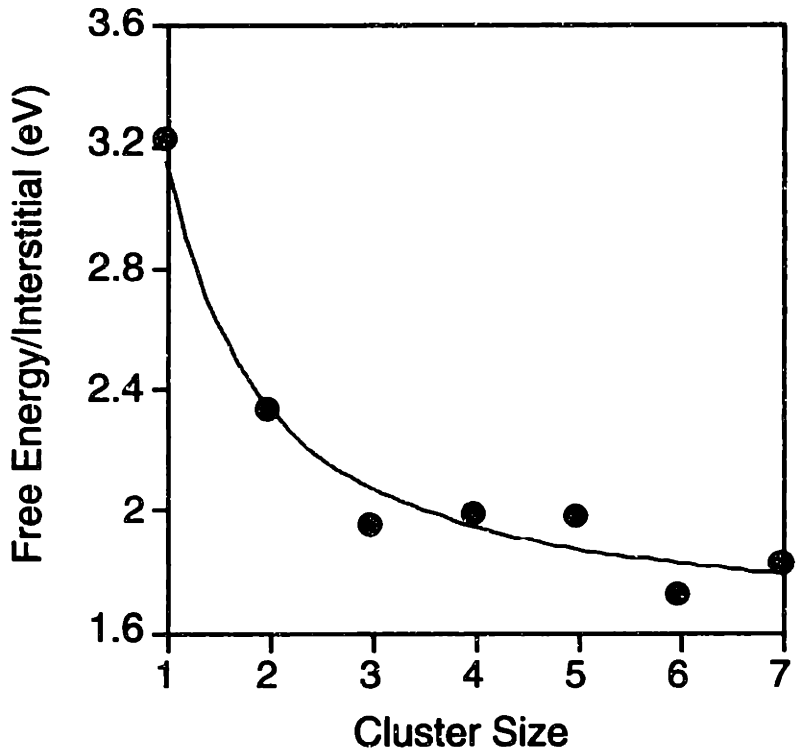

Figure 4-8: Free energy of formation per point defect for interstitial clusters as a function of cluster size.

The simulations were repeated at temperatures ranging from 500 K - 1600 K. At 500 K, the reference free energy of each cluster was computed using the Quasi-Harmonic Approximation (Lutsko *et al.*, 1988; Ashcroft and Mermin, 1979) in the same way as described in Section 3.5.1. The free energies across the entire temperature range were then computed using thermodynamic integration (Allen and Tildesley, 1992) as

$$\left(\frac{G}{kT}\right)_2 - \left(\frac{G}{kT}\right)_1 = \int_{T1}^{T2} \left(\frac{H}{kT}\right) \frac{dT}{T}. \tag{4.4}$$

where H is the *enthalpy* of the system. The free energy of formation per point defect in a cluster of size n, (G_n/n), as a function of cluster size is shown in Figure 4-8 for interstitial and Figure 4-9 for vacancy clusters at 1300 K. The data points represent the simulation results, while the lines are polynomial fits of the data and are only

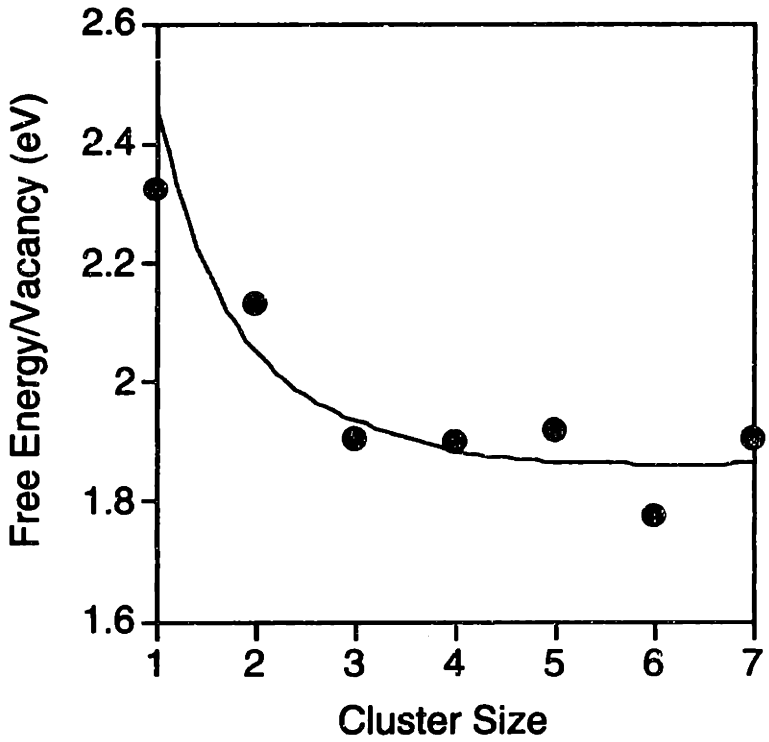

Figure 4-9: Free energy of formation per point defect for vacancy clusters as a function of cluster size.

meaningful at discrete cluster sizes. For both interstitial and vacancy clusters at this temperature, the SW potential predicts free energies of formation per point defect that rapidly asymptote to a final value, suggesting that clusters quickly reach a state that can be regarded to be macroscopic. The very size of these clusters is surprising; these clusters are far too small to have bulk and interfacial regions, as expected for clusters described by continuum theories.

The predicted free energy values were used to compute equilibrium distributions of clusters in systems containing a fixed numbers of either interstitials or vacancies according to

$$G_{TOTAL} = \sum_{i=1}^{7} n_i G_i^f - kT \left[Q \ln Q - (Q-\gamma)\ln(Q-\gamma) - \sum_{i=1}^{7} n_i \ln n_i \right] \qquad (4.5)$$

where γ is the total number of clusters, Q is the total number of point defects available for clustering, and n_i is the number of clusters of size i. In equation 4.5, the first and second terms represent the vibrational and configurational contributions to the free

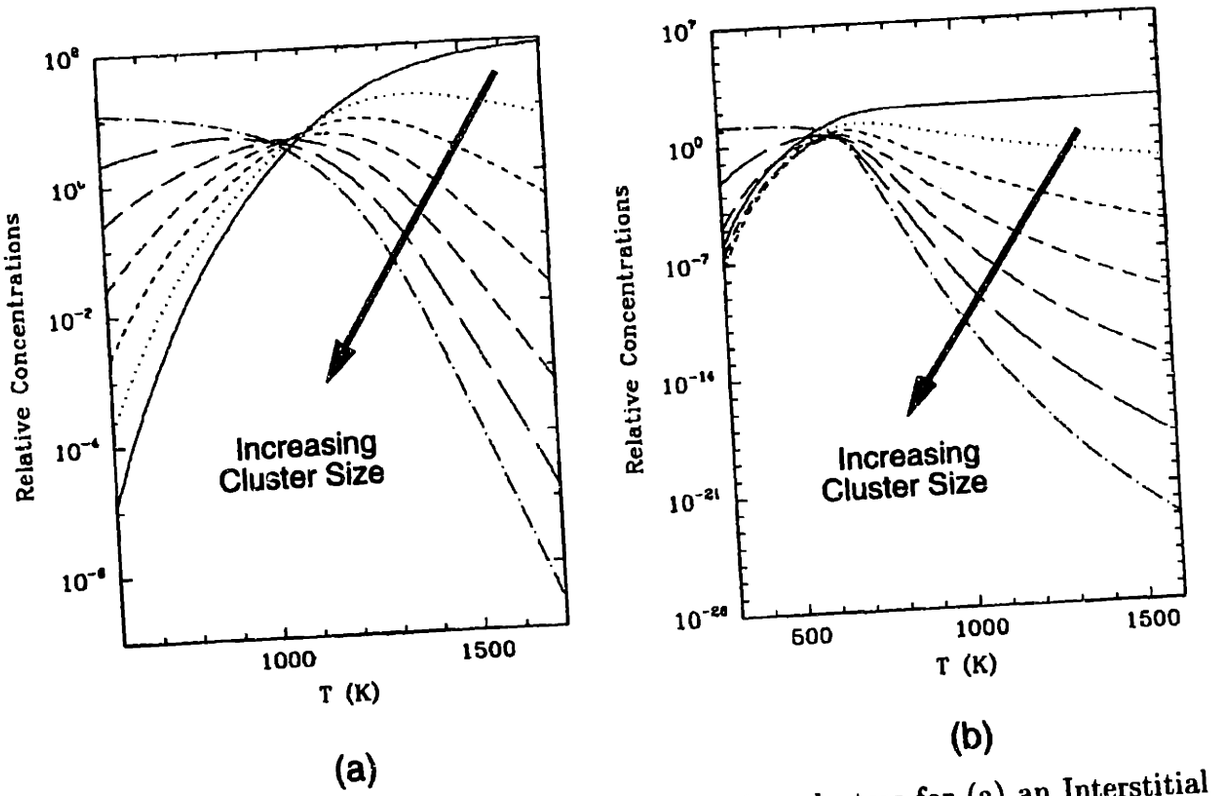

Figure 4-10: Equilibrium distributions of point defect clusters for (a) an Interstitial system, and (b) a Vacancy system.

energy respectively. The configurational contribution was derived by computing the product of the number of ways that γ clusters can be distributed over N lattice sites, $N!/(N-\gamma)!\gamma!$, and the number of ways the *distinguishable* clusters can be arranged among themselves.

For each point defect species, the distribution of 100 point defects among the seven cluster sizes was determined by minimizing the free energy (G_{TOTAL}) of the system subject to the constraint that the total number of point defects (Q) remains constant so that

$$Q = \sum_{i=1}^{7} i \times n_i \tag{4.6}$$

The free energy minimization was performed for temperatures ranging from 400 K to 1700 K using a Lagrange multiplier to satisfy the point defect constraint which gives

$$G_{TOTAL} = \sum_{i=1}^{7} n_i G_i^f - kT \left[Q \ln Q - (Q-\gamma)\ln(Q-\gamma) - \sum_{i=1}^{7} n_i \ln n_i \right] + \lambda \sum_{i=1}^{7} i n_i \tag{4.7}$$

where λ is the Lagrange multiplier.

The results of these calculations are shown in Figure 4-10. For both interstitial and vacancy clusters, the dominant species shifts from single point defects at high temperature to the largest clusters at low temperatures. This trend is readily explained by the fact that the free energy release due to aggregation is greatest at lower temperatures. The main difference in the two distributions is the critical temperature at which concentrations of clusters become appreciable. In the case of the interstitial aggregates, clusters begin to be favored at a temperature of about 1050 K, whereas for vacancies, this temperature is about 600 K. This result is due to the larger free energy release during interstitial agglomeration predicted by the results in Figures 4-8 and 4-9.

4.4.3 Simulation of Large Vacancy Clusters

A perfect silicon lattice containing 46,656 atoms was used to create a spherical void consisting of 275 vacancies. The discussion in Section 2.2.2 of Chapter 2 highlighted

the current technological interest in void defects in crystalline silicon due to their detrimental effect on device performance. The single vacancy results of Chapter 3 did not appear to be quantitatively accurate, but cluster simulations can still provide qualitatively useful information.

The void and surrounding lattice were allowed to relax at 1500 K while the energy and shape of the defect structure were monitored. The shape was determined by computing the radius of gyration in each of the three cartesian dimensions. The simulation shows that the void is structurally stable for the length of time simulated, about 0.2 *ns*. Neither the shape nor the energy changed appreciably once the surrounding lattice was relaxed. The enthalpy of formation for the void at 1500 K was calculated to be 0.5 eV per vacancy. Entropy contributions were not computed as for the smaller defects because of memory limitation problems with the QHA approach. The estimate for the enthalpy of formation per vacancy for the 275 vacancy void is far lower than that for the smaller vacancy clusters (2.6-1.9 eV per vacancy for clusters of up to seven vacancies). The free energy calculations for the small clusters described in the previous section suggested that a continuum limit in some sense had been approached by clusters containing more than four vacancies (see Figure 4-9).

However, if the concept of dangling bonds is used (Pantelides, 1988), it can be shown that the enthalpy of formation also approaches a continuum limit in a similar fashion to the free energy result.

The creation of a vacancy results in the neighboring atoms having lower coordination numbers and thus unbonded molecular orbitals. Although the SW potential is not capable of explicitly accounting for such effects, its empirical form imposes strong energy penalties for deviations from a complete tetrahedral lattice. The number of dangling bonds for each of the vacancy clusters was computed and the formation enthalpy distributed among them. Figure 4-11 shows a plot of the formation enthalpy per dangling bond as a function of cluster size. The rapid transition to a limiting value is clearly seen for clusters containing four or more vacancies all the way up to the 275 vacancy cluster. This result implies that the enthalpy of formation is more directly related to the *interface* between the cluster and the lattice than to the volume

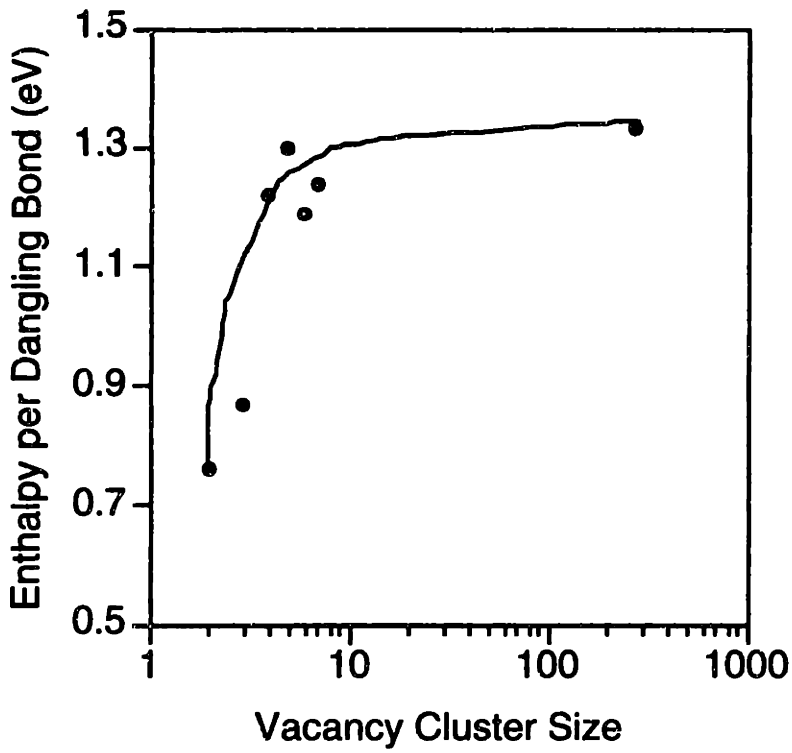

Figure 4-11: Enthalpy of formation per dangling bond as a function of cluster size.

of the cluster. In the case of voids this is clearly expected, as no forces can exist in the empty space within the void. In the case of large voids, analogies to surfaces can be made. This is particularly useful since silicon surfaces have been studied extensively. Nakamura et al. (1996) have estimated the silicon $< 111 >$ surface energy to be about 1.0 J/m^2, which is an average of several values that lie within 0.2 J/m^2 of each other.

$$E_{111} = 28.0\pi \left(\frac{6.0 \times 10^{-2} N_V}{4\pi} \right)^{2/3} \quad eV \tag{4.8}$$

for the energy of a void containing N_V vacancies and whose interface is similar to a $< 111 >$-oriented surface, or about 110 eV for a spherical 275 vacancy defect. This value is in surprisingly good qualitative agreement with the Stillinger-Weber MD result of about 135 eV. The assumption that large voids possess $< 111 >$ surface characteristics is reasonable due to the fact that this surface possesses the highest density and therefore the least number of dangling bonds. Figure 4-12 shows that the simulation does in fact predict a faceted void with the left sides clearly lying

on < 111 > planes, even though the initial configuration was spherical. Recent

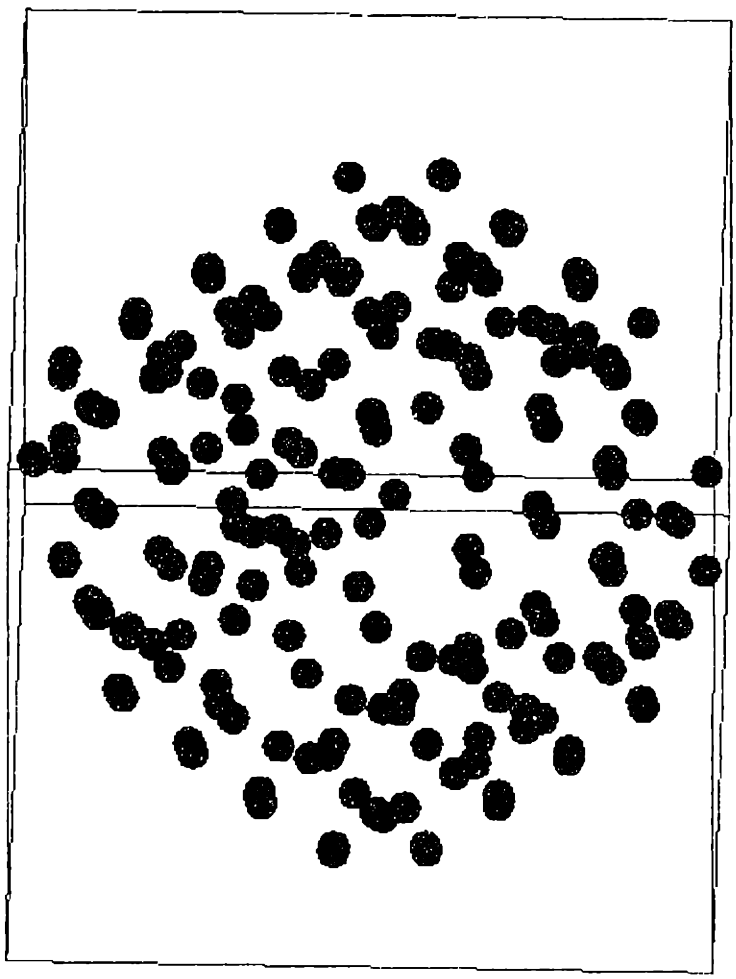

Figure 4-12: Atoms in the vicinity of the 275-vacancy void showing the faceted edges of the defect.

experimental studies (Itsumi *et al.*, 1995; Itsumi *et al.*, 1996) have shown that voids appearing in crystals grown by the Czochralski method are octahedral in shape, with sides lying along < 111 > planes. Further simulations with larger systems will give more direct visualization for the < 111 > faceting shown in Figure 4-12 because of better resolution. One of the more difficult aspects of characterizing simulated voids is actually visualizing them. Figure 4-12 shows the atoms that are adjacent to vacant

lattice sites, which is an indirect method of showing the void. A possible alternative is the construction of a surface based on the coordinates of these atoms so as to resolve the shape of the void more clearly.

4.4.4 Modeling Interstitial Aggregation Kinetics

The aggregation kinetics of point defects in silicon are the focus of much interest from the point of view of microdefect engineering. Quantitative estimates for the rate of aggregation of point defects as a function of temperature will ultimately allow for the detailed modeling of microdefect formation in a growing crystal. There are several fundamental problems with attempting to investigate these issues with atomistic simulation. Experimentally observed microdefects are at least 50 nm in size and can be as large as one micron. For stacking-fault type defects which are essentially two dimensional defects, these contain $O(10^5 - 10^6)$ point defects. In the case of three dimensional defects such as the voids mentioned previously the observed sizes can contain upto $O(10^9)$ point defects. Simulations with this many particles are not feasible especially if kinetic information is required which generally requires simulation over lengthy time scales.

Despite these difficulties, it is possible to investigate the initial nucleation process with far smaller systems. The lengthy time scales needed for homogeneous nucleation to proceed can be shortened by imposing very high supersaturations of point defects. A perfect diamond lattice of 46,656 atoms was used to host 64 additional interstitial atoms placed at regularly spaced intervals in tetrahedral locations. The NPT statistical ensemble was used and the pressure was set to zero as in all previous calculations. The system was allowed to relax without constraints and several properties were monitored in time. At certain intervals (about 10,000 time steps), the atomic coordinates were rapidly quenched to 0 K. Atoms involved in defects were isolated by subtracting an image of the perfect crystal from the quenched coordinates. Individual clusters were then identified using an algorithm developed by von Gottberg (von Gottberg, 1997). The algorithm is based upon the construction of neighbor lists for the system of particles that are associated with clusters. The form of the neighbor list is similar

to the one described in Section 4.3.2. An initial cluster is defined by assigning the first particle to it. The neighbors of this particle are extracted from the neighbor list and assigned to the same cluster. The algorithm recursively then searches for the neighbors of each of these new particles until no more particles can be assigned to the first cluster in this way. The next cluster is defined by the first particle that does not belong to the previous one. This procedure allows for the direct monitoring of the aggregate size distribution as a function of time.

Figure 4-13 shows the evolution of the average cluster size. The process of aggregation is clearly visible. The initial average size of four corresponds to single self-

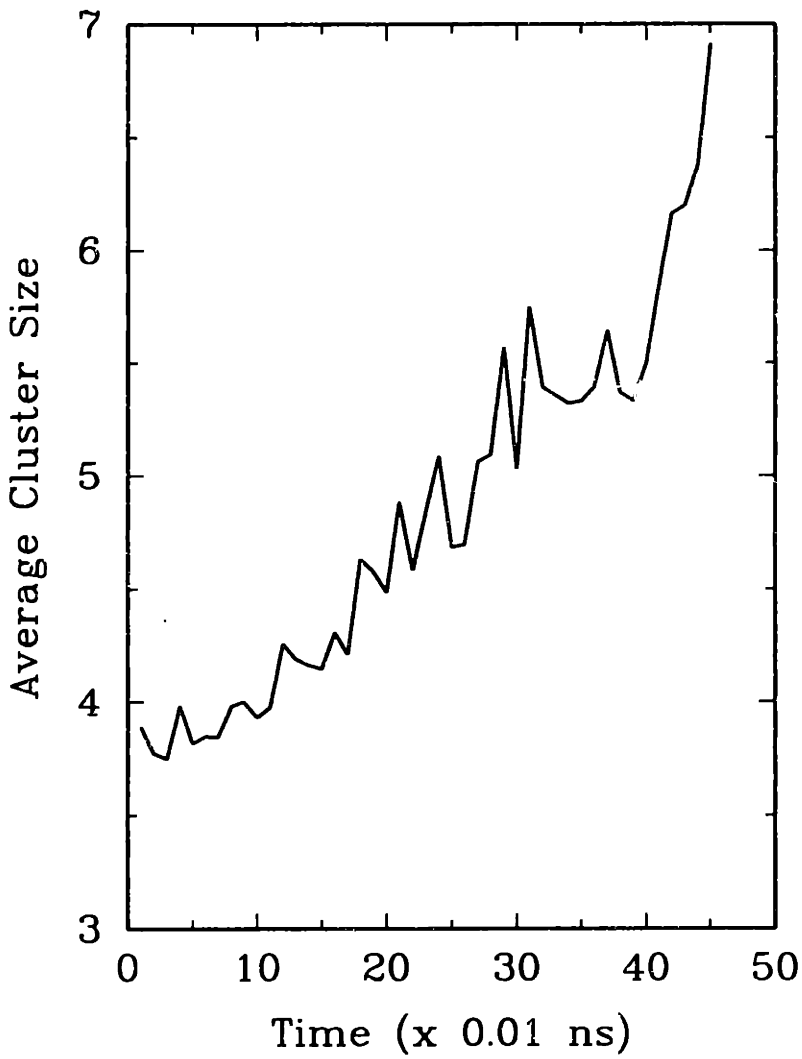

Figure 4-13: Evolution of the average cluster size in time.

interstitials; the value reflects the dumbbell configuration described in Section 3.6.3

of Chapter 3. The average increases steadily to about 6 and then levels out before suddenly jumping to 7. The final jump corresponds to the coalescence of two moderately sized clusters. The occurrence of this event is due to the very high cluster concentration in the simulation cell; for crystals with microdefect concentrations of about 10^8 cm^{-3}, this is an unlikely mechanism for cluster growth. The coalescence event can be more explicitly viewed in Figure 4-14 which shows the evolution of the maximum cluster size. While most increases (and decreases) in maximum cluster

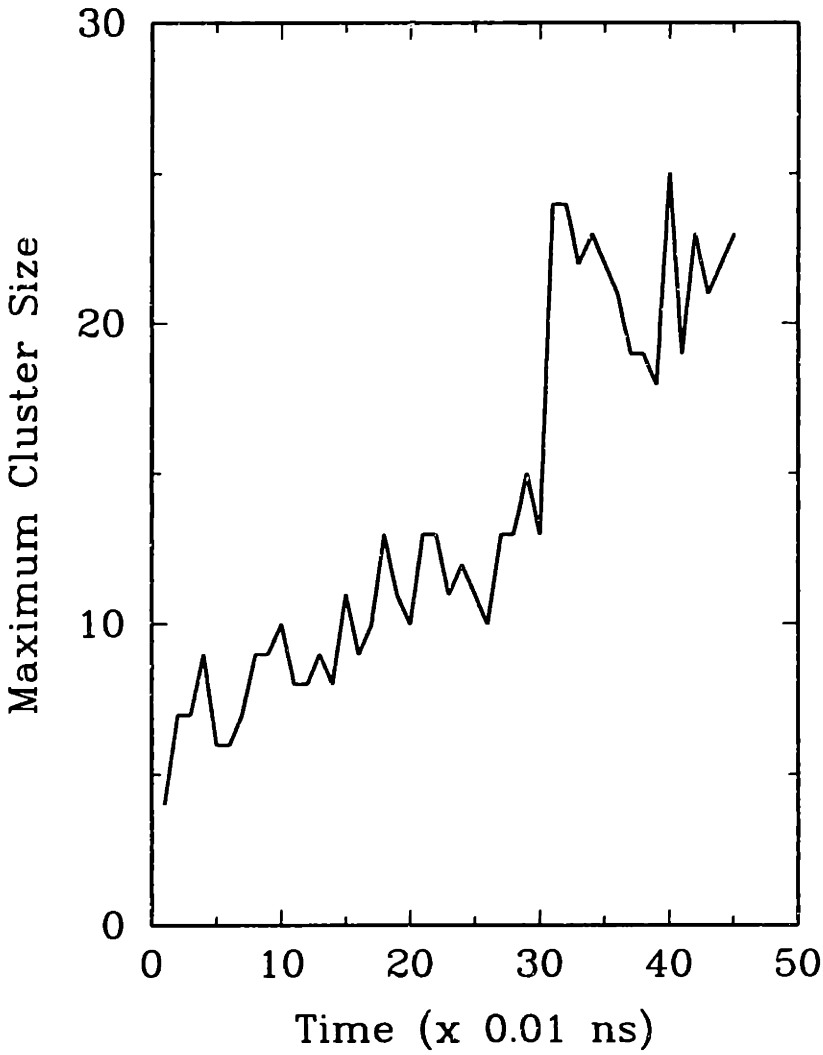

Figure 4-14: Evolution of the maximum cluster size in time.

size take place in jumps of two or three, some are much larger like the transitions at $t = 30$ and $t = 40$. In both Figures 4-13 and 4-14, the system appears to be evolving dynamically, and not monotonically; while clusters are clearly growing, not

every event contributes to this trend. This implies that the clusters are fairly loosely bound with respect to their surroundings and that a particular interstitial can move from one cluster to the other during the growth process, which allows the system to reach equilibrium much more quickly.

The process of cluster coalescence in this highly concentrated system occurs by point defect diffusion along the $< 110 >$ directions. Interstitial diffusion is known to occur along these directions because of the structure of the equilibrium interstitial which itself is oriented along this direction. Particles entering or leaving a cluster will therefore most likely do so along one of these directions.

Cluster interactions with other clusters makes analysis of coarsening kinetics more complicated. Many theories have been developed to predict so-called scaling laws for estimating the evolution of cluster size as a function of time. The first of these was the seminal work of Lifshitz, Slyozov, and Wagner (1971) who developed what is now known as the LSW theory. This theory predicts that cluster size evolves in time as $t^{1/3}$ which is found to agree well with experiment. The theory is based on the diffusion current of monomer to a cluster across the cluster boundary:

$$j = D\left(\frac{\partial C}{\partial r}\right)_{r=R} = \frac{D}{R}(C - C_R) = \frac{D}{R}\left(\Delta - \frac{\alpha}{R}\right) = \frac{dR}{dt}, \tag{4.9}$$

where C_R is the equilibrium concentration of monomer at the surface of the cluster, R is the radius of the cluster, and Δ is the supersaturation of the surrounding matrix defined as $C - C_{bulk}$. The quantity *alpha* is defined by,

$$C_R = C_{bulk} + \frac{2\sigma V C_{bulk}}{kTR} = C_{bulk} + \frac{\alpha}{R}. \tag{4.10}$$

In eq. (4.10), the equilibrium concentration of monomers at the cluster surface is not equal to the bulk equilibrium value because the free energy of the cluster which is a function of its surface tension, σ must be considered. Equation (4.9) is an expression for the rate of growth of a cluster in time and introduces the concept of a critical radius as defined by classical nucleation theory. Clusters that are larger than α/Δ will tend to grow ($dR/dt > 0$) while clusters smaller than this value will shrink and

disappear.

However, the $t^{1/3}$ scaling of cluster size with time is only observed in weakly supersaturated solutions that lead to a dilute concentration of clusters which are too far away to interact with each other and occupy a negligible fraction of the total volume. The LSW theory was extended to account for cluster interactions but not to account for finite cluster volume. Clusters can interact either by their strain fields or by overlapping diffusion layers. In systems where cluster growth proceeds rapidly as compared to monomer diffusion, depletion layers build up around the cluster, which have the effect of increasing the effective cluster interaction radius.

The salient point of the above discussion is that it is very difficult to characterize quantitatively the cluster growth kinetics for a very concentrated system such as the one described in this section. In general, we would not expect to see $t^{1/3}$ growth for these systems. More dilute systems with larger host lattices are needed before we can compute directly the aggregation kinetics of interstitial (or vacancy aggregates). While such information is extremely useful for modeling microdefect formation, the current code will have to be either run on a larger machine or significantly improved before such runs can be feasibly performed.

4.5 Conclusions

The parallel implementation of the molecular dynamics algorithm presented here makes possible large-scale simulations of microdefects using the three-body Stillinger-Weber interatomic potential. Although greater than 60 %, the parallel efficiency is hampered by the need to communicate two shells of cells between processors for the calculation of the three-body interactions. This communication is equivalent to reducing the discretization of the volume assigned to an individual processor and increasing the time spent in message passing relative to force and energy evaluation. Large efficiencies are obtained by increasing the number of cells assigned to each processor, with a concurrent increase in the number of particles or atoms in the simulation.

The calculations presented here for the free energies of small clusters of self-interstitials and vacancies in silicon are the first exploration of the dependence of the free energy of formation on cluster size, an important property in any model for microdefect precipitation. For $500 < T < 1600$ K, the SW potential predicts greater driving forces for the formation of interstitial clusters as compared to vacancy clusters. For both types of clusters, the driving force for a cluster of specific size increases with decreasing temperature shifting the equilibrium distribution towards the largest clusters at low temperatures. A significant result is the prediction that free energies of formation per atom reach asymptotic values for relatively small clusters, suggesting that it is not necessary to perform these detailed atomistic calculations for very large clusters. The free energies of formation for very large clusters can be estimated by simple extrapolation of the current results, as long as the morphology of the cluster is unchanged.

In this framework, these results are a first step toward understanding these complex systems. These results are of course subject to the limitations of the Stillinger-Weber interatomic potential (Balamane *et al.*, 1992). The SW potential has been shown to provide very good agreement with both experiment and *ab-initio* electronic structure calculations for the thermodynamics of the self-interstitial (Blöchl *et al.*, 1993). The extrapolation of this agreement to clusters of interstitials cannot be justified rigorously but other methods cannot as of yet model such systems.

The case for the vacancy calculations is less solid where there is significant evidence that electronic structure effects are important. These effects are observed in ab-initio calculations (Song *et al.*, 1993) as Jahn-Teller distortion. However, it is not yet practical to use any potential that is significantly more complex than the SW potential in simulations with large numbers of atoms. Despite the limitations of the Stillinger-Weber potential in modeling vacancy defects and surfaces (Balamane *et al.*, 1992), the simulation is able to give a qualitatively reasonable estimate for the energy of a void microdefect. Increases in computer speeds is beginning to make tight-binding calculations of cluster thermodynamics and structure more feasible. There is already some work on describing the thermodynamics of small point defect clusters appearing

in the literature (Rasband *et al.*, 1996b) and these are undoubtedly paving the way for future, larger calculations.

Chapter 5

Application of Microdefect Modeling to the Oxidation-Induced Stacking-Fault Ring in Czochralski Silicon

5.1 Introduction

One of the most highly studied microdefect structures in silicon is the ring-like distribution of oxidation-induced stacking faults (OSF-ring) that appears after oxidation of wafers grown by the Czochralski method (CZ). The OSF-ring is seen at a radius that depends exclusively on the growth conditions (von Ammon *et al.*, 1995; Hasebe *et al.*, 1989) e.g. the temperature field in the crystal and the growth rate; a typical example of OSF-ring response to growth rate is shown in Figure 5-1. The OSF-ring has been studied extensively for several reasons. The first is the ease of its delineation by oxidation of the wafer, which makes it readily visible. Furthermore, the OSF-ring exhibits unique and reproducible dynamics where it changes radial position depending on the crystal growth rate and the thermal environment during growth. These dynamics provide an excellent database for modeling. Numerous studies indi-

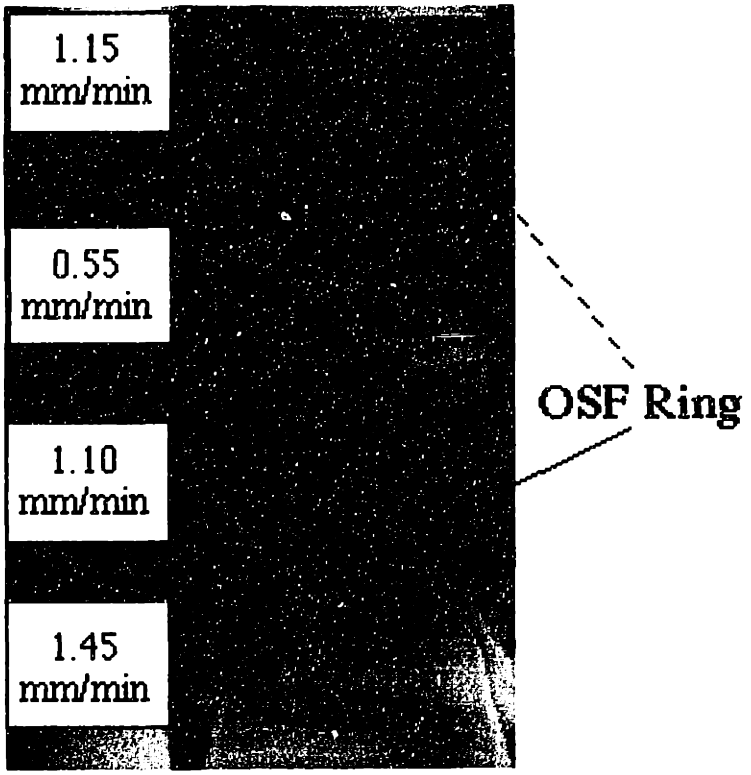

Figure 5-1: Oxidation-Induced Stacking-Fault Ring response to changes in crystal growth rate (Dornberger and von Ammon, 1996)

.

cate that microdefects which lie inside the ring region are formed by vacancies while those that lie outside the ring are self-interstitial in nature. In other words, the ring can be regarded as a visible boundary between two regions, each dominated by one of the two native point defect species. This observation provides a link between the location of the OSF-ring and the point defect distribution pattern within a crystal grown by the CZ method. Finally and most importantly, silicon wafers can be used for device fabrication only if the crystals are grown entirely and reproducibly in the vacancy rich regime, where the OSF-ring is located at the edge of the wafer.

A major technological challenge is to be able to grow the crystal in the vacancy dominated regime, while at the same time minimizing the number of vacancy related microdefects in the as-grown crystal. This objective can only be attained if the processes that control the location of the OSF-ring and control point defect aggregation are understood. We already know that microdefect densities are sensitively dependent on the growth conditions; however, there is great demand for a quantitative model which can *predict* these conditions.

Recently, several investigators have attempted to use point defect dynamic properties to develop continuum-scale models aimed at fully determining the point defect fields within a growing single-crystal in the hope of being able to predict at least qualitatively what growth parameters are responsible for OSF-ring dynamics (Brown *et al.*, 1994; Habu *et al.*, 1994; Wijaranakula, 1993). However, until now, no one model has been able to explain OSF-ring behavior consistently across a wide range of growth conditions. Reasons for this failure stem from two possible areas. Previous studies (Habu *et al.*, 1994) have shown that the results are always very sensitive on the point defect properties employed in the numerical simulations and poor estimates for the point defect properties lead to widely varying defect profiles. The other possible source of error is that the physics incorporated into the model, such as crystal geometry, thermal fields, and boundary conditions, are incorrect or oversimplified.

This chapter presents a point defect model that is capable of predicting OSF-ring dynamics over a wide range of experimental configurations. A two-dimensional finite element method is used to compute steady-state point defect concentrations.

Point defect properties are determined by a combination of atomistic simulation as described in Chapter 3, and fitting to experimental OSF-ring dynamics. The fitting process ensures that the best estimate for the point defect parameters is employed within the model. The test for validity of this approach is the models ability to *predict* OSF-ring dynamics for growth conditions very different to the one used in the fitting procedure by applying the model to crystals of different radii grown in different thermal environments. A comprehensive theoretical treatment of the governing equations is presented in Section 5.6 that provides an internally consistent picture for a mechanism for OSF-ring dynamics. The analysis is based on a two-scale asymptotic treatment of the coupled vacancy and self-interstitial conservation equations, resulting in an analytical expression for the location of the OSF-ring for a given growth environment.

5.2 Current Understanding of OSF-ring Dynamics

5.2.1 Experimental Studies

The OSF-ring has been studied extensively using experiments designed to characterize its behavior as a function of crystal growth operating conditions. Studies by Abe et al. (1981) and Hasebe et al. (1989) found that the *radius* of the OSF-ring depended exclusively on the crystal growth conditions and not the subsequent thermo-chemical annealing, even though the OSF-ring was not visible with X-ray topography until thermal oxidation was performed. Freeland et al. (1977) used transmission electron microscopy (TEM) tc show that the OSF-ring contained extrinsic (interstitial-type) stacking-fault loops lying on $< 111 >$ crystallographic planes, that grow isotropically from a nucleus containing oxygen, postulated to be an oxygen precipitate. These studies led to the conclusion that oxygen precipitates first form during CZ growth which then provide the nuclei for interstitial aggregation in stacking-faults during the steam oxidation of the wafers. Steam oxidation of silicon wafers is known to result in

the formation of an SiO_2 surface on the wafer, which acts as a source of interstitials into the bulk through a strain releasing mechanism (Hasebe *et al.*, 1989).

A series of recent experiments have characterized in detail the dynamics of the OSF-ring with respect to changes in the crystal pull rate and heat shield configurations in a CZ system on the radial location of the ring (Dornberger and von Ammon, 1996). The thermal environment was set for each experiment by adjusting the heat shield configuration in the CZ system. During growth, the crystal pull rate was varied continuously across a range known to lead to OSF-ring formation. The temperature field in the crystal for each thermal environment and pull rate was estimated by numerical simulation using an axisymmetric finite element analysis (Dupret *et al.*, 1990). It was found empirically that the variation in the temperature field is well characterized by the axial temperature gradient in the crystal at the melt/solid interface, $G(r)$, where $0 \leq r \leq R$ is the radial location in a crystal of radius R. The experiments showed that the location of the OSF-ring, $r = R_{OSF}$, correlates very well with a constant value for the ratio of the pull rate (V) to the interfacial axial temperature gradient, $G(R_{OSF})$. This value, $V/G(R_{OSF})$ which is referred to as the critical value was observed to be constant across a wide range of pull rates and heat transfer configurations.

Experimental data on the position of the OSF-ring as a function of pull rate and thermal configuration are reproduced in Figure 5-2. The correlation,

$$(V/G(R_{OSF}))_{crit} \approx 1.34 \times 10^{-3} \ cm^2/minK \tag{5.1}$$

also is plotted. A major goal of this chapter is to provide rigorous understanding of this correlation in the context of point defect dynamics.

5.2.2 Linking OSF-ring behavior to Point Defect Dynamics

Although experimental observations of silicon are focused on microdefects of size 50 nm or larger, it is believed that the critical step in differentiating between self-interstitial and vacancy dominated microdefects is the formation of point defect con-

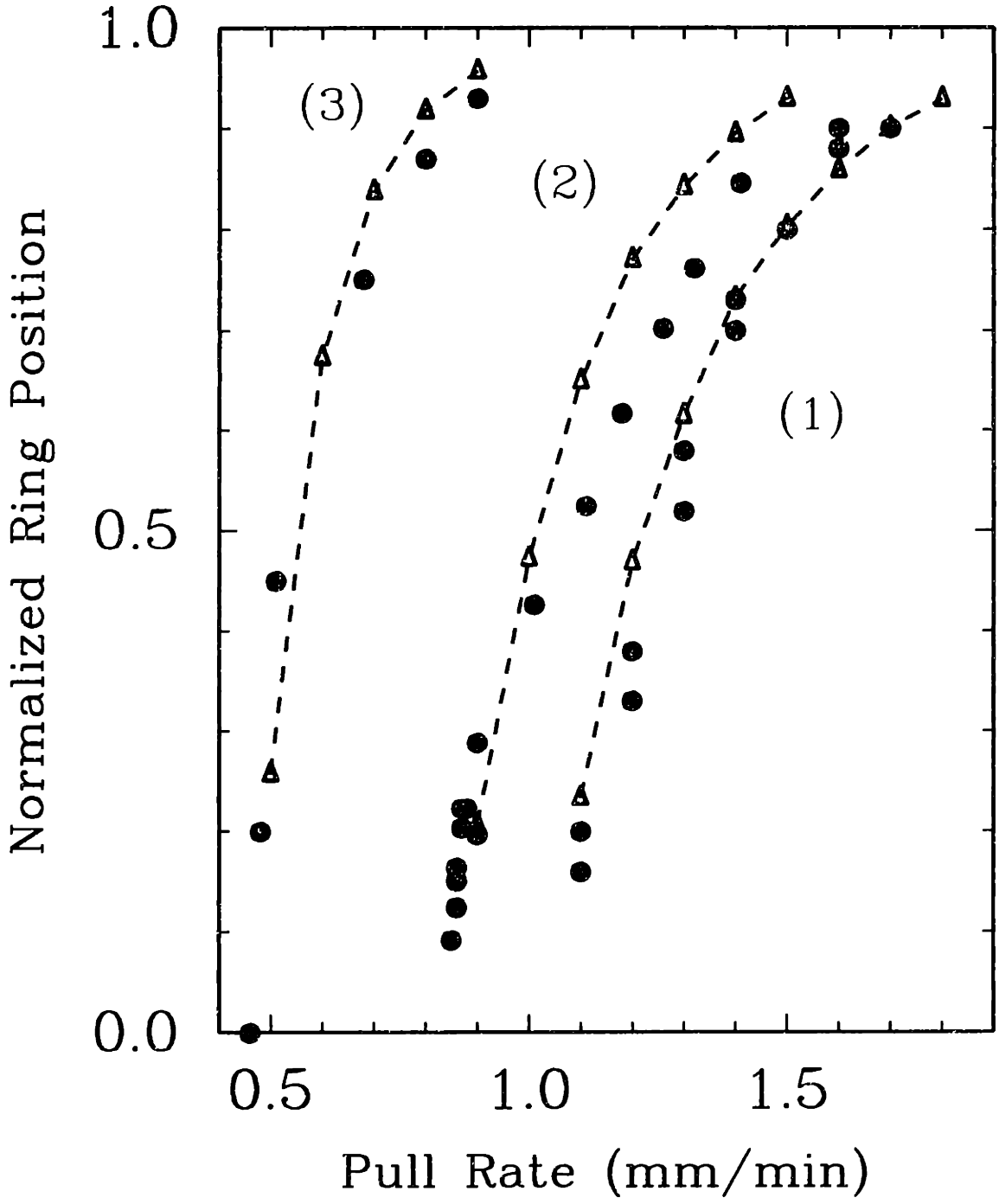

Figure 5-2: OSF-Ring response to changes in crystal growth rate (Dornberger and von Ammon, 1996): (1) HS1/4", (2) HS1/6", (3) HS2/8"; • = experimental results; $\triangle$ = correlation for $(V/G(R_{OSF}))_{crit} \approx 1.34 \times 10^{-3}\ cm^2/minK$

.

centrations, $C_I(r,z)$ and $C_V(r,z)$, respectively, where (r,z) describes a point in the crystal in the cylindrical coordinate system shown in Figure 5-3. This model is based

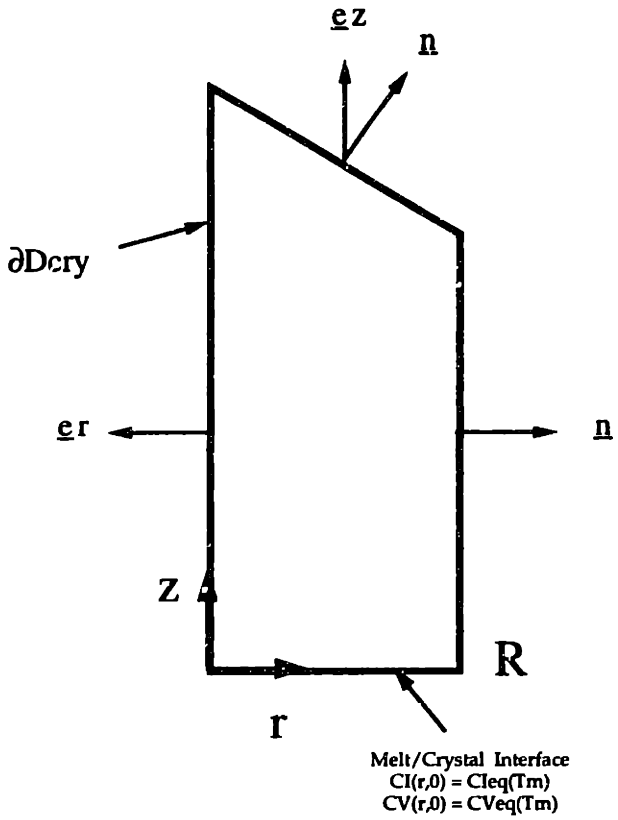

Figure 5-3: Cylindrical coordinate system for describing the axisymmetric crystal during growth from the melt.

on the hypothesis that microdefects from either point defect type grow at rates that are proportional to the driving force for point defect aggregation due to supersaturation of the surrounding bulk crystal and was first used in a semi-quantitative analysis by Voronkov (1982). Voronkov's contributions are discussed in Section 5.2.3.

Consider a growing crystal into which point defects are incorporated at the melt/solid interface. In addition to the convection of the point defects by crystal translation, they are also transported by diffusion and are either formed or depleted by the reversible reaction described in Chapter 2. As the temperature cools in the upper regions of the crystal, the thermally activated diffusion and reaction processes decay exponentially leading to a "frozen-in" defect profile that is simply translated in the direction of growth. If one considers the final frozen-in point defect concentrations in the cold upper portion of the crystal, the above picture can be used to predict

the dominant microdefect nuclei in the as-grown crystal. As a fixed point within the crystal cools during growth, large supersaturations of both point defect species are created locally due to the exponentially falling equilibrium concentrations; this supersaturation is released via transport, recombination and aggregation processes.

The validity of this picture in explaining microdefect dependence on crystal growth conditions relies on the presence of rapid point defect recombination at high temperature. The effect of fast recombination in the high temperature portion of the crystal is to drive the less concentrated species to annihilation while allowing the other point defect type to exist in large supersaturation. The annihilated species is driven to low concentrations relative to its equilibrium level at a given temperature and does not aggregate significantly, while the dominant species must form clusters to reduce its supersaturation because point defect recombination is no longer viable.

According to this interpretation, one can define a quantity, Δ, in the cold upper portion of the crystal where the point defect profiles are no longer changing axially (Voronkov, 1982),

$$\Delta(r,z) = C_I(r,z) - C_V(r,z) \tag{5.2}$$

Then, the tendency for self-interstitial type microdefects to dominate would correspond to $\Delta > 0$, while vacancy microdefects would correspond to regions where $\Delta < 0$.

There exists significant evidence that such a model is qualitatively accurate. The region inside the OSF-ring has been shown to contain vacancy rich microdefects while the region outside the ring typically consists of interstitial based microdefects. The radial distribution of these microdefects is qualitatively similar to microdefect distributions observed in floating-zone (FZ) crystals. Despite the lack of an OSF-ring structure, it is generally accepted that so-called D-defects which originate from the axis of the crystal are vacancy related while microdefects that originate from the sidewall (A and B defects) are interstitial related. For growth conditions that lead to both of these defect types simultaneously, a defect-free zone is observed in be-

tween which separates the two regions called the *neutral zone*. Several experiments have shown that the neutral zone in FZ crystals and the OSF-ring in CZ crystals possess very similar responses to changes in crystal growth rate and thermal environment (Voronkov, 1982). The dramatically different oxygen impurity levels in the two crystals thus affects the *structure* of the boundary between vacancy and interstitial microdefect regions, but not the *position* of the boundary as a function of operating conditions. This simple picture suggests that microdefect growth is essentially controlled by the dynamics of intrinsic point defects in the crystal. Both the neutral zone in FZ crystals and the OSF-ring region, denoted by $r = R_{OSF}$, would therefore correspond to $\Delta \sim 0$ or $\Delta(R_{OSF}, z) = 0$.

5.2.3 The Model of Voronkov (1982)

Voronkov (1982) was the first to apply the above picture in a quantitative investigation of the response of A,B and D defect distributions to changes in the growth rate of FZ crystals. His analysis was based on a net flux of point defects entering the crystal by considering convection due to growth and diffusion in one dimension

$$J = \left(VC_I - D_I \frac{\partial C_I}{\partial z}\right) - \left(VC_V - D_V \frac{\partial C_V}{\partial z}\right), \tag{5.3}$$

where C_I and C_V are the interstitial and vacancy concentrations, respectively, D_I and D_V are the diffusivities, and z is the axial coordinate within the crystal. For low growth rates, the diffusive flux dominates, and integration of eq (5.3) gives

$$D_I C_I - D_V C_V = k_1. \tag{5.4}$$

For high growth rates, the convective flux is dominant so that

$$C_I - C_V = k_2, \tag{5.5}$$

where k_1 and k_2 are constants of integration In eq. (5.4), the point defect species with higher contribution to self-diffusion, DC, will dominate, while at high growth rates,

eq. (5.5), the species that is incorporated into the crystal at the melt/solid interface in higher concentration will dominate. This picture provides the essential ingredients for explaining the effect of growth rates on FZ microdefect distribution.

In order to predict trends that are in qualitative agreement with experimental observations, Voronkov required that $C_I^{eq} < C_V^{eq}$ and $D_I C_I^{eq} > D_V C_V^{eq}$ at very high temperatures. Assuming that point defects are incorporated at the melt/solid interface at their equilibrium concentrations, A and B defects will be favored at lower growth rates, while D defects will dominate at higher growth rates, in accordance with experimental observations.

Previous experimental work has shown that the growth conditions at which A defects disappear from an FZ crystal are related to a constant value for the ratio of V/G, where G is the axial temperature gradient near the melt/solid interface (Abe and Harada, 1983b). These observations are very similar in nature to the results of Dornberger et al. (1995) and give more confidence in the hypothesis that intrinsic point defects control the dynamics in both cases. The axial temperature gradient, G, was incorporated into the analysis of Voronkov by assuming a linear form for the temperature gradient near the melt/solid interface and assuming that point defect recombination is fast enough that mass-action is satisfied:

$$C_I C_V = C_I^{eq} C_V^{eq}. \tag{5.6}$$

According to the above arguments, the point at which neither A,B or D defects dominate corresponds to the net point defect flux being zero in eq. (5.3). Rewriting eq. (5.3) in terms of activation energies of formation, E_I and E_V, he obtains

$$\frac{V}{G} = \frac{E_I + E_V}{2kT_m^2} \frac{D_I C_I^{eq}(T_m) - D_V C_V^{eq}(T_m)}{C_V^{eq}(T_m) - C_I^{eq}(T_m)}, \tag{5.7}$$

where G is the axial temperature gradient near the melt/solid interface, which shows that the transition between defect type occurs at a constant value for the ratio of the pull rate and the axial temperature gradient and is determined by the high temperature point defect properties. By comparing the results of his analysis to experiments,

Voronkov was able to establish rough estimates for the point defect properties, and a reasonably good estimate for the value of V/G that corresponds to the transition between A,B and D defects.

Although qualitatively correct in some aspects, Voronkov's analysis relied on several unjustified assumptions. The most critical of these was the assumption that a one-dimensional model is capable of predicting the radial distribution of microdefects in FZ crystals. By not considering a two-dimensional thermal environment, the model required the super-position of a radial problem in which the cylindrical surface of the crystal was assumed to act as a sink for point defects; in this respect the model is only semi-quantitative in that it would not be able to predict the radial variations in microdefect distributions. Furthermore, the model required the assumption of constant diffusivities, equal enthalpies of formation for interstitials and vacancies and infinitely fast recombination kinetics. These assumptions will be discussed further in the context of the perturbation analysis presented in Section 5.6.

5.3 Point Defect Model for OSF-ring Dynamics

Point defect dynamics are modeled in a crystal grown under steady state conditions. The model includes bulk convection of point defects due to crystal motion during growth, Fickian diffusion and thermodiffusion, and solid-state reactions for self-interstitial and vacancy recombination and creation through Frenkel pair formation (Fahey *et al.*, 1989). The steady-state point defect balance equations are (Brown *et al.*, 1994):

$$\nabla \cdot \left(-D_I(T) \nabla C_I + \frac{D_I(T) C_I Q_I^{*\prime}}{kT^2} \nabla T \right) + V \frac{\partial C_I}{\partial z} + k_{IV}(T)(C_I C_V - C_I^{eq}(T) C_V^{eq}(T)) = 0, \tag{5.8}$$

and

$$\nabla \cdot \left(-D_V(T) \nabla C_V + \frac{D_V(T) C_V Q_V^{*\prime}}{kT^2} \nabla T \right) + V \frac{\partial C_V}{\partial z} + k_{IV}(T)(C_I C_V - C_I^{eq}(T) C_V^{eq}(T)) = 0, \tag{5.9}$$

for self-interstitials and vacancies, respectively. In eqs. (5.8) and (5.9), $\nabla = \mathbf{e}_r(\partial/\partial r) + \mathbf{e}_z(\partial/\partial z)$ for an axisymmetric system and the subscripts, I and V, denote properties of self-interstitials (I) and vacancies (V). Also, $Q_i^{*\prime}$ ($i = I, V$) are the reduced heats of transport for point defects (deGroot and Mazur, 1984), $D_i(T)$ ($i = I, V$) are their Fickian diffusivities, and $k_{IV}(T)$ is the rate coefficient for the recombination reaction, which is expressed in eqs. (5.8) and (5.9) in terms of the law of mass action.

5.3.1 Phenomenological Theory of Conjugate Forces and Fluxes

Equations (5.8) and (5.9) contain expressions for the thermodiffusion of point defects in a temperature gradient. Thermal gradients near the melt/solid interface can be quite large and some investigators have put forth the idea that many aspects of microdefect behavior can be interpreted in terms of thermodiffusive transport (Habu *et al.*, 1993b; Habu *et al.*, 1993a). The nature of thermodiffusion is less well understood than that of Fickian diffusion and this section will describe the origin of the thermodiffusion expression in the point defect balance equations given above.

Mathematically, expressions for Fickian and thermodiffusion are both derived from generalized phenomenological theory. The starting point for obtaining these expressions is the *dissipation function* which is proportional to the entropy production in a non-equilibrated system (deGroot and Mazur, 1984)

$$\Psi = T\sigma. \tag{5.10}$$

Ψ is the dissipation function and σ is the entropy production rate. Ψ can be expressed in terms of conjugate thermodynamic fluxes and forces:

$$\Psi = \sum_i J_i X_i, \tag{5.11}$$

where X_i are the forces (e.g. a concentration gradient) and J_i are the corresponding fluxes (e.g. Fickian diffusion). Each of the thermodynamic fluxes can be represented

as a sum of contributions arising from each of the forces, so that

$$J_i = \sum_k L_{ik} X_k. \tag{5.12}$$

Equations (5.12) are called the *phenomenological equations*, and the coefficients L_{ik} are the phenomenological coefficients. The term phenomenological is used because these linear relationships have been obtained purely empirically.

In order to develop explicit expressions for the L_{ik}, it is necessary to consider the entropy balance for our system. As with any property, a conservation equation can be written for the entropy,

$$\rho \frac{ds}{dt} = -\nabla \cdot \mathbf{J}_s + \sigma, \tag{5.13}$$

where $\sigma \geq 0$ and $\mathbf{J}_s$ is the entropy flux. The entropy is a function of the internal energy, density, and composition of the system and if "local" equilibrium can be assumed (even though "global" equilibrium is not satisfied) then, we can write

$$T \frac{ds}{dt} = \frac{du}{dt} + p \frac{dv}{dt} - \sum_{k=1}^{n} \mu_k \frac{c_k}{dt}. \tag{5.14}$$

Applying mass and energy balances and neglecting pressure work (a good assumption for solid state systems), the above expression gives

$$\rho \frac{ds}{dt} = \frac{-\nabla \cdot \mathbf{J}_q}{T} + \frac{1}{T} \sum_{k=1}^{n} \mu_k \nabla \cdot \mathbf{J}_k - \frac{1}{T} \sum_{j=1}^{r} J_j A_j. \tag{5.15}$$

where the last term represents the *reactive* contribution to the entropy change, for each reaction, j, and the A_j's are called the chemical affinities and are defined by

$$A_j = \sum_{k=1}^{n} \nu_{kj} \mu_k \qquad (j = 1, 2, \cdots, r). \tag{5.16}$$

The ν_{kj} are stoichiometric coefficients for each component k and reaction j. Equation (5.16) can be cast into the form of a balance equation such as eq. (5.13) by

differentiation by parts:

$$\rho\frac{ds}{dt} = -\nabla\cdot\left(\frac{\mathbf{J}_q - \sum_k \mu_k \mathbf{J}_k}{T}\right) - \frac{1}{T^2}\mathbf{J}_q\cdot\nabla T - \frac{1}{T}\sum_{k=1}^{n}\mathbf{J}_k\cdot\left(T\nabla\frac{\mu_k}{T}\right) - \frac{1}{T}\sum_{j=1}^{r}J_jA_j. \quad (5.17)$$

Comparing eq. (5.17) with eq. (5.13) the following statements can be made for the entropy flux and production:

$$\mathbf{J}_s = \frac{1}{T}\left(\frac{\mathbf{J}_q - \sum_k \mu_k \mathbf{J}_k}{T}\right), \quad (5.18)$$

and

$$\sigma = -\frac{1}{T^2}\mathbf{J}_q\cdot\nabla T - \frac{1}{T}\sum_{k=1}^{n}\mathbf{J}_k\cdot\left(T\nabla\frac{\mu_k}{T}\right) - \frac{1}{T}\sum_{j=1}^{r}J_jA_j, \quad (5.19)$$

and the expression for the dissipation function now becomes

$$\left[\Psi = -\frac{1}{T}\mathbf{J}_q\cdot\nabla T - \sum_{k=1}^{n}\mathbf{J}_k\cdot\left(T\nabla\frac{\mu_k}{T}\right) - \sum_{j=1}^{r}J_jA_j\right]. \quad (5.20)$$

Written this way, eq. (5.20) gives expressions for the thermodynamic "fluxes" and "forces" in the phenomenological equations. Note that eq. (5.20) is not the only way to distribute the fluxes, $\mathbf{J}_q$ and $\mathbf{J}_k$ among the temperature and chemical potential driving forces. In fact, there are three common ways in which the flux terms can be expressed. The first is as shown in eq. (5.20). The other two are developed by using the thermodynamic relationships (deGroot and Mazur, 1984)

$$Td\left(\frac{\mu_k}{T}\right) = (d\mu_k)_T - \frac{h_k}{T}dT, \quad (5.21)$$

$$T\nabla\left(\frac{\mu_k}{T}\right) = \nabla\mu_k - \left(\frac{h_k}{T}\right)\nabla T, \quad (5.22)$$

where h_k in eq. (5.21) is defined as dh/dn_k (i.e. the partial specific enthalpy of component k). If eq. (5.21) is used, the expression for the dissipation function becomes

$$\left[\Psi = -\frac{1}{T}\mathbf{J}_q'\cdot\nabla T - \sum_{k=1}^{n}\mathbf{J}_k\cdot[(\nabla\mu_k)_T]\sum_{j=1}^{r}J_jA_j\right], \quad (5.23)$$

while if eq. (5.22) is used we obtain

$$\left[\Psi = -\mathbf{J}_s \cdot \nabla T - \sum_{k=1}^{n} \mathbf{J}_k \cdot \nabla \mu_k - \sum_{j=1}^{r} J_j A_j\right]. \tag{5.24}$$

In eq. (5.23), $\mathbf{J}'_q$ is given by

$$\mathbf{J}'_q = \mathbf{J}_q - \sum_{k=1}^{n} h_k \mathbf{J}_k, \tag{5.25}$$

and in eq. (5.24),

$$\mathbf{J}_s = \frac{1}{T}\mathbf{J}'_q + \sum_{k=1}^{n} s_k \mathbf{J}_k. \tag{5.26}$$

where $s_k \equiv -(\mu_k - h_k)/T$ and is the partial specific entropy of component k.

Eqs. (5.20), (5.23), and (5.24) give expressions for the entropy production (or dissipation function) which are equivalent, differing only in the way the flux terms are distributed among the forces. The choice of representation depends on what phenomena are present in the system; certain choices can make particular cases simpler to treat.

Mathematical Expressions for Thermodiffusion

This section describes the application of the phenomenological equations to obtain an expression for thermodiffusion. Consider first a system in which only vacancies diffuse in a lattice under a temperature and concentration gradient in the absence of reaction (Schaal, 1983), so that

$$\sigma = J_V X_V + J_A X_A + J_q X_q, \tag{5.27}$$

where A represents a lattice atom, V is a vacancy and q represents heat. Rewriting,

$$\sigma = J_V(X_V - X_A) + J_q X_q + (J_V + J_A)X_A. \tag{5.28}$$

For vacancy diffusion, a flow of vacancies necessarily implies a reverse flow of lattice atoms $J_A = -J_V$ so,

$$\sigma = J_V(X_V - X_A) + J_q X_q. \tag{5.29}$$

Here, the convention of eq. (5.20) is adopted; i.e. the specific partial enthalpy of the diffusing species is contained in the heat flux, J_q. The phenomenological equations, eqs. (5.12), can be expressed as

$$J_V = L_{VV}(X_V - X_A) + L_{Vq}X_q, \tag{5.30}$$

$$J_A = -J_V, \tag{5.31}$$

$$J_q = L_{qV}(X_V - X_A) + L_{qq}X_q. \tag{5.32}$$

A "heat of transport", $Q^* \equiv L_{Vq}/L_{VV}$ can now be defined and the above phenomenological equations can be written as

$$J_V = L_{VV}(X_V - X_A + Q_V^* X_q), \tag{5.33}$$

$$J_A = -J_V, \tag{5.34}$$

$$J_q = Q_V^* J_V + (L_{qq} - L_{Vq}Q_V^*)X_q. \tag{5.35}$$

Rewriting the last equation,

$$J_q - L_{qq}X_q = Q_V^*(J_V - L_{Vq}X_q), \tag{5.36}$$

and so the physical meaning of Q_V^* is given by the ratio of the heat flux to the mass flux,

$$Q_V^* = \frac{J_q}{J_V}, \tag{5.37}$$

when $\nabla T = 0$. The thermodiffusion terms in eqs. (5.8) and (5.9) are expressed in terms of the *reduced* heats of transport, $Q^{*'}$. The relationship between Q^* and $Q^{*'}$ is obtained by simply adopting the convention of equation (5.23) instead of equa-

tion (5.20). Let

$$X'_V \equiv X_V + h_V X_q, \tag{5.38}$$

$$X'_A \equiv X_A + h_A X_q, \tag{5.39}$$

$$X'_q \equiv X_q. \tag{5.40}$$

where h_V and h_A are the partial specific enthalpies of A and V. In order that σ be invariant under this transformation, the following fluxes are defined

$$J'_V \equiv J_V, \tag{5.41}$$

$$J'_A \equiv J_A, \tag{5.42}$$

$$J'_q \equiv J_q - (h_V - h_A)J_V. \tag{5.43}$$

The phenomenological equations can now be written as

$$J_V = L_{VV}(X'_V - X'_A) + L'_{Vq}X_q, \tag{5.44}$$

$$J'_q = L'_{Vq}(X'_V - X'_A) + L'_{qq}X_q, \tag{5.45}$$

and in the same way as before, we can define a (reduced) heat of transport by considering the $\nabla T = 0$ case so that

$$L'_{Vq} = Q^{*'}_V L_{VV}, \tag{5.46}$$

and the relationship between the two heats of transport is seen to be

$$Q^{*'}_V = Q^*_V - (h_V - h_A). \tag{5.47}$$

In the convention of eq. (5.23), the concentration gradient driving force is given by $(\nabla\mu)_T$. Application of the Gibbs-Duhem equation,

$$n_A(\nabla\mu_A)_T + n_V(\nabla\mu_V)_T = 0, \tag{5.48}$$

and the equation for chemical potential,

$$\mu_V = \mu_V^0 + kT \log \frac{n_V}{n_V + n_A}, \tag{5.49}$$

we obtain the following expressions for the fluxes:

$$J_V = -\frac{kL_{VV}}{n_V}\left(\nabla n_V + \frac{n_V Q_V^{*\prime}}{kT^2}\nabla T\right), \tag{5.50}$$

$$J_q' = -\frac{kL_{VV}}{n_V}Q_V^{*\prime}\nabla n_V - \frac{L_{qq}'}{T^2}\nabla T. \tag{5.51}$$

In other words, the reduced heat of transport is

$$Q_V^{*\prime} = \frac{J_q'}{J_V}, \tag{5.52}$$

for $\nabla T = 0$. Rewriting the vacancy concentration in terms of mole fraction, $C_V = n_V/(n_A + n_V)$, and the flux as $j_V = J_V/(n_A + n_V)$ gives

$$j_V = -D_V\left(\nabla C_V + \frac{C_V Q_V^{*\prime}}{kT^2}\nabla T\right). \tag{5.53}$$

This is the expression for the transport flux in eq. (5.9).

Now, a similar expression can be developed for self-interstitial transport. Analogously to equation (5.53) we can write

$$j_I = -D_I\left(\nabla C_I + \frac{C_I Q_I^{*\prime}}{kT^2}\nabla T\right), \tag{5.54}$$

and the phenomenological equations as

$$J_I = L_{II}X_I + L_{Iq}'X_q, \tag{5.55}$$

$$J_q' = L_{Iq}'X_I' + L_{qq}'X_q, \tag{5.56}$$

where

$$X_I' = X_I + h_I X_q. \tag{5.57}$$

Note that in the case of self-interstitial transport, the species flux is not related to the flux of lattice atoms as it is in the case of the vacancy so the expression relating the reduced heat of transport to the heat of transport is given by

$$Q_I^{*\prime} = Q_I^* - h_I. \tag{5.58}$$

The relationship between the reduced heats of transport and the original quantities can be viewed from an intuitive point of view. Associated with the migration of a species is the implicit transfer of the energy needed to form it. In other words, when a self-interstitial (or vacancy) moves from one part of the crystal to another, it automatically transports its energy of formation with it. The reduced quantities subtract this effect out.

Estimation of the (Reduced) Heats of Transport

Estimates for the reduced heats of transport ($Q_I^{*\prime}$ and $Q_V^{*\prime}$) for point defects in silicon are not available in the literature. These quantities are extremely difficult to measure experimentally and can be estimated theoretically only with significant assumptions. Some attempts at computing $\{Q_i^{*\prime}\}$ in solids involve atomistic simulations and employing the Irving-Kirkwood equation (Gillan, 1977; J. and Finnis, 1978; Jones *et al.*, 1996) using detailed knowledge of the paths of point defect diffusion in the lattice. Jones et al. (1996), for example, have used this approach to consider the thermodiffusion of vacancies in solid argon. They concluded that, for this particular system, the reduced heat of transport was about -0.2 eV, computed from an estimate for the non-reduced quantity of about -0.11 eV and the formation energy for a vacancy in solid argon which has been estimated previously to be about 0.09 eV (Burton and Jura, 1967). This example gives an estimate for both the reduced and non-reduced heats of transport that is on the order of the energy of formation. However, it has been pointed out that there is no way to determine *a priori* for a given system how these quantities relate to the formation or migration energies.

The Wirtz model (Wirtz, 1943) provides a qualitative framework for describing

thermodiffusion in terms of the energy of migration. Here, the diffusion path is decomposed into three distinct phases; jump initiation, the saddle point, and jump conclusion. The heat of transport is computed by estimating which of the above steps requires the most energy for diffusion to occur. The Wirtz model has not been found to give quantitatively reliable estimates in most cases and there is controversy over whether its interpretation of the thermodiffusion process is correct.

In conclusion, estimates for heats of transport derived from theory and experiment follow no particular trend relative to either the migration energy or the formation energy. They can be positive or negative, even for the same compound investigated by different researchers (Bokshtein *et al.*, 1985) It will be assumed that thermodiffusion plays a small role in shaping OSF-ring dynamics and that the *reduced* heats of transport can be neglected. This assumption implies that the non-reduced quantity is equal to the enthalpy of formation for both point defect species. The sensitivity analysis in Section 5.5.2 demonstrates the validity of this assumption. Furthermore, it will be shown asymptotically that, given an upper bound on the magnitude of the reduced heats of transport, the effects of thermodiffusion are of secondary importance relative to other transport mechanisms.

5.3.2 Boundary Conditions

Point defect concentrations in the crystal are predicted from eqs. (5.8)-(5.9) with a specific set of boundary conditions on the crystal surfaces. We assume that the self-interstitial and vacancy concentrations are in equilibrium at the melt/solid interface; for a flat interface this gives

$$C_I(r,0) = C_I^{eq}(T_m), \tag{5.59}$$

$$C_V(r,0) = C_V^{eq}(T_m). \tag{5.60}$$

The axis of the crystal ($r = 0$) is taken as an axis of symmetry, so that

$$\mathbf{e}_r \cdot \mathbf{J}_i = 0, \quad i = I, V \tag{5.61}$$

where the flux of the point defect i is written as

$$\mathbf{J}_i \equiv -D_i(T)\nabla C_i + \frac{D_i C_i Q_i^{*\prime}}{kT^2}\nabla T. \tag{5.62}$$

Along the exposed crystal surface we also assume that the flux of point defects is zero, i.e.

$$\mathbf{n} \cdot \mathbf{J}_i|_{\partial D_{cry}} = 0, \quad i = I, V, \tag{5.63}$$

where $\mathbf{n}$ is the outward pointing normal on the cylindrical crystal surface, and $\mathbf{n} = \mathbf{e}_z$ at the top of the crystal. Applying eq. (5.63) removes radial outdiffusion of point defects as a mechanism for changing the point defect concentrations and simplifies the interpretation of the results described here. Other boundary conditions are possible. For example, Brown et al. (1994) and Habu and Tomiura (1996) assume that point defects along the surface of the crystal are in equilibrium at the local temperature; i.e.,

$$C_i|_{\partial D_{cry}} = C_i^{eq}[T(r,z)]|_{\partial D_{cry}}, \quad i = I, V. \tag{5.64}$$

Hu (1985) has proposed a Robin condition that lies between the no-flux condition, eq. (5.63), and the infinitely fast source/sink model, eq. (5.64):

$$\mathbf{n} \cdot \mathbf{J}_i|_{\partial D_{cry}} + \gamma[C_i - C_i^{eq}(T)]|_{\partial D_{cry}} = 0, \quad i = I, V, \tag{5.65}$$

where γ is an unknown constant that represents a surface recombination velocity and depends on the surface environment of the crystal. We pursue the sensitivity of our results to the constant γ in Section 5.5.1.

5.3.3 Other Models of Point Defect Dynamics

The Model of Habu et al. (1993),(1994),(1996)

The use of point defect dynamics to model OSF-ring behavior has been investigated previously by other researchers. Habu et al. (1993) have developed a similar approach of using experimental data related to the OSF-ring to "tune" point defect properties

in a model of point defect dynamics. Despite the similarity in the approach there are fundamental differences between the present model and the model of Habu et al. (1993) (from here on referred to as the Habu model) which deserve discussion. The details of the Habu model are important because its results, namely estimates for the point defect thermophysical properties, are directly comparable to our results.

The point defect dynamic model of Habu et al. (1993) considers convection, Fickian and thermodiffusion, and reaction of point defects in a growing crystal with the following balance equations for self-interstitials and vacancies:

$$\frac{dN_I}{dt} = \nabla \cdot \left[D_I \nabla N_I - \left(v + \frac{D_I (H_I^f - Q_I^*)}{RT^2} \nabla T \right) N_I \right] \tag{5.66}$$

$$- \frac{4\pi (D_I + D_V) R_c N_I N_V}{E_F} \exp\left(-\frac{\Delta G}{RT} \right) RT \ln \frac{N_I N_V}{N_I^{eq} N_V^{eq}}, \tag{5.67}$$

$$\frac{dN_V}{dt} = \nabla \cdot \left[D_V \nabla N_V - \left(v + \frac{D_V (H_V^f - Q_V^*)}{RT^2} \nabla T \right) N_V \right] \tag{5.68}$$

$$- \frac{4\pi (D_I + D_V) R_c N_I N_V}{E_F} \exp\left(-\frac{\Delta G}{RT} \right) RT \ln \frac{N_I N_V}{N_I^{eq} N_V^{eq}}, \tag{5.69}$$

where $N_{I,V}$ are the interstitial and vacancy concentrations, ΔG is the recombination barrier, E_F is the heat released by the recombination reaction, R_c is the recombination distance.

There are several notable differences between this representation and the point defect model of eqs. (5.8) and (5.9). The first and most significant of these is the extra factor, $H_{I,V}^f$, in the thermodiffusion term present in the Habu model. The origin of this term is best understood by following the derivation in Habu et al. (1993b).

The discussion of the generalized phenomenological theory in Section 5.3.1 described how entropy production can be expressed equivalently in three representations, the only difference being in the distribution of the fluxes among the forces.

Habu's formulation of thermodiffusion begins with an expression for the chemical potential driving force for a species, D, using the convention of eq. (5.24) (Habu *et al.*,

1993b),

$$J_D = -M\nabla\mu_D, \tag{5.70}$$

where it is assumed that $\nabla T = 0$. The chemical potential of D, μ_D, is given by

$$\mu_D = RT \ln \frac{C_D}{C_{eq}}, \tag{5.71}$$

and

$$C_{eq} = C_D^0 \exp\left(-\frac{H_D^f}{RT}\right). \tag{5.72}$$

The derivative of the chemical potential with respect to temperature and concentration is then given by

$$J_D = -\frac{MRT}{C_D}\nabla C_D + MRT\frac{H_D^f}{RT^2}\nabla T. \tag{5.73}$$

The second term in eq. (5.73) defines the so-called "uphill diffusion" (diffusion against a concentration gradient) on which Habu bases his analysis of OSF-ring dynamics. For $\nabla T = 0$, eq. (5.73) must lead to an expression of Ficks Law, so

$$M = \frac{C_D D_D}{RT}, \tag{5.74}$$

and

$$J_D = -D_D\nabla C_D + \frac{D_D C_D H_D^f}{RT^2}\nabla T. \tag{5.75}$$

Again, the second term represents the so-called "uphill diffusion" effect. The problem with the Habu formulation appears in a second publication which is denoted here as Paper II (Habu *et al.*, 1993a) where the general form of the phenomenological equations is written as

$$J_D = \alpha_{DD}[F_D - (\nabla\mu_D)_T] - \frac{\alpha_{DQ}}{T}\nabla T, \tag{5.76}$$

$$J_Q = \alpha_{QD}[F_D - (\nabla\mu_D)_T] - \frac{\alpha_{QQ}}{T}\nabla T, \tag{5.77}$$

where F_D represents an external force which, in this case, is zero. These equations are written in the convention of eq. (5.23) not (5.24). Habu uses the result of paper I (Habu *et al.*, 1993b) for $(\nabla\mu_D)_T$ term but Paper I evaluated $\nabla\mu_D$ **not** $(\nabla\mu_D)_T$. So, eq. (16) in paper II (Habu *et al.*, 1993a) written as,

$$J_D = -D_D\nabla C_D + \frac{D_D C_D(H_D^f + \delta Q_D)}{RT^2}\nabla T \tag{5.78}$$

is inconsistent because it is derived with two different representations of the phenomenological equations. The correct expression for $(\nabla\mu_D)_T$ should be evaluated at constant temperature and would give

$$J_D = -D_D\nabla C_D + \frac{D_D C_D \delta Q_D}{RT^2}\nabla T, \tag{5.79}$$

which is equivalent to the expression in eqs. (5.8) and (5.9). If $(\nabla\mu_D)$ were to be used, i.e. the convention in Paper I, then the correct conjugate flux to use with the temperature gradient is the entropy flux, $\mathbf{J}_s$, where

$$\mathbf{J}_s = \frac{1}{T}\mathbf{J}_q' + \sum_{k=1}^{n} s_k \mathbf{J}_k, \tag{5.80}$$

and $s_k = -(\mu_k - h_k)/T$ is the partial specific entropy of component k.

The incorrect expression for thermodiffusion in the formulation of the Habu model gives a very large estimate for thermodiffusion, which, under certain growth conditions, can lead to a net flux of point defects *towards* the melt/solid interface. Thus, at low crystal growth rates, thermodiffusion determines which species survives in excess in the crystal. By adjusting the thermophysical constants in the model, Habu et al. are able to reproduce qualitatively the radial position of a self-interstitial peak which is postulated to correspond to the OSF-ring. The presence of this peak is due to the radial outdiffusion of point defects in the region near the cylindrical surface of the crystal. The correlation of an interstitial peak to the OSF-ring is at variance with experimental observations that D-defects exist in the region immediately inside the OSF-ring, which implies that the OSF-ring exists at the boundary between interstitial

and vacancy-dominated regions.

The Habu model also does not consider the effects of a thermal field on these predictions. Thus, it is impossible to determine whether the model reproduces correctly the empirical V/G condition of eq. (5.1). Furthermore, the model parameters are tuned so that qualitative agreement for pull rate variations is obtained without accounting for the fact that the thermal gradient varies radially by over 100%. This variation is bound to change the dynamic response of the model significantly. One of the key assumptions in the Habu model is that the pull rate at which the OSF-ring disappears into the center of the crystal is a fixed quantity. In fact, this value is completely dependent on the thermal profile within the crystal making the above assumption incorrect.

In conclusion, the mechanism for OSF-ring dynamics in the Habu model is very different to the one presented in this chapter, and it is primarily based on an incorrect formulation of the thermodiffusion effect. Comparisons between the resulting fitted thermophysical properties from the two models will be discussed in detail in Section 5.4.2. It is clear that the Habu model is incapable of reproducing the empirical V/G relationship in its current form because it does not consider the radial thermal profile. The results of this chapter will provide very strong evidence that the radial profile of the thermal field is sufficient to explain OSF-ring dynamics for a wide range of crystal pull rates and geometries.

5.3.4 Finite Element Analysis

Profiles of the point defect concentrations, $(C_I(r,z), C_V(r,z))$ and of $\Delta(r,z)$ are computed by solving eqs. (5.8) and (5.9) by the Galerkin finite element method with the temperature field $T(r,z)$ as input. The crystal is partitioned into quadrilateral finite elements as shown in Figure 5-4. The concentration fields are approximated as summations of unknown coefficients and bilinear finite element basis functions. The field equations and boundary conditions are discretized by Galerkin's method to yield a large set of nonlinear algebraic equations for the unknown coefficients. These equations are solved by Newton's method using LU-decomposition with frontal storage

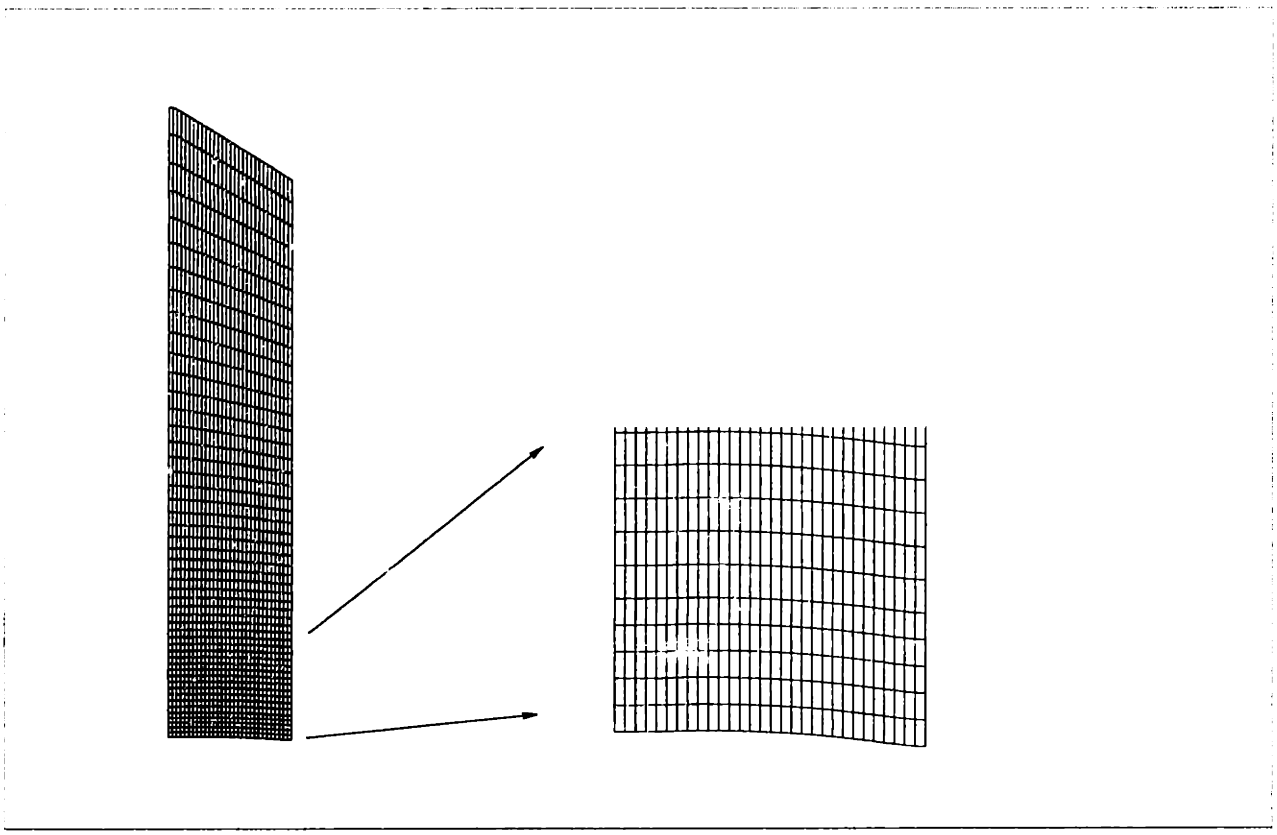

Figure 5-4: Finite element mesh used to compute C_I and C_V. The melt/solid boundary shape (expanded view) is determined from large-scale heat transfer simulations of the actual crystal growth system.

methods (Hood, 1976).

The finite element mesh is adapted to yield accurate solutions in regions where the defect concentrations change rapidly. As shown below, such rapid variation occurs adjacent to the melt/crystal interface. A typical mesh used in our analysis has 50 elements arranged axially and 30 elements distributed radially.

Temperature fields used in defect dynamics simulations are taken directly from finite element analysis of heat transfer in the crystal and throughout the Czochralski system for a particular geometry and growth rate (Bornside *et al.*, 1990; Dupret *et al.*, 1990). These simulations account for high temperature heat transfer, solidification and capillarity in the Czochralski crystal growth system to predict the thermal profiles and crystal geometry in the system. The heat transfer model accounts for conduction and diffuse-grey radiation which couples the heat transfer between surfaces and the crystal radius through view factors. Capillary action at the gas/melt interface is described with the Young-Laplace equation (Bornside *et al.*, 1990) which balances capillary and hydrostatic forces at the melt meniscus. The entire problem is reduced to a set of non-linear algebraic equations which are solved by Newton's method.

The crystal temperature fields are interpolated onto the bilinear finite element basis used for the defect concentrations and used as input in the point defect simulations. The heat transfer calculations also are used to compute the shape of the melt/crystal interface for a particular growth rate and geometry, which is used to define the shape of the mesh in the point defect simulations.

5.3.5 Point Defect Property Estimation

The predictions of the point defect dynamic model are critically sensitive on the values of the thermophysical properties – equilibrium concentrations, heats of transport, diffusivities and the rate coefficient for recombination – of the defects. Simultaneously, as discussed in Chapter 3, there is considerable disagreement on the values of these coefficients and their dependence on temperature.

The present analysis begins with the values of these parameters predicted using the atomistic simulations described in Chapter 3 based on the Stillinger-Weber inter-

atomic potential (Stillinger and Weber, 1985). The estimates for the diffusivities and equilibrium concentrations are given below in Arrhenius form as:

$$D_i(T) = D_i^0 \exp\left(-\frac{H_i^m}{kT}\right), \quad i = I, V \tag{5.81}$$

$$C_i^{eq}(T) = \rho \exp\left(-\frac{S_i^f(T)}{k}\right) \exp\left(-\frac{H_i^f(T)}{kT}\right), \quad i = I, V \tag{5.82}$$

where ρ is the density of lattice sites, $\{H_i^m\}$ are the migration energies, $\{H_i^f\}$ the formation enthalpies, and $\{S_i^f\}$ the formation entropies, respectively. As described previously, the enthalpies and entropies of formation were fit to linear relationships, $H_i^f(T) = A_i + B_i T$ and $S_i^f(T) = X_i + Y_i T$, respectively. The thermophysical point defect properties predicted from the Stillinger-Weber interatomic potential are summarized below:

$$D_I(T) = 1.76 \times 10^{-2} \exp\left(-\frac{0.94}{kT}\right) \ cm^2/s \tag{5.83}$$

$$D_V(T) = 1.70 \times 10^{-3} \exp\left(-\frac{0.46}{kT}\right) \ cm^2/s \tag{5.84}$$

$$C_I^{eq}(T) = 2.97 \times 10^{23} \exp\left(\frac{1.53 + 3.85 \times 10^{-3} T}{k}\right) \exp\left(-\frac{3.46 + 3.08 \times 10^{-4} T}{kT}\right) \ cm^{-3} \tag{5.85}$$

$$C_V^{eq}(T) = 4.97 \times 10^{22} \exp\left(\frac{0.76 + 3.53 \times 10^{-3} T}{k}\right) \exp\left(-\frac{2.48 + 2.33 \times 10^{-4} T}{kT}\right) \ cm^{-3} \tag{5.86}$$

$$k_{IV}(T) = 1.2 \times 10^{-6} \exp\left(-\frac{0.58 - T(2.29 + 7.38 \times 10^{-3} T)}{kT}\right) \tag{5.87}$$

5.3.6 Fitting Point Defect Properties to Experimental Data

The thermophysical properties for point defects predicted from the Stillinger-Weber interatomic potential are a starting point for simulation of point defect dynamics. However, the wide variation in experimental estimates for these parameters and the empirical nature of the Stillinger-Weber potential make it doubtful that quantitative predictions can be expected in the context of the present simulations. To produce quantitatively accurate simulations for the dynamics of the OSF-ring the preexponential factors (D_I^0, D_V^0, X_I, X_V) are taken as adjustable parameters to fit experiments.

The remainder of the functional forms for the thermophysical properties are kept at the values reported in eqs. (5.83)-(5.87).

The four parameters (D_I^0, D_V^0, X_I, X_V) were fit to a set of measurements, $R_{OSF}^{exp}(V)$, for a particular crystal radius and growth system; we used the system denoted as HS1/4" by Dornberger et al. (1996). for growth of a 4" diameter crystal. The location of the OSF-ring was computed by finite element analysis from the condition $\Delta(R_{OSF}, L) = 0$ where L is far from the melt/crystal interface. The goodness of fit between the defect dynamic analysis and the experimental measurements was quantified by the objective function:

$$\chi \equiv \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left[R_{OSF}(V_i) - R_{OSF}^{exp}(V_i)\right]^2} \tag{5.88}$$

where $\{V_i\}$ are a set of growth rates within a range for which the OSF-ring is observed experimentally.

The sensitivity of the predictions for $R_{OSF}(V)$ to the parameters (D_I^0, D_V^0, X_I, X_V) was probed before an automated optimization scheme was developed. Interestingly, the response of $R_{OSF}(V)$ was equal and opposite to changes in X_I and X_V; i.e. increasing the interstitial equilibrium concentration and decreasing the vacancy equilibrium concentration have the same effects. This observation is used to reduce the number of adjustable parameters to three; $X_I - X_V$, D_I^0, and D_V^0. Also, the contours of $\Delta(r, z)$ are extremely sensitive to the values of these parameters; variations frequently led to the contour disappearing into the center of the crystal or at the crystal edge, where R_{OSF} is no longer physically defined. This constraint made the application of a multi-dimensional optimization scheme very difficult. Instead, the parameter fitting was performed as a series of one-dimensional Golden-Section searches (Press *et al.*, 1989) to determine optimum values of D_I^0 for preselected values of $X_I - X_V$ and D_V^0, which were selected to give physically reasonable predictions.

Combining the constants (X_I, X_V) into the single parameter $X_I - X_V$ forces one of the point defect concentrations to be set independently of the present simulations. X_V (and therefore C_V^{eq}) was selected from literature estimates of the vacancy equilibrium

concentration and so that the interstitial and vacancy self-diffusion contributions, $D_I C_I^{eq}$ and $D_V C_V^{eq}$ agree well with literature estimates which are known to within an order-of-magnitude at high temperature (Frank *et al.*, 1984).

5.4 Simulation Results

Experimental results for two crystal growth systems and three crystal diameters are compared to the results of the point defect simulations; the two systems are denoted as HS1 and HS2 and the three experimental data sets (HS1/4", HS1/6", and HS2/8"), where the second number denotes the crystal diameter.

5.4.1 Parameter Fitting and Comparison with Experiments

The case HS1/4" was used for optimizing the parameters ($X_I - X_V$, D_I^0, D_V^0). The comparison between experimental results $R_{OSF}^{exp}(V)$ and the predictions of simulations are shown in Figure 5-5 for the three cases using the best estimates for the point defect properties. These values of the parameters are:

$$D_I = 0.242 \exp\left(-\frac{0.937}{kT}\right) \quad cm^2/s \tag{5.89}$$

$$D_V = 1.0 \times 10^{-3} \exp\left(-\frac{0.457}{kT}\right) \quad cm^2/s \tag{5.90}$$

$$C_I^{eq} = 2.97 \times 10^{23} \exp\left(\frac{1.40 + 3.85 \times 10^{-3}T}{k}\right) \exp\left(-\frac{3.46 + 3.08 \times 10^{-4}T}{kT}\right) \quad cm^{-3} \tag{5.91}$$

$$C_V^{eq} = 4.97 \times 10^{22} \exp\left(\frac{-3.70 + 3.53 \times 10^{-3}T}{k}\right) \exp\left(-\frac{2.48 + 2.33 \times 10^{-4}T}{kT}\right) \quad cm^{-3} \tag{5.92}$$

$$k_{IV}(T) = 1.2 \times 10^{-6} \exp\left(-\frac{0.61 - T(-2.30 + 7.38 \times 10^{-3}T)}{kT}\right) \tag{5.93}$$

The goodness of fit of the experimental results is similar for both the data set used in the parametric search (HS1/4") and the other two data sets. The sensitivity of the data fit to the choice of parameters is demonstrated in Figure 5-6 by plotting contours of χ from eq. (5.88) as a function of $X_I - X_V$ and D_V^0 for the data. There is a fairly

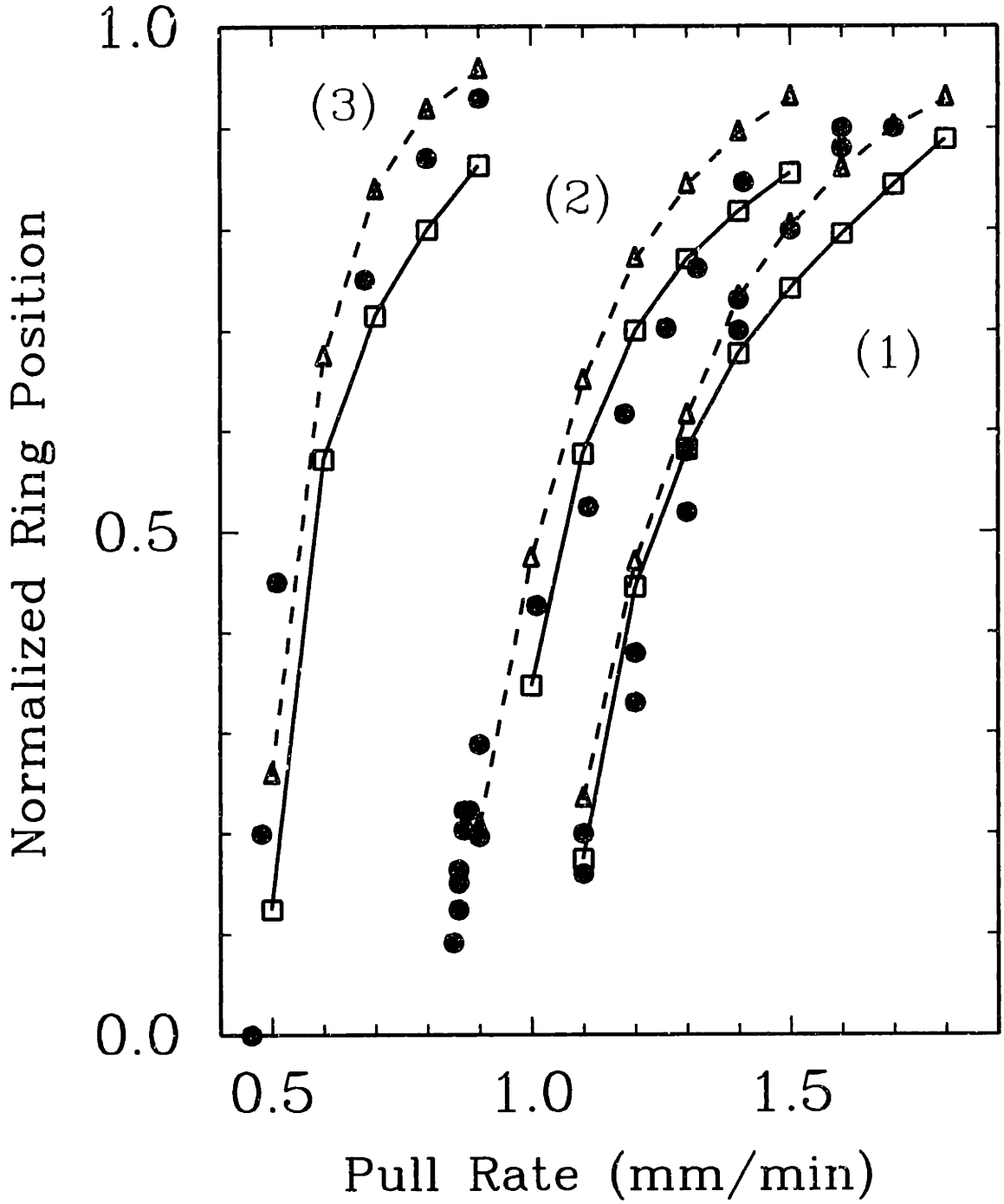

Figure 5-5: OSF-Ring response to changes in crystal growth rate (Dornberger and von Ammon, 1996): (1) HS1/4", (2) HS1/6", (3) HS2/8"; □ = Simulation Results; • = Experimental Results; △ = Correlation for $(V/G(R_{OSF}))_{crit} \approx 1.34 \times 10^{-3}\ cm^2/minK$

.

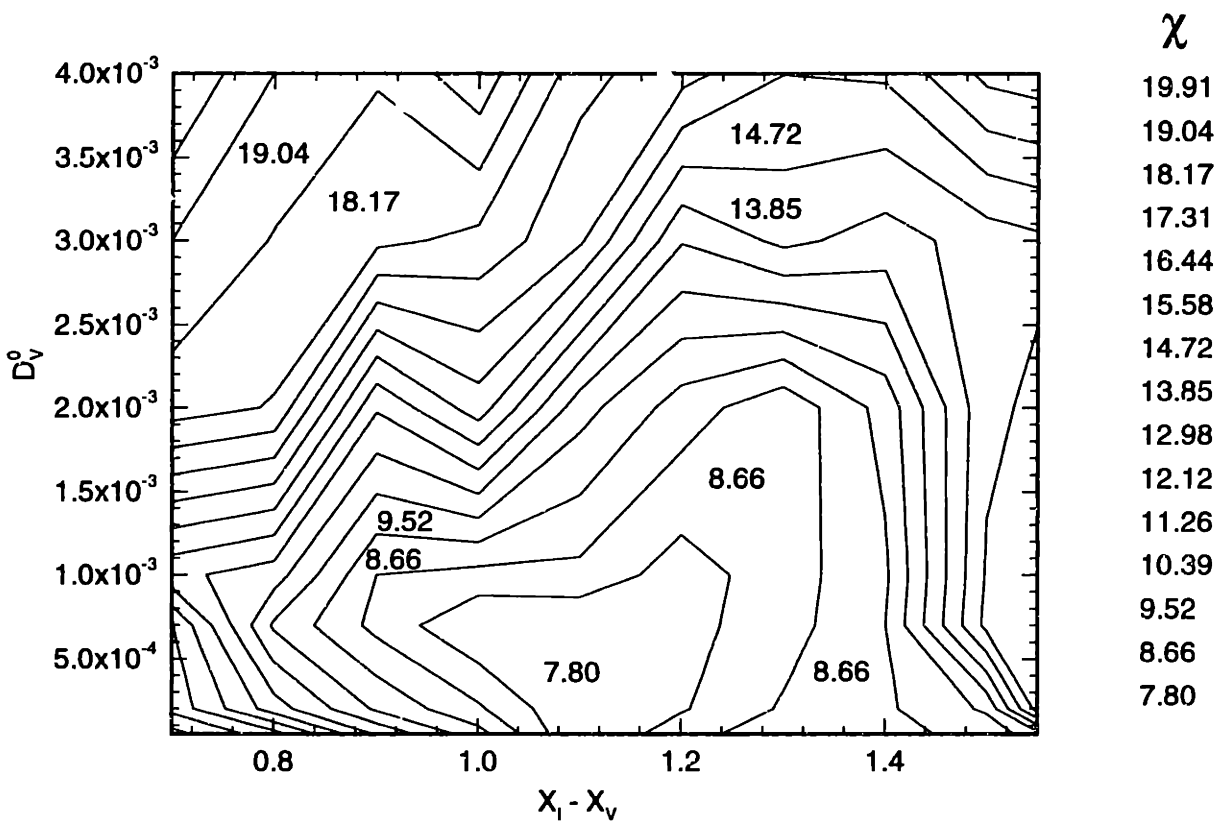

Figure 5-6: Contours of χ as a function of D_V^0 and $(X_I - X_V)$

broad region of parameters that yield similar quality in the parametric fitting. This region is roughly bound by $4.7 \leq (X_I - X_V) \leq 5.2$ and $2\times10^{-4} \leq D_V^0 \leq 2\times10^{-3}$; these bounds correspond to $0.1 \leq D_I^0 \leq 0.7$. Any parameter set within these bounds leads to good agreement between the point defect simulations and the OSF-ring data in terms of the displacement between the computed and experimental values of $R_{OSF}(V)$.

The non-uniqueness of the point defect parameters that lead to good agreement with OSF-ring data stems from two sources. The experimental results of Dornberger et al. (1996) are uncertain to within 5% of the radius, the uncertainty becoming larger as the OSF-ring is found nearer to the center of the crystal because of the more sensitive response. The optimization scheme described above is very sensitive to such scatter and the result is broadening of the parameter subspace that satisfies the experimental data. The second source of uncertainty arises from the definition of the objective function, χ. As described in eq. (5.88), it is sensitive to displacement between the two response curves but not to their *shapes*. Figure 5-7 shows the predicted OSF-ring dynamics for two distinct sets of point defect thermophysical properties, which although exhibit similar values of χ, lead to qualitatively different behavior.

It was found in an initial study that including a gradient term in the objective function gave no additional benefits since the slope was not sensitive to the small changes in D_I^0 during each of the one-dimensional optimizations. In other words, the slope of the response curve, $\Delta = 0$, is much less sensitive to changes in the parameters than the position. This implies that any optimization scheme will be strongly biased to parameters that minimize the displacement rather than a combination of both. However, it is possible to use the gradient information of the response curves to narrow the choice within the optimum subspace defined above. The slopes for the curves $R_{OSF}^{exp}(V)$ and $R_{OSF}(V)$ were computed via the polynomial fits that were used to compute χ in eq. (5.88) and we define a second objective function in terms of the gradients as

$$\chi_G \equiv \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left[\frac{dR_{OSF}(V_i)}{dV} - \frac{dR_{OSF}^{exp}(V_i)}{dV}\right]^2} \tag{5.94}$$

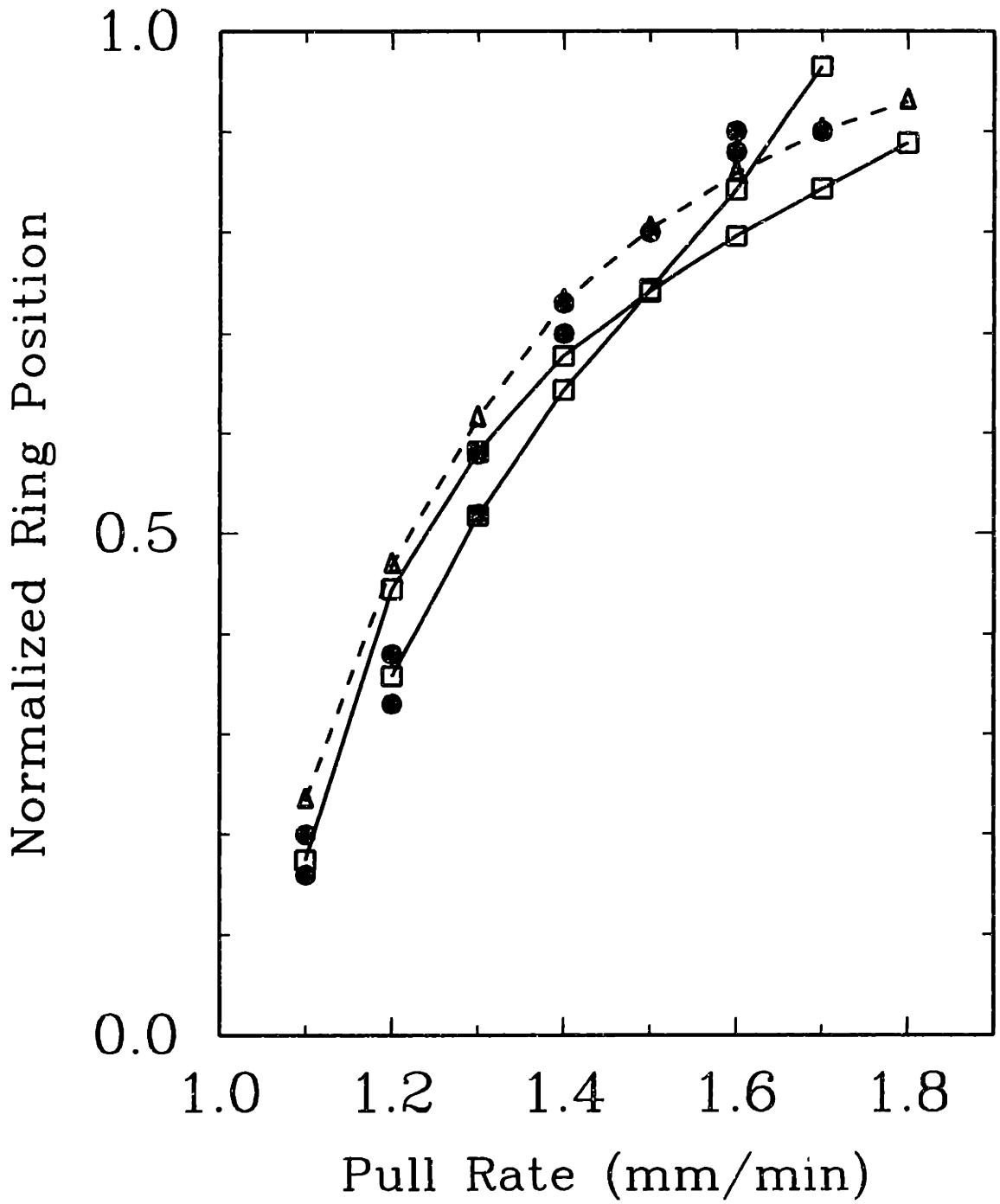

Figure 5-7: Similar values of χ for two different point defect property sets give very different results; • = experimental data, △ = V/G correlation, □ = simulation.

The point defect thermophysical property sets in the region of good fit were used to compute and compare slopes of the two curves using χ_G. The set that led to the lowest value of χ_G was then taken as the best estimate and is given above in eqs. (5.89)-(5.93). Values for χ_G ranged from 0.15 to about 0.24 in the optimum parameter subspace, providing good separation between the parameter sets in this region. The two step optimization process described here assumes that we can decouple the two objective functions, χ and χ_G, and perform optimizations with respect to each sequentially. Unfortunately, there is no rigorous method to estimate the uncertainty introduced by this approach.

The temperature field, interstitial and vacancy concentration fields, and contours of $\Delta(r,z)$ are shown in Figures 5-8 and 5-9 for calculations with system HS1 and a 6" diameter crystal grown at growth rates of (a) 1.0 mm/min and (b) 1.4 mm/min. General characteristics of these results are typical for point defect dynamics in silicon. First, C_I and C_V decrease rapidly from their equilibrium values with distance into the crystal. As described in Section 5.6, the recombination reaction accounts for this decrease. At axial distances greater than one radius the temperature has fallen to less than 0.75 T_m and the point defect distributions are effectively frozen, as represented by the parallel vertical structure of the contours for C_I, C_V, and Δ. The contour $\Delta = 0$ appears for both growth rates and moves to larger r with increasing V, as expected.

Because the boundary condition, eq. (5.63), has been used on the crystal sidewall the only driving force for *radial* structure in the point defect concentrations is the radial structure of the temperature field. Hence, a starting point for understanding the defect concentration profiles is to examine the solutions near $r = 0$, the crystal axis, and to ignore radial variations. At $r = 0$ and close to the interface there are more vacancies than interstitials; compare $C_V^{eq}(T_m) = 1.2 \times 10^{15}\ cm^{-3}$ to $C_I^{eq}(T_m) = 9.6 \times 10^{14}\ cm^{-3}$. Eventhough the recombination rate coefficient is high, the reaction rate is initially zero because the concentrations are in equilibrium. The falling temperature and decreasing equilibrium concentrations, and the differing values of the diffusivities ($D_I(T_m) = 3.8 \times 10^{-4}\ cm^2/s$ and $D_V(T_m) = 4.3 \times 10^{-5}\ cm^2/s$) upsets this balance

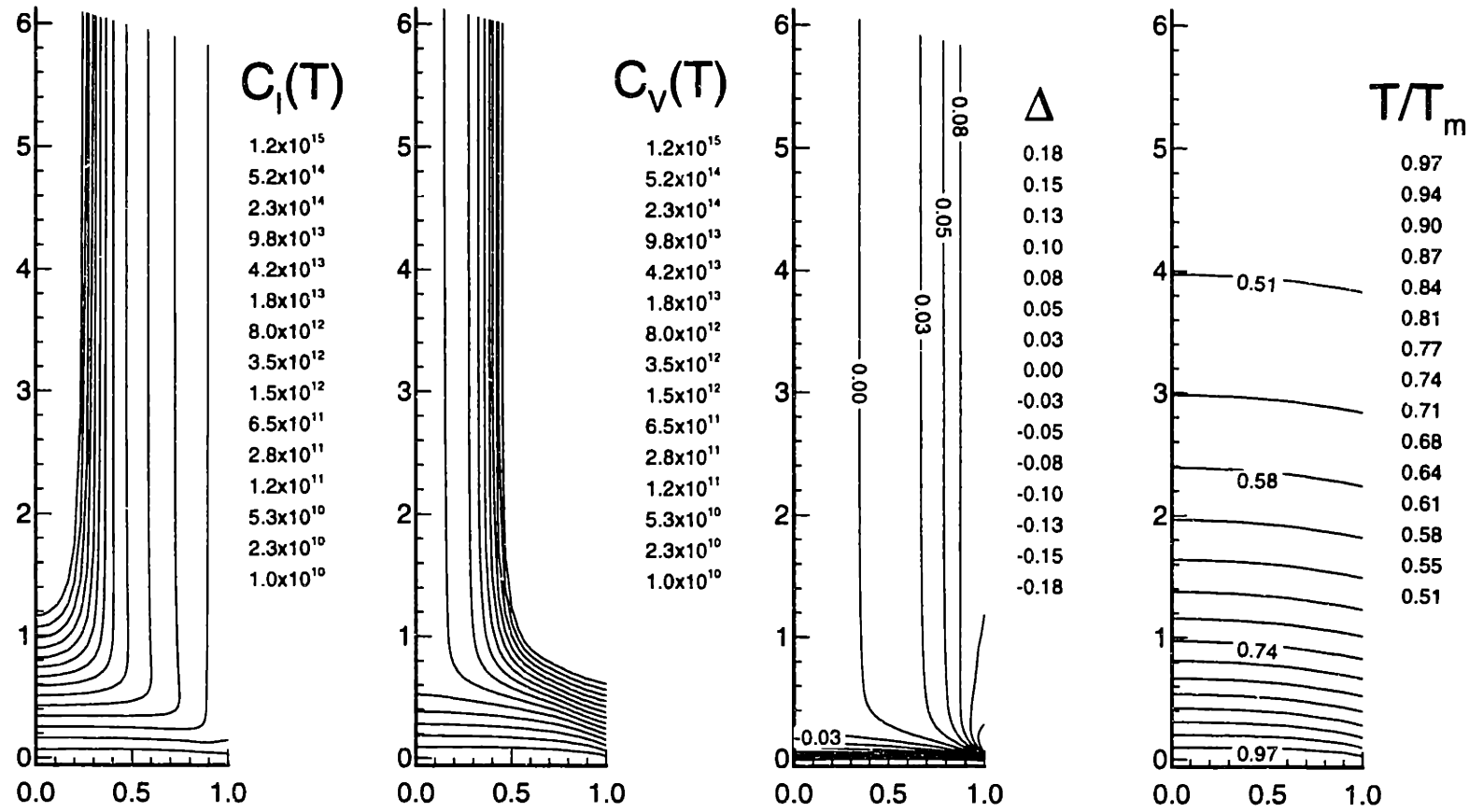

Figure 5-8: Profiles for C_I, C_V, Δ, and T for the case HS1/6" at a pull rate of 1.0 mm/min. Axial and radial coordinates are scaled with the radius of the crystal.

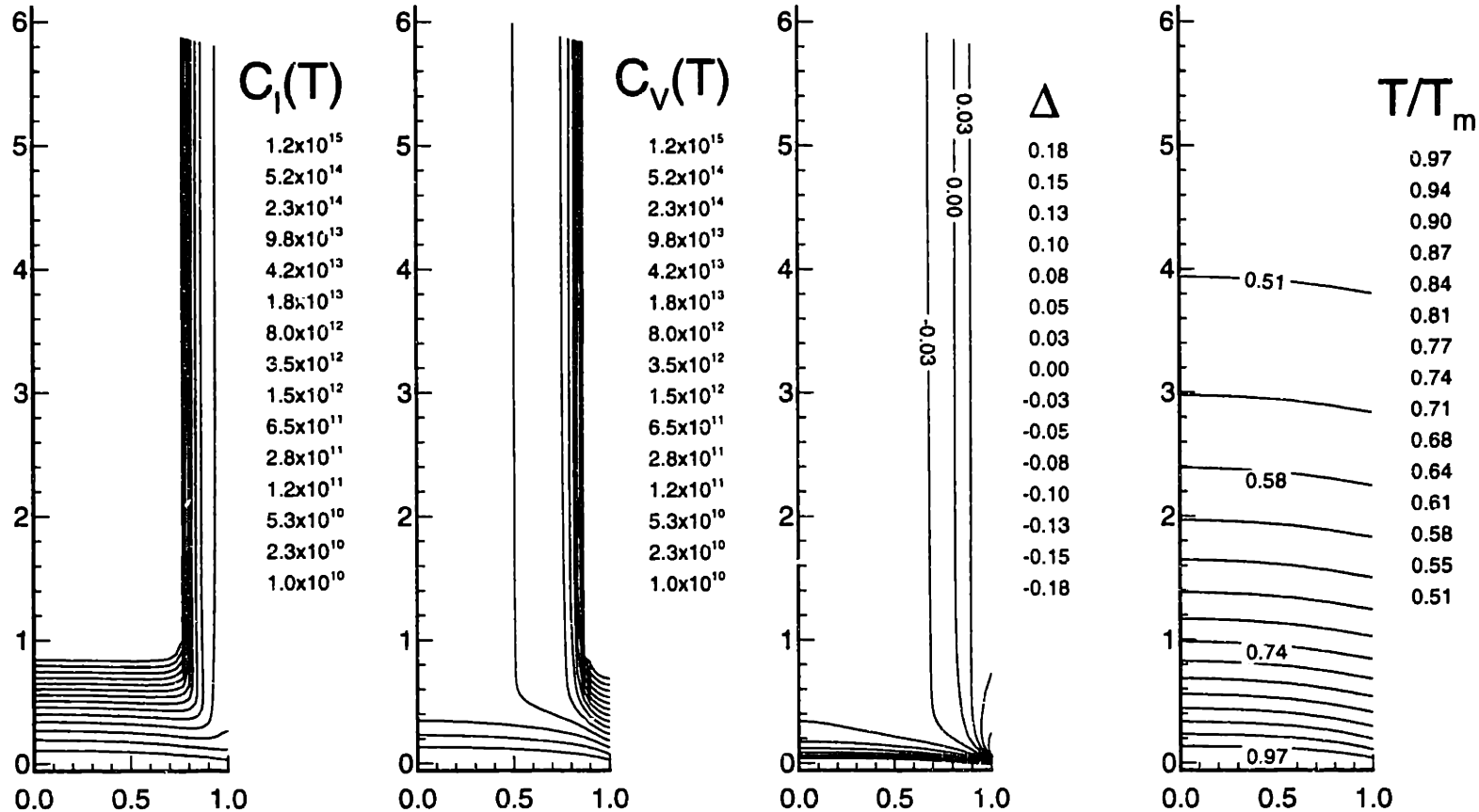

Figure 5-9: Profiles for C_I, C_V, Δ, and T for the case HS1/6" at a pull rate of 1.4 mm/min. Axial and radial coordinates are scaled with the radius of the crystal.

for $z > 0$. The increasing driving force for recombination leads to rapid recombination within a small distance of the interface and the excess vacancy concentration rapidly dominates.

The parameter $\Delta(0, z)$ has its minimum value at the interface due to the unequal point defect incorporation at the melt/crystal interface. It increases away from the interface as the faster diffusing interstitials increase their presence. Along the crystal axis ($r = 0$) the interstitials are eventually annihilated and $\Delta(0, z_{TOP}) \rightarrow C_V(0, z_{TOP})$ because $C_I(0, z_{TOP}) \rightarrow 0$.

The radial structure of the point defect concentration fields is driven entirely by the temperature field. The axial temperature gradient near the melt/crystal interface increases monotonically from the center of the crystal ($r = 0$) to the cylindrical surface ($r = R$) because of heat loss from the surface. The details of the axial temperature gradient near the melt/crystal interface are central to the formation of the point defect concentration profiles because of the driving force for recombination supplied by the decreasing equilibrium concentrations with temperature. Near the centerline ($r = 0$), the axial temperature gradient is shallow, so that the equilibrium point defect concentrations decrease relatively slowly limiting the recombination rate and permitting both defect species time to diffuse from the interface. Hence, the vacancies, which are incorporated into the crystal at a higher concentration eventually dominate the interstitial concentration. Closer to the crystal surface ($r = R$) the temperature gradient is steeper and the equilibrium concentrations fall more rapidly driving the recombination rate up through increased point defect supersaturation. This effect enhances the diffusion-limited point defect recombination where the slower moving vacancies become the less concentrated species and eventually leads to interstitial concentrations that are largest. At a critical radial position these two scenarios are balanced and the point defect concentrations are essentially equal, so that $\Delta \sim 0$.

An important prediction of the analysis is the final vacancy concentrations inside of the defect neutral region ($\Delta \sim 0$). Commercial crystals are always grown under conditions that result in the entire crystal being vacancy dominated, with the OSF-ring at the crystal cylindrical surface. However, minimization of void defects, and thus

vacancy supersaturation, is highly desirable as voids have been linked to increased device failure rates (von Ammon *et al.*, 1996). The values of the concentrations at the top of the crystal, $C_V(0, z_{TOP})$, are plotted in Figure 5-10 as a function of OSF-ring location, and in Figure 5-11 as a function of growth rate, for each crystal diameter and Czochralski system studied here. Characteristic values of $C_V(0, z_{TOP})$ are between 1×10^{12} cm^{-3} and 1×10^{14} cm^{-3} inside the OSF-ring for the point defect properties described in eqs. (5.89)-(5.93). The estimates for $C_V(0, z_{TOP})$ in Figure 5-10 all collapse onto each other when plotted as a function of normalized OSF-ring position. This concentration of vacancies would be available to aggregate into the voids that are observed experimentally (Itsumi *et al.*, 1995). Understanding the relationship between the final vacancy concentration and the size distribution of vacancy aggregates is critical for void defect control. Our current predictions indicate that the crystal should be grown with the lowest possible growth rate for which the OSF-ring is still located at the periphery of the crystal. The present calculations will allow us to minimize the final vacancy concentration with respect to a particular thermal field, but will not give quantitative information on the actual sizes of the resulting vacancy aggregates. The final distribution of aggregates is a complicated function of the thermal history of the crystal and cannot be inferred from the frozen-in vacancy concentration alone. Chapter 6 is devoted to the development of a detailed nucleation model for point defects that will allow for the computation of aggregate size distribution.

5.4.2 Thermophysical Properties of Point Defects

The values of the point defect properties in eqs. (5.89)-(5.93) are compared here to the Stillinger-Weber estimates described in Chapter 3 and to the results of the OSF-ring analysis of Habu et al. (1996). The estimates for the diffusivities of self-interstitials and vacancies are compared in Figures 5-12 and 5-13, respectively. The fitting to OSF-ring data results in little change in the vacancy diffusivity but increases the interstitial diffusivity by over an order-of-magnitude. The fitting procedure described here predicts that vacancy and interstitial diffusion proceeds at roughly equal rates

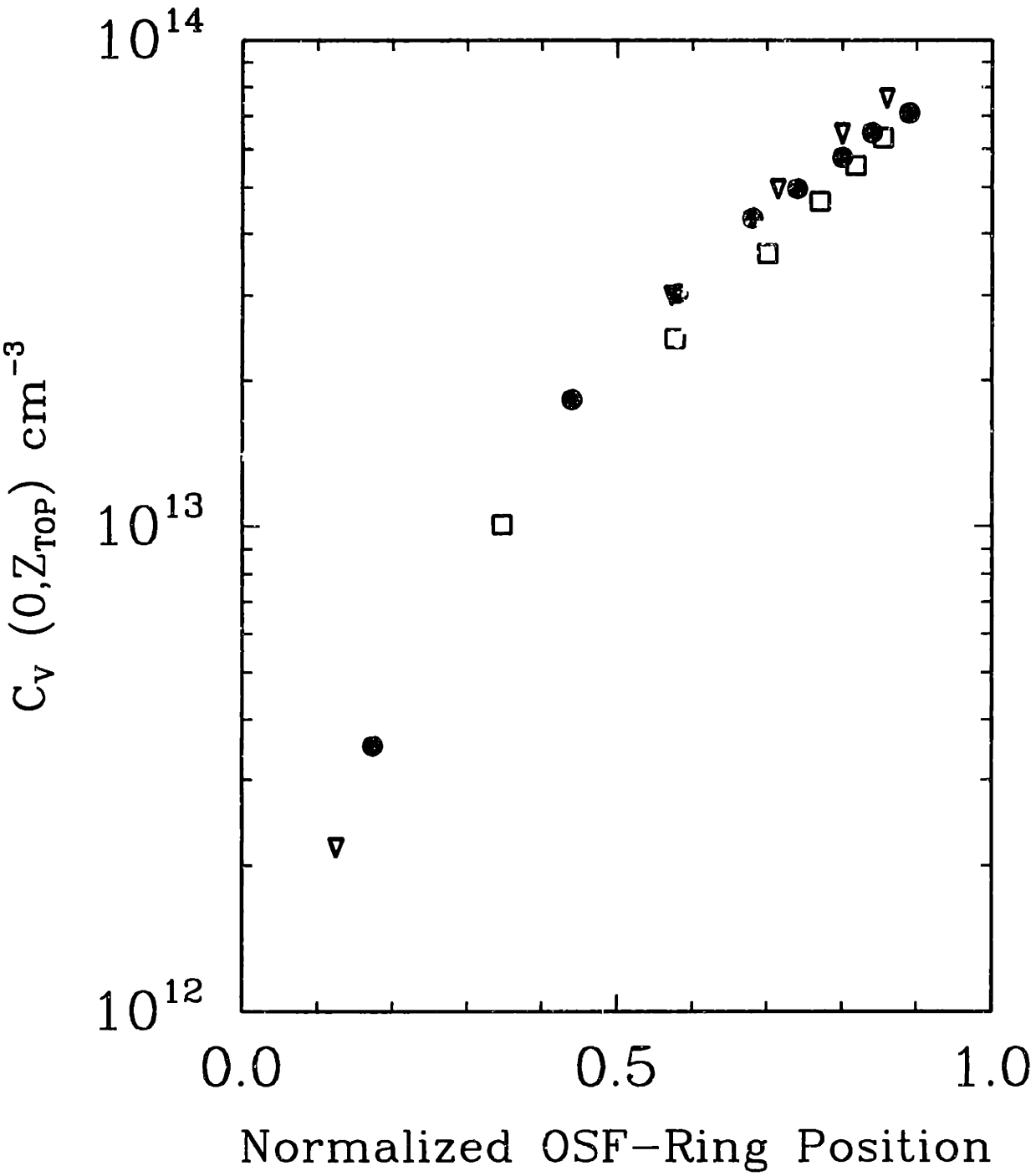

Figure 5-10: Frozen-in concentrations of vacancies at the top of the crystal centerline as a function of OSF-ring position for the three experimental cases; • = HS1/4", □ = HS1/6", ▽ = HS2/8".

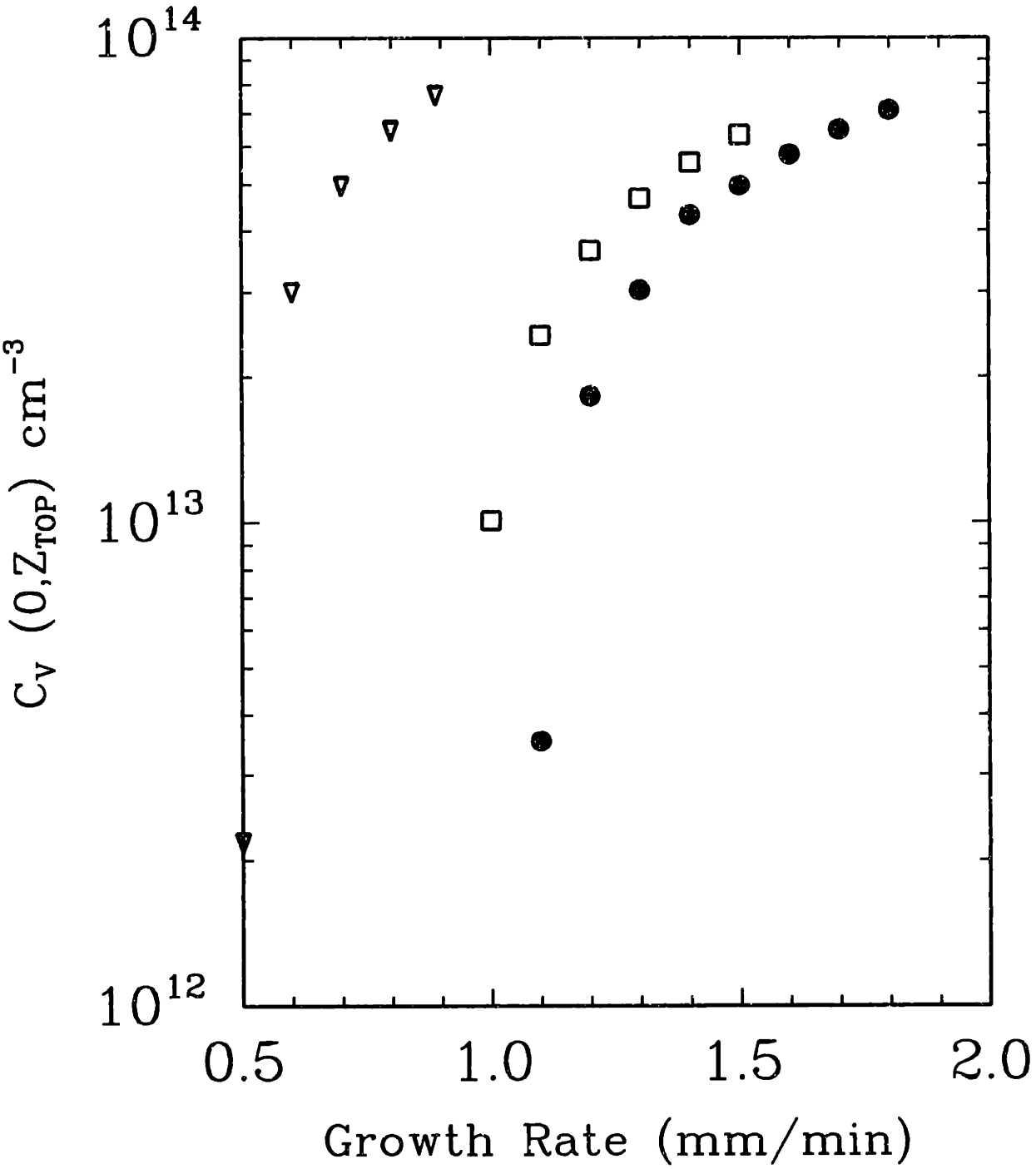

Figure 5-11: Frozen-in concentrations of vacancies at the top of the crystal centerline as a function of growth for the three experimental cases; • = HS1/4", □ = HS1/6", ▽ = HS2/8".

near the melting temperature, $D_I(T_m) \sim D_V(T_m) \sim O(10^{-4})\ cm^2/s$. Although these estimates are generally at the high end of the range defined by other estimates in the literature, the vacancy diffusion coefficient used by Habu et al. (1996) is still higher by over one order of magnitude, reaching $3 \times 10^{-3}\ cm^2/s$ at T_m. This is in contrast to the very close agreement for the interstitial diffusivity across the entire temperature range.

In the case of the equilibrium concentrations, adjusting the SW estimates to OSF-ring data did not change the interstitial value but resulted in a two order-of-magnitude drop in the vacancy estimate. In both cases, there is close agreement to the estimates of Habu et al. (1996), especially at high temperature. For both point defects, the current estimates represent the lower end of the range found experimentally although the agreement is generally good. The equilibrium concentrations are plotted in Figures 5-14 and 5-15 for self-interstitials and vacancies, respectively.

Comparison with the self-diffusion components, $D_I C_I^{eq}$ and $D_V C_V^{eq}$, shows more clearly the improvement to the SW estimates that the present analysis has provided. The parameter values used in the analysis here give for $D_I C_I^{eq}$,

$$D_I C_I^{eq} = 7.2 \times 10^{22} \exp\left(-\frac{1.4 + 3.85 \times 10^{-3} T}{k}\right) \exp\left(-\frac{4.4 + 3.08 \times 10^{-4} T}{kT}\right) \quad cm^{-1}s^{-1} \tag{5.95}$$

At the melting temperature our estimate $D_I(T_m)C_I^{eq}(T_m) = 3.6 \times 10^{11}\ cm^{-1}s^{-1}$ agrees well with the prediction of Stolwijk et al. (1986), $D_I(T_m)C_I^{eq}(T_m) = 1.5 \times 10^{11}\ cm^{-1}s^{-1}$. The predictions of $D_I(T)C_I^{eq}(T)$ shown in Figure 5-16 all agree well at high temperature with the exception of the Stillinger-Weber value. The result of Habu et al. (1996) also is slightly lower at lower temperatures.

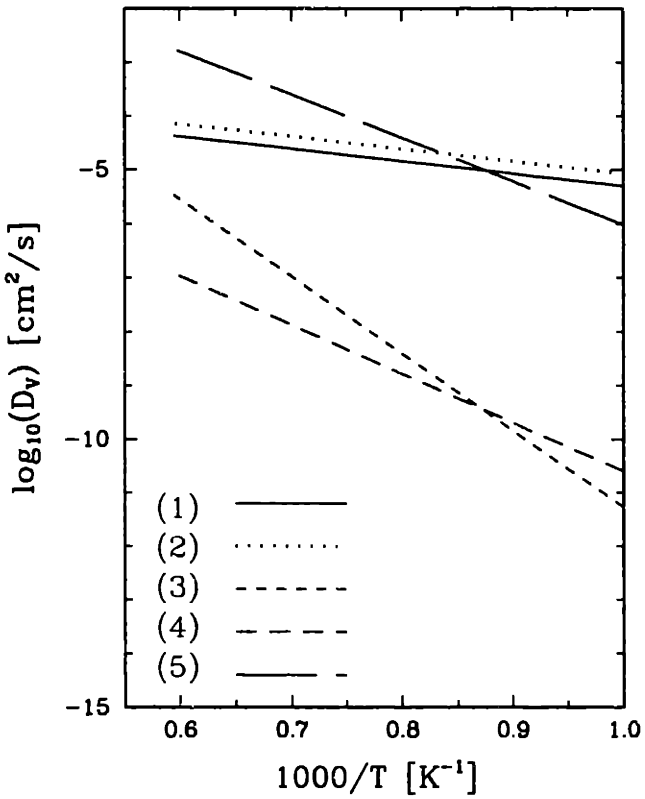

Figure 5-13: Vacancy diffusivities as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) Zinc Diffusion (Bracht *et al.*, 1995);
(5) OSF-ring Fitting (Habu and Tomiura, 1996)

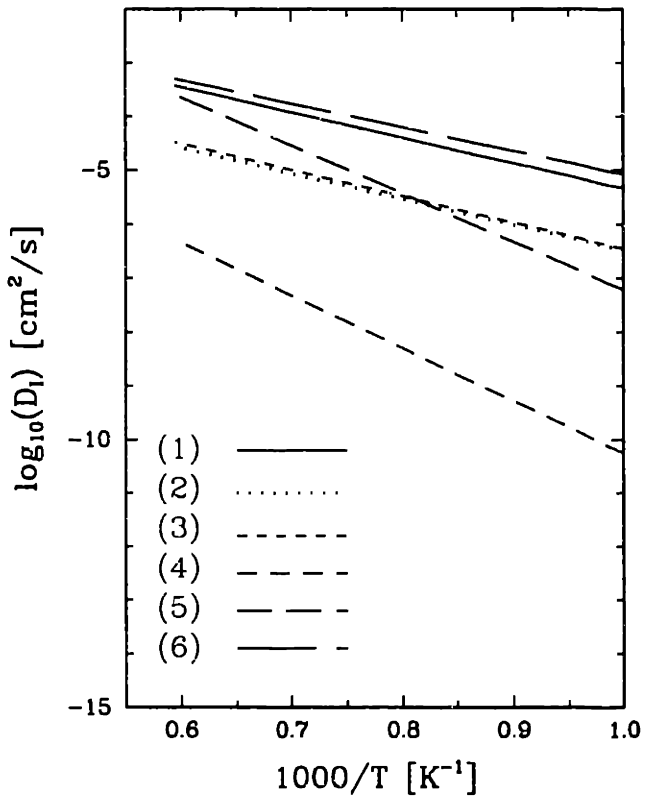

Figure 5-12: Self-interstitial diffusivities as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) OED/ORD (Wijaranakula, 1990);
(5) Zinc Diffusion (Bracht *et al.*, 1995);
(6) OSF-ring Fitting (Habu and Tomiura, 1996)

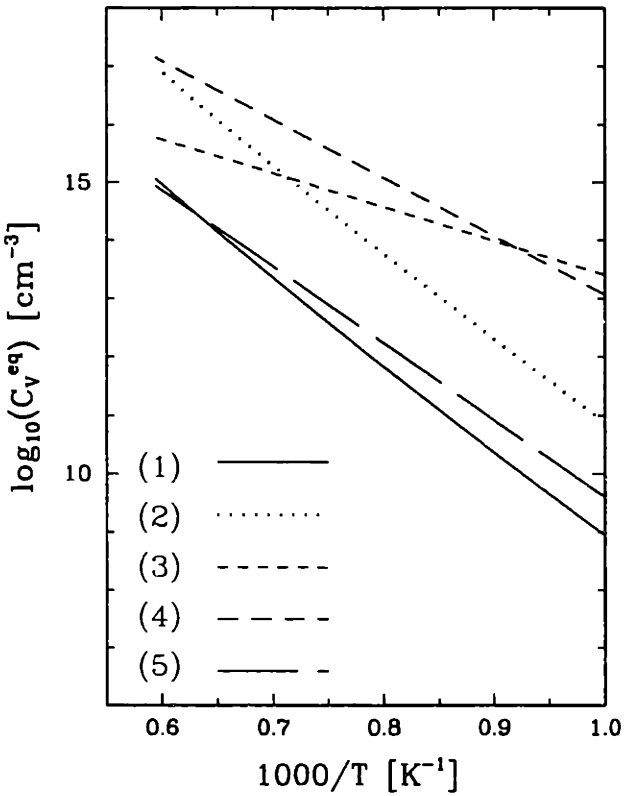

Figure 5-15: Vacancy equilibrium concentrations as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) Zinc Diffusion (Bracht *et al.*, 1995);
(5) OSF-ring Fitting (Habu and Tomiura, 1996)

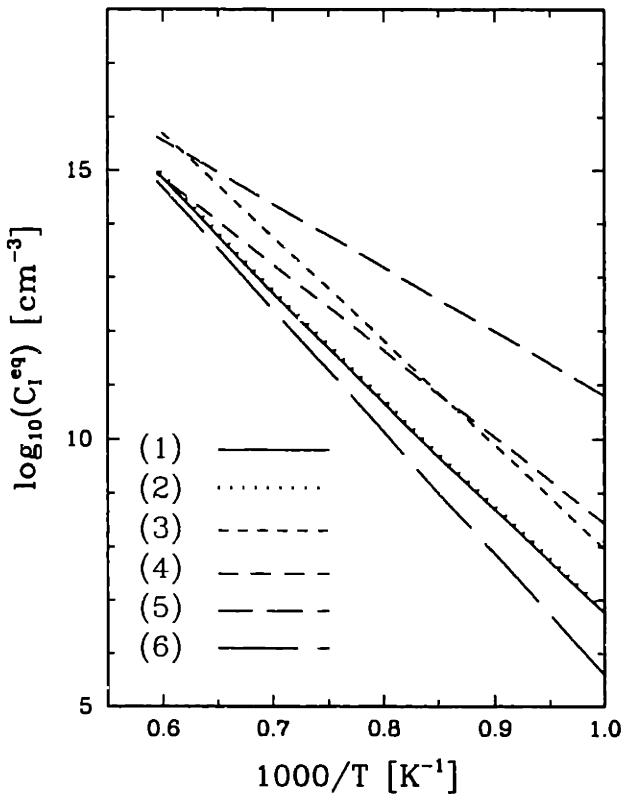

Figure 5-14: Self-interstitial equilibrium concentrations as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) Zinc Diffusion (Bracht *et al.*, 1995);
(5) OED/ORD Experiments (Park and Law, 1992);
(6) OSF-ring Fitting (Habu and Tomiura, 1996)

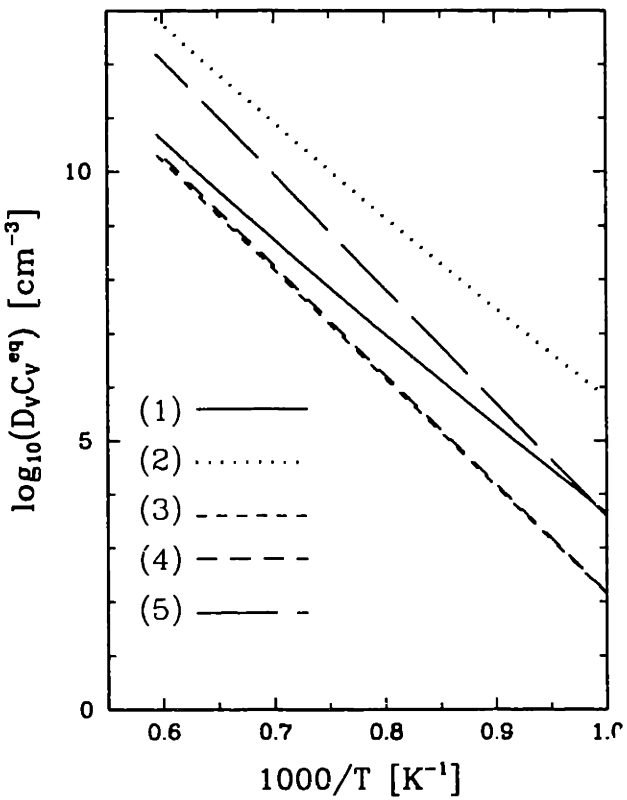

Figure 5-17: Vacancy self-diffusion component as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) Gold Diffusion (Gösele and Tan, 1983);
(5) OSF-ring Fitting (Habu and Tomiura, 1996)

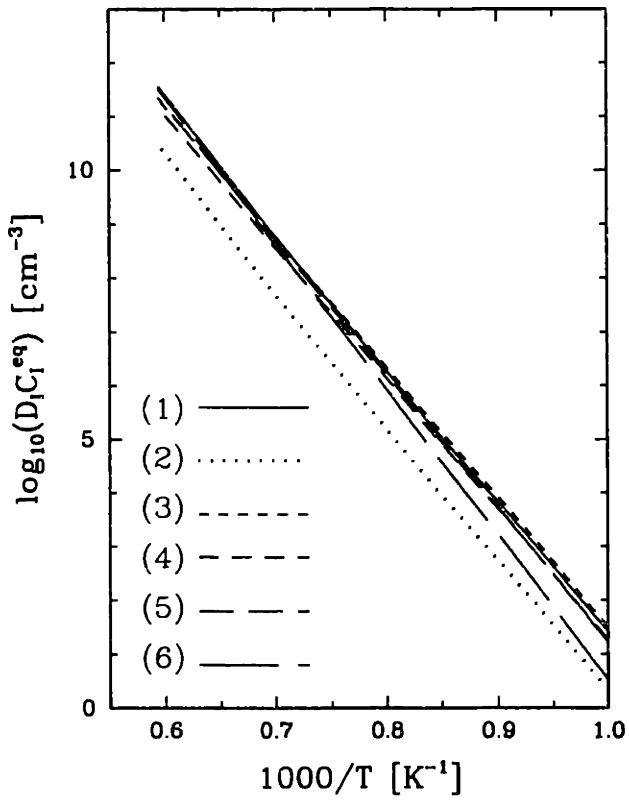

Figure 5-16: Interstitial self-diffusion component as a function of temperature:
(1) Current OSF-ring dynamics fitting;
(2) Stillinger-Weber Potential;
(3) Gold Diffusion (Zimmermann and Ryssel, 1992);
(4) Gold Diffusion (Stolwijk *et al.*, 1986);
(5) Zinc Diffusion (Bracht *et al.*, 1995);
(6) OSF-ring Fitting (Habu and Tomiura, 1996)

Our estimate for $D_V C_V^{eq}$, is consistent with predictions from metal gettering, especially at high temperature. The value used by Habu et al. (1996) disagrees by approximately two orders-of-magnitude, as shown in Figure 5-17. This discrepancy results in a larger vacancy contribution to self-diffusion at the melting temperature, and is at variance with most estimates today. It is generally accepted that $D_I C_I^{eq} > D_V C_V^{eq}$ near the melting temperature while $D_I C_I^{eq} < D_V C_V^{eq}$ at lower temperature with the cross-over taking place in the temperature range 1350-1400 K. The predictions of Habu et al. (1996) for point defect dynamics depend critically on this parameter. The large estimate for the vacancy contribution leads to radial outdiffusion of vacancies, determining the radial structure of the point defect fields *independently* of the radial structure of the temperature field. This large diffusivity is also important for increasing the magnitude of the vacancy thermodiffusion effect relative to the interstitial term which provides the mechanism for the transition between interstitial and vacancy dominated regions in the model of Habu et al. (1996).

The ratio of the interstitial and vacancy contributions to self-diffusion, $D_I C_I^{eq}/D_V C_V^{eq}$ is valuable for comparing our thermophysical properties to the experimental values at the melting temperature. This quantity is useful because it is not a function of the absolute concentrations but rather their ratio, which is an independent parameter in our model as discussed in Section 5.4.1. The thermophysical properties described in eqs. (5.89)-(5.93) give $D_I C_I^{eq}/D_V C_V^{eq} \approx 7.2$ at $T = T_m$, which compares well with the experimental value of 5.7 (Tan and Gösele, 1985; Stolwijk *et al.*, 1986).

Interestingly, if we compare two parameter sets of equally good fit for the objective function, χ, eq. (5.88) but with the objective functions, χ_G, eq. (5.94) we find that while our best estimate gives a value of 7.2 ($\chi_G = 0.15$), another data set predicts $D_I C_I^{eq}/D_V C_V^{eq} \sim 22$ ($\chi_G = 0.2$). In other words, the parameter set that best reproduces the OSF-ring data, i.e. minimizes both χ and χ_G, also leads to the best agreement with the literature estimate for $D_I C_I^{eq}/D_V C_V^{eq}$.

The results shown in Figures 5-15 and 5-17 point out that, in general, agreement between estimates for the vacancy properties become worse at lower temperatures. As will be shown in Section 5.6, OSF-ring dynamics are almost entirely determined

by high temperature properties so that the analysis here is not sensitively dependent on low temperature point defect properties. This motivates the use of the migration and formation energies computed by atomistic simulation. Clearly, in the case of the interstitial there is good agreement for these energies. However, vacancy activation energies are much more uncertain. Most estimates for the vacancy formation energy lie in the range of 2.7-3.5 eV as discussed in Chapter 3, which is consistent with our estimate of about 3.1 eV at $T = T_m$. There is less agreement on the vacancy migration energy; if the self-diffusion activation energy, $H_V^m + H_V^f$ is assumed to be accurate, the vacancy formation energies would leave about 0.5-1.4 eV for the migration energy. Our value of 0.46 eV is at the lower bound of this range, but there have been independent estimates that are lower by 0.3 eV (Watkins *et al.*, 1979; Tang *et al.*, 1997).

The results presented in Section 5.4 are not sensitively dependent on the magnitude of the recombination rate. As will be demonstrated in Section 5.6, the temperature dependence of the recombination rate does not critically affect the predictions of the model. In fact, the present model only requires the presence of "fast" recombination kinetics in the region near the melt/solid interface; this concept will be defined quantitatively in Section 5.6.

5.5 Sensitivity Analysis

5.5.1 Effect of Equilibrium Boundary Conditions at the Radial Surface

The excellent agreement between the predictions of the model and the experimental measurements of the OSF-ring require only the radial dependence of the temperature field introduced through the heat transfer conditions in the furnace. The exact nature of the boundary conditions on the cylindrical surface of the crystal, as described by either of eqs. (5.63), (5.64), or (5.65), does not matter. Previous analyses, such as Brown et al. (1994) or Habu and Tomiura (1996), assume that equilibrium con-

ditions hold on the crystal surface to produce a radial gradient for out-diffusion of point defects, which induces radial structure. This mechanism is not required in our description, Calculations reported above were performed with the no-flux condition on point defects, eq. (5.63).

The effect of introducing additional radial structure on our predictions was tested by performing calculations with the boundary condition eq. (5.64) for the HS1/4" system with the same set of thermophysical parameters. The prediction for the radial location for $\Delta(R_{OSF}, z_{TOP}) = 0$ is shown in Figure 5-18 and compared to the result predicted for the no-flux condition. For $R_{OSF} < 0.7$ the predictions are in very close agreement, but diverge as the OSF-ring approaches the crystal surface.

The causes for this divergence are understood by considering the structure of the defect profiles near the surface. With the no-flux conditions, the interstitial concentration is generally highly supersaturated near the surface while the vacancies are depleted. Forcing point defect equilibrium at the surface creates a driving force for interstitial out-diffusion while *injecting* vacancies, thus shifting the radial location for $\Delta = 0$ towards the surface. The penetration distance of this effect is thus related to the point defect diffusivities. We estimate the diffusion length of each point defect species by computing its mean displacement during crystal growth as

$$L_i \approx 2\sqrt{\int_{T_m}^{T} D_i(T)dt}\,, \qquad i = I, V. \tag{5.96}$$

For the point defect properties and growth conditions considered in this study, eq. (5.96) gives diffusion lengths of about 4 mm and 1 mm for interstitials and vacancies, respectively. Clearly, the effect of the boundary condition penetrates further, as shown in Figure 13. This longer range effect is due to the very fast recombination kinetics, which forces the point defect concentrations to approximately satisfy the mass action law, $C_I C_V = C_I^{eq} C_V^{eq}$. As the point defects become supersaturated in the cooling crystal, recombination becomes very rapid; thus wherever an imbalance in the concentrations exists, the species present in smaller quantities will be driven towards extinction. Only in the special case of roughly equal concentrations will both species

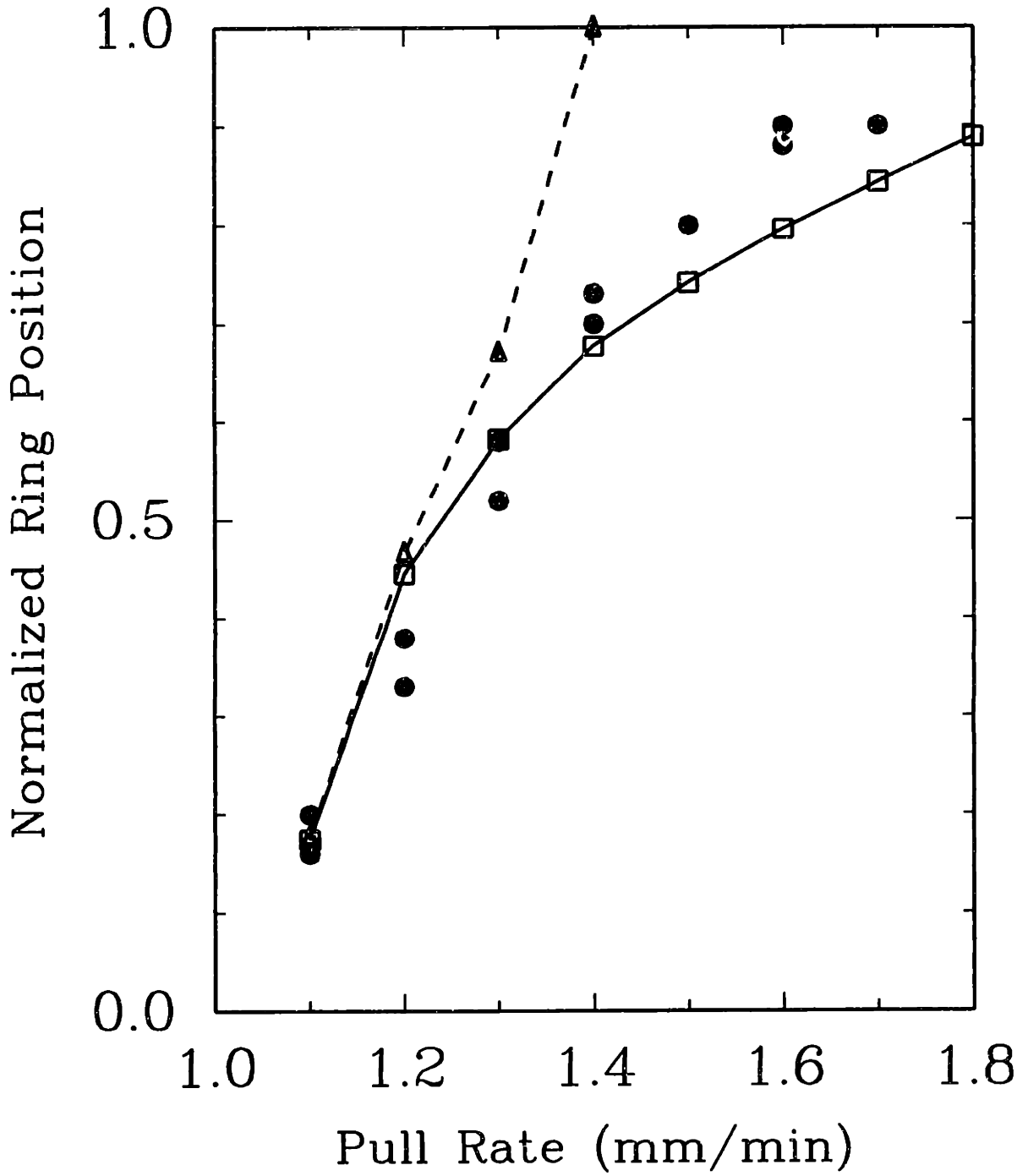

Figure 5-18: Effect of equilibrium concentration boundary conditions on the radial surface. • = Experimental Results; □ = No-flux Conditions; △ = Equilibrium Conditions

survive with significant concentrations. The location of this balance is very sensitive to the point defect concentrations, as is observed by the rapid variation of the OSF-ring radius with small changes in the growth rate. Hence, the small change introduced by imposing equilibrium at the surface propagates farther into the crystal.

Undoubtedly, the appropriate boundary conditions on the surface of the crystal is neither the no-flux nor equilibrium state, but described on a mixed condition such as eq. (5.65) where the parameters $\{\gamma_i\}$ represent kinetic barriers for recombination at the crystal surface. For example, such a barrier may be due to the limitations of surface transport or the presence of an oxide film.

5.5.2 Effect of Thermodiffusion on the Point Defect Model

Without reasonable estimates for the reduced heats of transport, $\{Q_i^{*\prime}\}$, $i = I, V$, the influence of point defect transport by thermodiffusion can only be investigated by calculations that explore parametric sensitivity. We demonstrate the insensitivity of the predictions to the values of $Q_I^{*\prime}$ and $Q_V^{*\prime}$ by calculations with the point defect properties used in Section 5.4.1 and with $Q_I^{*\prime} = Q_V^{*\prime} = 1.0\ eV$. Although arbitrary, this value is the same order-of-magnitude as the formation energies for each species.

The prediction of the variation of the OSF-ring, given by $\Delta(R_{OSF}(V), z_{TOP}) = 0$, for the system HS1/4" is shown in Figure 13 along with the experimental results and the predictions using $Q_I^{*\prime} = Q_V^{*\prime} = 0$. The incorporation of the thermodiffusion effect moves $\Delta(R_{OSF}(V))$ outwards by about 0.5 R until it disappears at the cylindrical surface for $V = 1.4mm/min$. This observation is explained by fact that the thermodiffusion effect is proportional to the point defect diffusivities; the higher interstitial diffusivity leads to greater interstitial depletion due to the thermodiffusion of interstitials towards the melt/solid interface. Once again, the high recombination kinetics lead to high sensitivity to perturbations in the relative point defect concentrations.

The predictions with and without thermodiffusion are brought in much closer agreement by adjusting the point defect properties to maintain the best fit to the reference set of experiments. Simply adjusting the interstitial equilibrium concentration gives the results shown in Figure 13, very close to the original predictions;

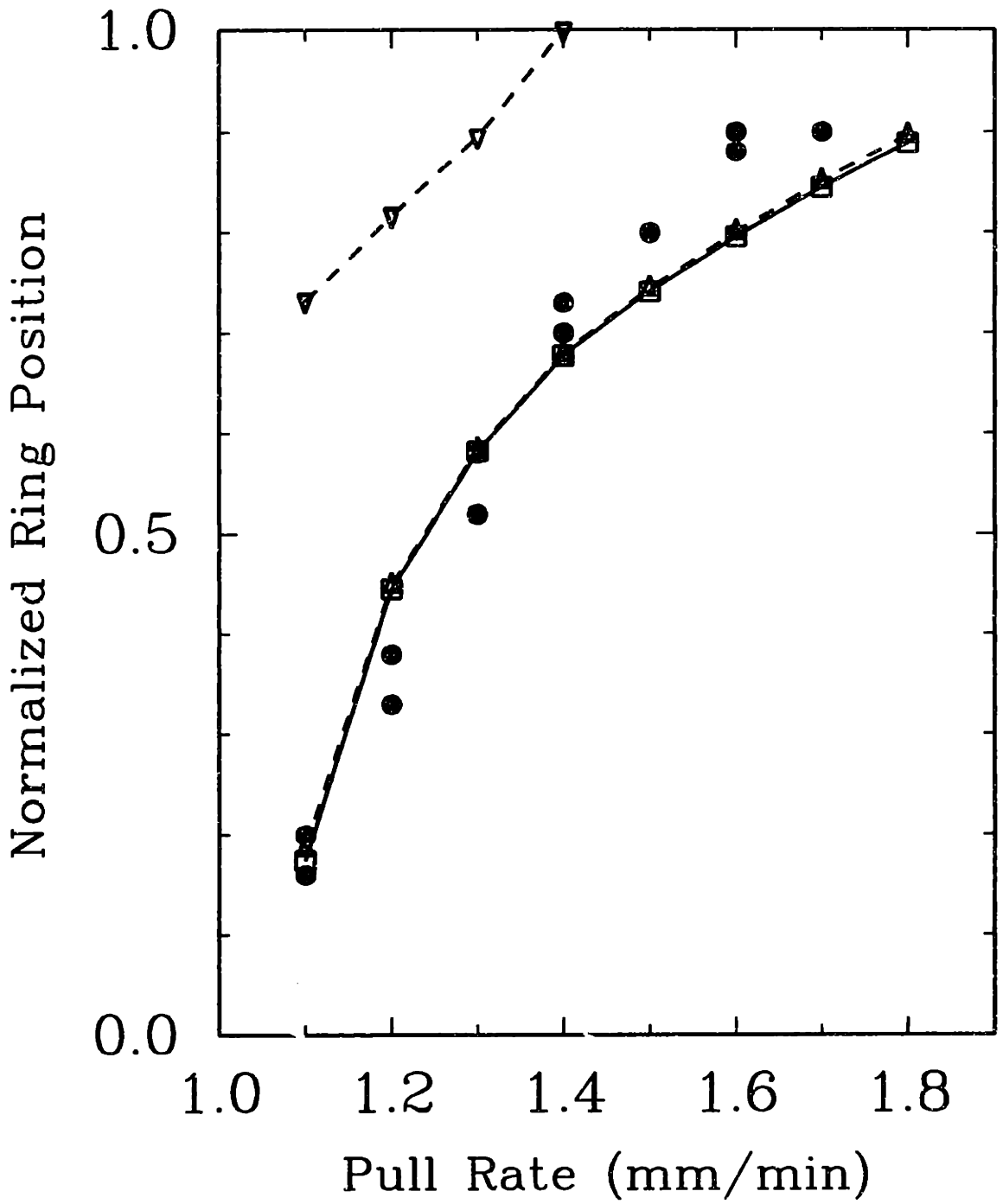

Figure 5-19: Effect of assigning $Q_I^{*\prime} = Q_V^{*\prime} = 1\ eV$ on the model prediction after re-optimization of the interstitial equilibrium concentration. • = experimental results, $\square \rightarrow Q_I^{*\prime} = Q_V^{*\prime} = 0$; $\triangledown \rightarrow Q_I^{*\prime} = Q_V^{*\prime} = 1\ eV$ with no adjustment in $C_I^{eq}(T_m)$; and $\triangle \rightarrow Q_I^{*\prime} = Q_V^{*\prime} = 1\ eV$ with a 5% increase in $C_I^{eq}(T_m)$

.

only a 5% increase in C_I^{eq} is predicted from the value used neglecting thermodiffusion. This result shows that the predictions of the model for $\Delta(R_{OSF}(V)$ are much more sensitive on the equilibrium concentrations than to the values of the reduced heats of transport. While potentially having a quantitative influence on the predictions of the analysis and on the values of the thermophysical properties used in the analysis, adding thermodiffusion does not qualitatively change the predictions of the analysis.

5.6 Perturbation Analysis of the Point Defect Model

The success of the comparison between the numerical predictions of the point defect dynamics model and the response of the OSF-ring location to the temperature field and growth rate in the crystal give confidence in the physical picture of defect transport and recombination represented by the model. Several results from this study, as well as the simple version of the model proposed by Voronkov (1982) suggest that asymptotic analysis of the model may be feasible and lead to very simple results.

The foremost clue that an asymptotic structure exists is the structure of the point defect concentration fields, as shown in Figures 5-8 and 5-9. These clearly demonstrate boundary layers adjacent to the melt/crystal interface in which the majority of the axial variation of the concentration fields occurs, independent of the radial position. The scaling analysis in Section 5.6.1 suggests the occurrence of the boundary layer as a consequence of the very high recombination rate at elevated temperatures and the exponential dependence of the equilibrium concentrations, eqs. (5.91) and (5.92), and the kinetic rate coefficient, eq. (5.93), on temperature. These facts lead to the matched asymptotic perturbation analysis described in Section 5.6.2.

Finally, the tantalizingly simple correlation, $(V/G)_{crit} \approx 1.34 \times 10^{-3}\ cm^2/minK$, eq. (5.1), suggests that the location of the OSF-ring, i.e. the radial location for $\Delta = 0$, can be reduced to a simple expression in terms of the interface temperature gradient. Such an expression is derived in Section 5.6.4 and linked to the model of Voronkov (1982), discussed in Section 5.2.3.

5.6.1 Scaling Analysis

The point defect dynamics balance equations, eqs. (5.8) and (5.9), and the interfacial boundary condition, eqs. (5.59) and (5.60), are put in dimensionless form by introducing scaled variables. Using the coordinate system shown in Figure 3, where the melt/crystal interface corresponds to $z = 0$ and $0 \leq r \leq R$, the concentration fields are scaled with the interstitial equilibrium concentrations at the interface, which corresponds to the melting temperature, the radial coordinate with R, and the axial coordinate with a yet unknown length scale, L, for the length scale for axial variation of the concentration fields. These variables are

$$\hat{z} \equiv z/L, \qquad \hat{r} \equiv r/R, \tag{5.97}$$

$$\Theta(\hat{r},\hat{z}) \equiv \frac{T(\hat{r},\hat{z})}{T_m}, \qquad a(\hat{r},\hat{z}) \equiv \frac{C_I(\hat{r},\hat{z})}{C_I^{eq}(1)}, \qquad b(\hat{r},\hat{z}) \equiv \frac{C_V(\hat{r},\hat{z})}{C_I^{eq}(1)} \tag{5.98}$$

Introducing these scales into eqs. (5.8), (5.9), (5.59), and (5.60) yields

$$\frac{1}{Da}\left[-Pe\frac{\partial a}{\partial \hat{z}} + \frac{\partial}{\partial \hat{z}}\left(d_I\frac{\partial a}{\partial \hat{z}}\right) + \omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}d_I\frac{\partial a}{\partial \hat{r}}\right)\right]$$
$$-\frac{1}{Da}\left[\frac{\partial}{\partial \hat{z}}\left(\Phi_I d_I a\right) + \frac{\omega}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}\Phi_I d_I\frac{\partial\Theta/\partial r}{\partial\Theta/\partial z}a\right)\right] = K(ab - a^{eq}(\Theta)b^{eq}(\Theta)) \tag{5.99}$$

$$\frac{1}{Da}\left[-Pe\frac{\partial b}{\partial \hat{z}} + \frac{\partial}{\partial \hat{z}}\left(d_V\frac{\partial b}{\partial \hat{z}}\right) + \omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}d_V\frac{\partial b}{\partial \hat{r}}\right)\right]$$
$$-\frac{1}{Da}\left[\frac{\partial}{\partial \hat{z}}\left(\Phi_V d_V b\right) + \frac{\omega}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}\Phi_V d_V\frac{\partial\Theta/\partial r}{\partial\Theta/\partial z}b\right)\right] = K(ab - a^{eq}(\Theta)b^{eq}(\Theta)) \tag{5.100}$$

and

$$a(\hat{r},0) = 1, \qquad b(\hat{r},0) = \frac{C_V^{eq}(1)}{C_I^{eq}(1)}, \qquad 0 \leq \hat{r} \leq 1, \tag{5.101}$$

where,

$$\omega \equiv \frac{L}{R} \tag{5.102}$$

is the ratio of the length scales, the Damköhler number, Da,

$$Da \equiv \frac{L^2 k_{IV}(1) C_I^{eq}(1)}{D_I(1)}, \tag{5.103}$$

scales the recombination rate over diffusion at the melting temperature and the Peclet number, Pe,

$$Pe \equiv \frac{VL}{D_I(1)}, \tag{5.104}$$

scales the convection of point defects by crystal motion against diffusion. Dimensionless, temperature-dependent equilibrium concentrations, diffusivities, and recombination rate coefficient are defined as

$$a^{eq}(\Theta) \equiv \frac{C_I^{eq}(\Theta)}{C_I^{eq}(1)}, \qquad b^{eq}(\Theta) \equiv \frac{C_V^{eq}(\Theta)}{C_I^{eq}(1)}, \tag{5.105}$$

$$d_I(\Theta) \equiv \frac{D_I(\Theta)}{D_I(1)}, \qquad d_V(\Theta) \equiv \frac{D_V(\Theta)}{D_I(1)}, \tag{5.106}$$

$$K(\Theta) = \frac{k_{IV}(\Theta)}{k_{IV}(1)}. \tag{5.107}$$

The dimensionless group arising from the scaling of the thermodiffusive flux is

$$\Phi(\Theta) \equiv \frac{Q^{*\prime} L G^*}{k T_m^2} \left(\frac{1}{\Theta^2}\right), \tag{5.108}$$

where G^* scales the gradient in the temperature field in the crystal.

The cause of the boundary layer in the concentration fields is seen immediately from eqs. (5.99) and (5.100) and typical values of the dimensionless groups. For the thermophysical properties given in eqs. (5.89)-(5.93) they are

$$Da = O(10^6), \qquad Pe = O(1), \tag{5.109}$$

so that near the interface recombination is the dominant mechanism for modifying the point defect fields. In the limit $Da \to \infty$ all derivatives scale from the balance equations, eqs. (5.99) and (5.100), and the far field boundary conditions, eq. (5.63),

can only be satisfied by constructing a singular perturbation analysis in which the axial distance is rescaled so that diffusion balances recombination in a small region near the interface. This is performed in Section 5.6.2.

The analysis is complicated by the exponential dependence of each of the functions defined in eqs. (5.105)-(5.107) on temperature; this leads to very rapid decreases of these coefficients when T decreases significantly from T_m. We use concepts developed in combustion theory (Kapila, 1983) for high activation energies to include this effect in the asymptotic analysis and to define the axial length scale L.

The temperature field close to the interface is approximately given by a linear axial gradient that varies slowly with r, i.e.

$$T(r,z) = T_m - G^* g(r) z, \tag{5.110}$$

or in dimensionless form

$$\Theta(\hat{r},\hat{z}) = 1 - \left(\frac{G^* L}{T_m}\right) g(\hat{r})\hat{z}, \tag{5.111}$$

where $g(\hat{r})$ is a dimensionless function that describes the radial variation of the axial gradient. Assume that $g(\hat{r}) = O(1)$ so that the constant G^* scales the magnitude of the gradient.

Each function given by eqs. (5.105)-(5.107) has the form

$$Q(\Theta) = Q_0 \exp\left(-\frac{\gamma^*}{\Theta}\right), \tag{5.112}$$

where γ^* is a dimensionless activation energy, $\gamma^* \equiv \gamma/kT_m$. It is assumed that the temperature dependencies of Q_0 and γ^* are weak and that these quantities are constant across the boundary layers. Using eq. (5.111) in (5.112) and taking $G^*/T_m << 1$ gives

$$Q(\Theta) = Q_0 \exp\left[-\frac{\gamma^*}{1 - \frac{G^* L}{T_m} g(\hat{r})\hat{z}}\right] \approx \exp(-\gamma^*) \exp\left[-\gamma^* \left(\frac{G^* L}{T_m}\right) g(\hat{r})\hat{z}\right]. \tag{5.113}$$

For the functions given in eqs. (5.105)-(5.107) to be near unity gives

$$\frac{Q(\Theta)}{Q(1)} \sim \exp\left[-\gamma\left(\frac{G^* L}{T_m}\right) g(\hat{r})\hat{z}\right] \sim 1 \tag{5.114}$$

so that $\hat{z} = T_m/(\gamma^* G^* L) = 1$ is satisfied if

$$L \equiv \frac{T_m}{\gamma^* G^*}. \tag{5.115}$$

The new axial length scale is based on the length scale for the axial temperature gradient reduced by the high dimensionless activation energy. For typical Czochralski crystal growth conditions ($G^* = 80 \;\; K/cm$ and $R = 10 \;\; cm$), this scale becomes $L = 1.5 \;\; cm$ and the ratio of length scales ω is

$$\omega \equiv \frac{L}{R} = \frac{T_m}{\gamma^* R G^*} = 0.15. \tag{5.116}$$

It is assumed that $\omega << 1$ in the asymptotic analysis described in Section 5.6.2.

The magnitude of the thermodiffusive flux is analyzed in terms of L and G^* by substituting eq. (5.115) into eq. (5.108) to yield

$$\Phi_i(\Theta) = Q_i^{*\prime} \frac{T_m}{\gamma_{if}^* G^*} \cdot \frac{G^*}{k T_m^2}\left(\frac{1}{\Theta^2}\right) = \frac{Q_i^{*\prime}}{H_i^f}\frac{1}{\Theta^2}. \tag{5.117}$$

where $\gamma_{if}^* = H_i^f/kT_m$, $i = I, V$. Hence the importance of thermodiffusion scales as the ratio $Q_i^{*\prime}/H_i^f$. It is assumed that these scales are small and neglect thermodiffusion in the following analysis.

5.6.2 Matched Asymptotic Analysis

Intermediate Region: $z = O(L)$

An asymptotic solution valid for large Damköhler number ($Da >> 1$) and small ω ($\omega << 1$) is constructed by expanding $a(\hat{r}, \hat{z})$ and $b(\hat{r}, \hat{z})$ as Taylor series in both

parameters as

$$a(\hat{r},\hat{z}) = a_{0,0}(\hat{r},\hat{z}) + \epsilon a_{1,0}(\hat{r},\hat{z}) + \omega a_{0,1}(\hat{r},\hat{z}) + \epsilon\omega a_{1,1}(\hat{r},\hat{z}) + O(\epsilon^2,\omega^2) \qquad (5.118)$$

$$b(\hat{r},\hat{z}) = b_{0,0}(\hat{r},\hat{z}) + \epsilon b_{1,0}(\hat{r},\hat{z}) + \omega b_{0,1}(\hat{r},\hat{z}) + \epsilon\omega b_{1,1}(\hat{r},\hat{z}) + O(\epsilon^2,\omega^2) \qquad (5.119)$$

where $\epsilon \equiv Da^{-1/2}$. Substituting eqs. (5.118) and (5.119) into eqs. (5.99)-(5.101) gives

$$\begin{aligned} -\epsilon^2 Pe\left(\frac{\partial}{\partial\hat{z}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) + \epsilon^2\frac{\partial}{\partial\hat{z}}\left(d_I\frac{\partial}{\partial\hat{z}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) \\ +\epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_I\frac{\partial}{\partial\hat{r}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) \\ = K[(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1}) - a^{eq}b^{eq}] \end{aligned} \qquad (5.120)$$

$$\begin{aligned} -\epsilon^2 Pe\left(\frac{\partial}{\partial\hat{z}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) + \epsilon^2\frac{\partial}{\partial\hat{z}}\left(d_V\frac{\partial}{\partial\hat{z}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) \\ +\epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_V\frac{\partial}{\partial\hat{r}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) \\ = K[(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1}) - a^{eq}b^{eq}] \end{aligned} \qquad (5.121)$$

with boundary conditions

$$a_{0,0}(\hat{r},0) + \epsilon a_{1,0}(\hat{r},0) + \omega a_{0,1}(\hat{r},0) + O(\epsilon^2,\omega^2) = a^{eq}(\Theta=1) \qquad 0\le\hat{r}\le 1 \qquad (5.122)$$

$$b_{0,0}(\hat{r},0) + \epsilon b_{1,0}(\hat{r},0) + \omega b_{0,1}(\hat{r},0) + O(\epsilon^2,\omega^2) = b^{eq}(\Theta=1) \qquad 0\le\hat{r}\le 1 \qquad (5.123)$$

It is useful to introduce two additional concepts at this point. First, subtracting eq. (5.120) from eq. (5.121) gives

$$\begin{aligned} -\epsilon^2 Pe\left(\frac{\partial}{\partial\hat{z}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) + \epsilon^2\frac{\partial}{\partial\hat{z}}\left(d_I\frac{\partial}{\partial\hat{z}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) \\ +\epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_I\frac{\partial}{\partial\hat{r}}(a_{0,0}+\epsilon a_{1,0}+\omega a_{0,1})\right) = -\epsilon^2 Pe\left(\frac{\partial}{\partial\hat{z}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) \\ +\epsilon^2\frac{\partial}{\partial\hat{z}}\left(d_V\frac{\partial}{\partial\hat{z}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) + \epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_V\frac{\partial}{\partial\hat{r}}(b_{0,0}+\epsilon b_{1,0}+\omega b_{0,1})\right) \end{aligned} \qquad (5.124)$$

which eliminates the recombination reaction and is simply the scaled version of a flux balance at a point in the crystal. Also, the scaling of axial distance with L implies that for $z >> L$, all exponential functions, eqs. (5.105)-(5.107), decay to zero and eqs. (5.120)-(5.121) reduce to

$$\frac{\partial a}{\partial \hat{z}}(\hat{r}, \infty) = 0, \tag{5.125}$$

$$\frac{\partial b}{\partial \hat{z}}(\hat{r}, \infty) = 0. \tag{5.126}$$

These relations are used as far field boundary conditions for the leading order solutions in the Taylor expansions, eqs. (5.118) and (5.119). Computation of higher-order terms would require more careful consideration of the far field and the effect of the crystal radius.

At leading order ($\epsilon \to 0, \omega \to 0$) both eqs. (5.120) and (5.121) reduce to

$$K(a_{0,0}b_{0,0} - a^{eq}b^{eq}) = 0 \tag{5.127}$$

which simply relates the point defect concentrations by the law of mass action, and eq. (5.124) becomes

$$Pe\frac{\partial a_{0,0}}{\partial \hat{z}} - \frac{\partial}{\partial \hat{z}}\left(d_I \frac{\partial a_{0,0}}{\partial \hat{z}}\right) = Pe\frac{\partial b_{0,0}}{\partial \hat{z}} - \frac{\partial}{\partial \hat{z}}\left(d_V \frac{\partial b_{0,0}}{\partial \hat{z}}\right) \tag{5.128}$$

Most interestingly, at leading order radial transport by diffusion has scaled out of the analysis, so that point defect transport can be treated as a one-dimensional process with radial variation entering the analysis only through the temperature field, as represented by the function $g(\hat{r})$. Voronkov assumed one-dimensional expressions for the fluxes in his analysis (Voronkov, 1982).

Equations (5.127) and (5.128) define the leading order concentration field in the intermediate region defined by axial distances of $O(L)$. At leading order this solution also satisfies the far field conditions, eqs. (5.125)-(5.126).

Inner Region: $z = O(Da^{-1/2}L)$

The equations, (5.120),(5.121), and (5.124), are valid on the length scale L; however, this scaling does not capture the steep variation in point defect concentrations near the melt/crystal interface caused by the high Damkohler number, $\epsilon << 1$. From similar analyses for reaction-diffusion problems (Aris, 1975) we expect diffusion and reaction to both be important within this asymptotically small region. This dominant balance is accomplished by rescaling $\hat{z}$ as

$$\xi \equiv \hat{z}/\epsilon \tag{5.129}$$

so that $\xi = O(1)$ corresponds to $\hat{z} = O(Da^{-1/2})$.

The rescaled equations (5.99)-(5.101) are

$$-\epsilon Pe\frac{\partial \hat{a}}{\partial \xi} + \frac{\partial}{\partial \xi}\left(d_I\frac{\partial \hat{a}}{\partial \xi}\right) + \epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}d_I\frac{\partial \hat{a}}{\partial \hat{r}}\right) = K(\hat{a}\hat{b} - a^{eq}(\Theta)b^{eq}(\Theta)) \tag{5.130}$$

$$-\epsilon Pe\frac{\partial \hat{b}}{\partial \xi} + \frac{\partial}{\partial \xi}\left(d_V\frac{\partial \hat{b}}{\partial \xi}\right) + \epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial \hat{r}}\left(\hat{r}d_V\frac{\partial \hat{b}}{\partial \hat{r}}\right) = K(\hat{a}\hat{b} - a^{eq}(\Theta)b^{eq}(\Theta)) \tag{5.131}$$

and the boundary conditions are

$$\hat{a}(\hat{r},0) = 1, \qquad \hat{b}(\hat{r},0) = \frac{C_V^{eq}(1)}{C_I^{eq}(1)}, \qquad 0 \le \hat{r} \le 1, \tag{5.132}$$

where $\hat{a} \equiv \hat{a}(\hat{r},\xi)$ and $\hat{b} \equiv \hat{b}(\hat{r},\xi)$. Writing the solutions to eqs. (5.130)-(5.132) as Taylor series

$$\hat{a}(\hat{r},\xi) \approx \hat{a}_{0,0}(\hat{r},\xi) + \epsilon\hat{a}_{1,0}(\hat{r},\xi) + \omega\hat{a}_{0,1}(\hat{r},\xi) + O(\epsilon^2,\omega^2,\epsilon\omega) \tag{5.133}$$

$$\hat{b}(\hat{r},\xi) \approx \hat{b}_{0,0}(\hat{r},\xi) + \epsilon\hat{b}_{1,0}(\hat{r},\xi) + \omega\hat{b}_{0,1}(\hat{r},\xi) + O(\epsilon^2,\omega^2,\epsilon\omega) \tag{5.134}$$

and introducing these forms into eqs. (5.130)-(5.132) gives

$$-\epsilon Pe\frac{\partial}{\partial \xi}(\hat{a}_{0,0} + \epsilon\hat{a}_{1,0} + \omega\hat{a}_{0,1}) + \frac{\partial}{\partial \xi}\left(d_I\frac{\partial}{\partial \xi}(\hat{a}_{0,0} + \epsilon\hat{a}_{1,0} + \omega\hat{a}_{0,1})\right)$$

$$+\epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_I\frac{\partial}{\partial\hat{r}}(\hat{a}_{0,0}+\epsilon\hat{a}_{1,0}+\omega\hat{a}_{0,1})\right)$$
$$= K[(\hat{a}_{0,0}+\epsilon\hat{a}_{1,0}+\omega\hat{a}_{0,1})(\hat{b}_{0,0}+\epsilon\hat{b}_{1,0}+\omega\hat{b}_{0,1})-a^{eq}b^{eq}] \qquad (5.135)$$

$$-\epsilon Pe\frac{\partial}{\partial\xi}(\hat{b}_{0,0}+\epsilon\hat{b}_{1,0}+\omega\hat{b}_{0,1})+\frac{\partial}{\partial\xi}\left(d_V\frac{\partial}{\partial\xi}(\hat{b}_{0,0}+\epsilon\hat{b}_{1,0}+\omega\hat{b}_{0,1})\right)$$
$$+\epsilon^2\omega^2\frac{1}{\hat{r}}\frac{\partial}{\partial\hat{r}}\left(\hat{r}d_V\frac{\partial}{\partial\hat{r}}(\hat{b}_{0,0}+\epsilon\hat{b}_{1,0}+\omega\hat{b}_{0,1})\right)$$
$$= K[(\hat{a}_{0,0}+\epsilon\hat{a}_{1,0}+\omega\hat{a}_{0,1})(\hat{b}_{0,0}+\epsilon\hat{b}_{1,0}+\hat{b}_{0,1})-a^{eq}b^{eq}] \qquad (5.136)$$

with boundary conditions

$$\hat{a}_{0,0}(\hat{r},0)+\epsilon\hat{a}_{1,0}(\hat{r},0)+\omega\hat{a}_{0,1}(\hat{r},0)+O(\epsilon^2,\omega^2,\epsilon\omega)=a^{eq}(\Theta=1) \qquad (5.137)$$

$$\hat{b}_{0,0}(\hat{r},0)+\epsilon\hat{b}_{1,0}(\hat{r},0)+\omega\hat{b}_{0,1}(\hat{r},0)+O(\epsilon^2,\omega^2,\epsilon\omega)=b^{eq}(\Theta=1) \qquad (5.138)$$

Collecting terms of $O(1)$ gives

$$\frac{\partial}{\partial\xi}\left(d_I\frac{\partial}{\partial\xi}\hat{a}_{0,0}\right)=K(\hat{a}_{0,0}\hat{b}_{0,0}-a^{eq}b^{eq}) \qquad (5.139)$$

$$\frac{\partial}{\partial\xi}\left(d_V\frac{\partial}{\partial\xi}\hat{b}_{0,0}\right)=K(\hat{a}_{0,0}\hat{b}_{0,0}-a^{eq}b^{eq}) \qquad (5.140)$$

$$\hat{a}_{0,0}(\hat{r},0)=a^{eq}(1), \qquad \hat{b}_{0,0}(\hat{r},0)=b^{eq}(1) \qquad (5.141)$$

Equations (5.139)-(5.141) define the concentration fields in the inner region at leading order in the expansion.

The composite concentration field at leading order is determined by matching the solution in the inner region, valid on a length scale $O(Da^{-1/2}L)$, to the solution valid on the length scale $O(L)$. These matching conditions are written formally as

$$\lim_{\xi\to\infty}\hat{a}(\hat{r},\xi)=\lim_{z\to 0}a(\hat{r},\hat{z}) \qquad (5.142)$$

$$\lim_{\xi\to\infty}\hat{b}(\hat{r},\xi)=\lim_{z\to 0}b(\hat{r},\hat{z}) \qquad (5.143)$$

Concentration Field in the Inner Region

Equations (5.139)-(5.141) are solved in closed form. This is accomplished by expanding the temperature field, diffusivities, and equilibrium concentrations for small axial distances by writing the temperature field as

$$\Theta(\xi) = 1 - \gamma^{*-1} g(\hat{r})\hat{z} = 1 - \epsilon\gamma^{*-1} g(\hat{r})\xi, \tag{5.144}$$

where γ^* is the characteristic dimensionless activation energy used in eq. (5.115). The diffusivities take the form, for $i = I, V$,

$$D_i(\Theta) = D_i^0 \exp\left(-\frac{\gamma_{im}^*}{1 - \epsilon\gamma^{*-1} g(\hat{r})\xi}\right) \approx D_i^0 \exp\left(-\gamma_{im}^* - \epsilon g(\hat{r})\xi\frac{\gamma_{im}^*}{\gamma^*}\right), \tag{5.145}$$

where $\gamma_{im}^* \equiv H_i^m/kT_m$ and it is assumed that $\gamma_{im}^*/\gamma^* = O(1)$. Therefore, using eq. (5.106), the dimensionless diffusivities are

$$d_I(\Theta) = \exp\left(-\epsilon g(\hat{r})\xi\frac{\gamma_{Im}^*}{\gamma^*}\right) \approx 1 - \epsilon g(\hat{r})\xi\frac{\gamma_{Im}^*}{\gamma^*} \approx 1, \tag{5.146}$$

and

$$d_V(\Theta) = \frac{D_V(1)}{D_I(1)} \exp\left(-\epsilon g(\hat{r})\xi\frac{\gamma_{Vm}^*}{\gamma^*}\right) \approx \frac{D_V(1)}{D_I(1)}\left(1 - \epsilon g(\hat{r})\xi\frac{\gamma_{Vm}^*}{\gamma^*}\right) \approx \frac{D_V(1)}{D_I(1)}. \tag{5.147}$$

Expressions for the equilibrium concentrations $\{a^{eq}(\Theta), b^{eq}(\Theta)\}$ and the rate coefficient, $K(\Theta)$, are derived in an analagous manner:

$$a^{eq}(\Theta) = \exp\left(-\epsilon g(\hat{r})\xi\frac{\gamma_{If}^*}{\gamma^*}\right) \approx 1 - \epsilon g(\hat{r})\xi\frac{\gamma_{If}^*}{\gamma^*} \approx 1 \tag{5.148}$$

$$b^{eq}(\Theta) = \frac{C_V^{eq}(1)}{C_I^{eq}(1)} \exp\left(-\epsilon g(\hat{r})\xi\frac{\gamma_{Vf}^*}{\gamma^*}\right) \approx \frac{C_V^{eq}(1)}{C_I^{eq}(1)}\left(1 - \epsilon g(\hat{r})\xi\frac{\gamma_{Vf}^*}{\gamma^*}\right) \approx \frac{C_V^{eq}(1)}{C_I^{eq}(1)} \tag{5.149}$$

$$K(\Theta) = \frac{D_I(\Theta) + D_V(\Theta)}{D_I(1) + D_V(1)} \exp\left(-\epsilon g(\hat{r})\xi\frac{\gamma_{IV}^*}{\gamma^*}\right) \approx \frac{D_I(\Theta) + D_V(\Theta)}{D_I(1) + D_V(1)}\left(1 - \epsilon g(\hat{r})\xi\frac{\gamma_{IV}^*}{\gamma^*}\right) \approx 1 \tag{5.150}$$

where $\gamma_{if}^* \equiv H_i^f/kT_m$ and $\gamma_{IV}^* \equiv \Delta G_{IV}/kT_m$ are the dimensionless formation energy and free energy barrier to recombination, respectively. To leading order, each expression (5.146)-(5.150) is a constant evaluated at the melting temperature.

The solution to eqs. (5.139)-(5.141) and (5.146)-(5.150) can be written down by inspection as

$$\hat{a}_{0,0}(\hat{r},\xi) = 1, \tag{5.151}$$

$$\hat{b}_{0,0}(\hat{r},\xi) = \frac{C_V^{eq}(1)}{C_I^{eq}(1)}. \tag{5.152}$$

so to a first approximation the concentrations remain at their equilibrium values in the inner region.

Concentration Field in the Intermediate Region

The solutions to equations (5.127) and (5.128) are developed by postulating Arrhenius behavior for $C_I(\hat{r},\hat{z})$ and $C_V(\hat{r},\hat{z})$ for $\hat{z} = O(1)$, so that

$$a(\hat{r},\hat{z}) = A\exp\left(-\frac{\gamma_{If}^{*\prime}}{\Theta}\right), \tag{5.153}$$

$$b(\hat{r},\hat{z}) = B\exp\left(-\frac{\gamma_{Vf}^{*\prime}}{\Theta}\right), \tag{5.154}$$

where $\gamma_{if}^{*\prime} \equiv \gamma_{if}'/kT_m$ is a dimensionless *effective* formation energy for the *actual* point defect concentrations in an axial region where $\hat{z} = O(1)$. Using eq. (5.144) and truncating at leading order gives

$$a(\hat{r},\hat{z}) \sim A\exp(-\gamma_{If}^{*\prime})\exp\left(-\frac{\gamma_{If}^{*\prime}}{\gamma^*}g(\hat{r})\hat{z}\right), \tag{5.155}$$

$$b(\hat{r},\hat{z}) \sim B\exp(-\gamma_{Vf}^{*\prime})\exp\left(-\frac{\gamma_{Vf}^{*\prime}}{\gamma^*}g(\hat{r})\hat{z}\right). \tag{5.156}$$

Defining

$$A' \equiv A\exp(-\gamma_{If}^{*\prime}), \qquad B' \equiv B\exp(-\gamma_{Vf}^{*\prime}), \tag{5.157}$$

and

$$\hat{\alpha}(\hat{r}) \equiv \frac{\gamma_{If}^{*\prime}}{\gamma^*} g(\hat{r}), \qquad \hat{\beta}(\hat{r}) \equiv \frac{\gamma_{Vf}^{*\prime}}{\gamma^*} g(\hat{r}), \tag{5.158}$$

gives the point defect concentrations as

$$a(\hat{r}, \hat{z}) = A' \exp\left[-\hat{\alpha}(\hat{r})\hat{z}\right] = A' \exp\left[-\alpha(\hat{r})z\right], \tag{5.159}$$

$$b(\hat{r}, \hat{z}) = B' \exp\left[-\hat{\beta}(\hat{r})\hat{z}\right] = B' \exp\left[-\beta(\hat{r})z\right], \tag{5.160}$$

where the second form is in terms of the dimensional length scale, z. The equilibrium concentrations and diffusivities are expressed in analogous manner as

$$a^{eq}(\Theta) = \exp[-\delta(\hat{r})z] \qquad b^{eq} = \frac{C_V^{eq}(1)}{C_I^{eq}(1)} \exp[-\phi(\hat{r})z] \tag{5.161}$$

$$d_I(\Theta) = \exp[-\mu(\hat{r})z] \qquad d_V(\Theta) = \frac{D_V(1)}{D_I(1)} \exp[-\nu(\hat{r})z] \tag{5.162}$$

where $\delta(\hat{r}) \equiv \gamma_{If}^* g(\hat{r})/L\gamma^*$, $\phi(\hat{r}) \equiv \gamma_{Vf}^* g(\hat{r})/L\gamma^*$, $\mu(\hat{r}) \equiv \gamma_{Im}^* g(\hat{r})/L\gamma^*$, and $\nu(\hat{r}) \equiv \gamma_{Vm}^* g(\hat{r})/L\gamma^*$. Substituting eqs. (5.159)-(5.162) into eqs. (5.127) and (5.128) and simplifying gives the following two relationships between the four unknowns, $\{\alpha(\hat{r}), \beta(\hat{r}), A', B'\}$:

$$\begin{aligned} \alpha(\hat{r}) \frac{V}{D_I(1)} \exp[-\alpha(\hat{r})z] + \alpha(\hat{r})^2 d_I \exp[-\alpha(\hat{r})z] - \alpha(\hat{r})\mu(\hat{r}) d_I \exp[-\alpha(\hat{r})z] \\ = \beta(\hat{r}) \frac{V}{D_I(1)} \frac{B'}{A'} \exp[-\beta(\hat{r})z] + \beta(\hat{r})^2 d_V \frac{B'}{A'} \exp[-\beta(\hat{r})z] \\ -\beta(\hat{r})\nu(\hat{r}) d_V \frac{B'}{A'} \exp[-\beta(\hat{r})z] \end{aligned} \tag{5.163}$$

and

$$\alpha(\hat{r}) + \beta(\hat{r}) = \delta(\hat{r}) + \phi(\hat{r}) \tag{5.164}$$

The other two unknowns can be determined by employing the asymptotic matching conditions, eqs. (5.142) and (5.143) which give

$$A' = 1, \qquad B' = \frac{C_V^{eq}(1)}{C_I^{eq}(1)} \tag{5.165}$$

and

$$0 = \left[D_I(1) - \frac{C_V^{eq}(1)}{C_I^{eq}(1)} D_V(1)\right] \alpha^2$$
$$+ \left[V\left(1 + \frac{C_V^{eq}(1)}{C_I^{eq}(1)}\right) + \mu D_I(1) + 2\frac{C_V^{eq}(1)}{C_I^{eq}(1)}(\delta + \phi) D_V(1)\right] \alpha$$
$$- \left[\frac{C_V^{eq}(1)}{C_I^{eq}(1)}(\delta + \phi)\left(V + D_V(1)\frac{D_V(1)}{D_I(1)}(\nu + \delta + \phi)\right)\right] \qquad (5.166)$$

In combination eqs. (5.164) and (5.166) completely determine $a(\hat{r}, \hat{z})$ and $b(\hat{r}, \hat{z})$ in the inner layer, $z \sim O(L)$.

5.6.3 Verification of the Leading-Order Solution by Comparison to Finite Element Analysis

The accuracy of the leading order solution is tested by direct comparison to the point defect simulations. First, the point defect simulation code was used to compute concentration fields for interstitials and vacancies for a given pull rate and growth geometry, in particular, the parameters obtained in Section 5.4.1 and given by eqs. (5.89)-(5.93) were used in a simulation of the HS1/4" crystal at a pull rate of 1.0 mm/min. The concentrations and diffusivities were numerically fitted to exponential functions of the form $P\exp(-\gamma z)$ in the region near the melt/solid interface at three radial positions; these results are shown in Table 1. In all cases, the correlation coefficient for the fits was greater than 0.9998, indicating that the solution fields were well approximated by this exponential form in this region of the crystal. The constants in the asymptotic form, eq. (5.164), of the law of mass action computed from these fits are compared in Table 5.1. The agreement is very good; this is another indication that the asymptotic analysis is justified. The values for the effective activation energies computed from the simulation results are compared directly to those obtained from eq. (5.166). These results are shown in Table 5.2. Once again, the leading order solution predicts the simulation results very well for $z \sim O(L)$ indicating that the picture of dual boundary layers is essentially correct.

5.6.4 Expressions for $(V/G)_{crit}$

The most significant result of Voronkov's analysis was the conclusion that the ratio (V/G) can be used to determine whether a vacancy-rich or interstitial-rich regime exists during crystal growth, in qualitative agreement with the experimental observations of Dornberger et al. (1996). Their expression for the critical value of (V/G), which separates these regions, was derived in terms of the point defect thermophysical properties at the melting temperature by requiring the point defect fluxes at the melt/solid interface to be equal.

Voronkov estimates a value for $(V/G)_{crit}$ by considering the pull rates at which the various microdefects appear and disappear in float-zone crystals. This estimate is used in conjunction with literature estimates for the silicon self-diffusion coefficient to compute values for the point defect thermophysical properties at the melting temperature. In the present work, we are essentially carrying out the reverse analysis, where a more rigorous form of Voronkov's $(V/G)_{crit}$ is derived based on the asymptotic approximations for $C_I(r,z)$ and $C_V(r,z)$ developed in Section 5.6.2. The point defect thermophysical properties given in eqs. (5.89)-(5.93) are used to compute an independent estimate for $(V/G)_{crit}$ which is compared to the experimental value of Dornberger et al. (1996)

It was shown in section 5.6.2 that to leading order the difference of the point defect balance equations led to a relationship between the point defect fluxes where the radial components of the differential equations were unimportant, as in eq. (5.128). If we define the generalized point defect fluxes as

$$J^I_{0,0}(\hat{r},\hat{z}) \equiv (Pe)a_{0,0}(\hat{r},\hat{z}) - d_I\frac{\partial a_{0,0}(\hat{r},\hat{z})}{\partial \hat{z}}, \tag{5.167}$$

$$J^V_{0,0}(\hat{r},\hat{z}) \equiv (Pe)b_{0,0}(\hat{r},\hat{z}) - d_V\frac{\partial b_{0,0}(\hat{r},\hat{z})}{\partial \hat{z}}, \tag{5.168}$$

we see that to leading order, eq. (5.128) gives

$$J^I_{0,0}(\hat{r},\hat{z}) - J^V_{0,0}(\hat{r},\hat{z}) = H(r), \qquad 0 \le \hat{z} \le \infty \tag{5.169}$$

where $H(r)$ is an integration constant for a given radial location. As $\hat{z} \rightarrow \infty$, the point defect fluxes reduce to

$$J^I_{0,0} \equiv (Pe)a_{0,0}, \qquad J^V_{0,0} \equiv (Pe)b_{0,0} \tag{5.170}$$

so that

$$Pe(a_{0,0} - b_{0,0}) = H(r), \tag{5.171}$$

or

$$\Delta = H'(r), \tag{5.172}$$

where $\Delta(\hat{r}, \hat{z}) \equiv C_I(\hat{r}, \hat{z}) - C_V(\hat{r}, \hat{z})$ and $H'(r) \equiv H(r)C^{eq}_I(1)/Pe$. Note that Δ varies axially until the point defect fluxes approach the forms given in eqs. (5.170). In other words, if $\Delta = 0$ in the cold, upper section of the crystal, $H'(r) = H(r) = 0$ and the point defect fluxes are equal for all $\hat{z}$ at a given radial position. Thus, according to our model for the location of the OSF-ring, the radial location of the ring corresponds to the radial position where the point defect fluxes into the crystal across the melt/solid interface are equal.

Thus, for $J^I_{0,0} = J^V_{0,0}$, substituting the leading order approximations to $a(\hat{r}, \hat{z})$ and $b(\hat{r}, \hat{z})$, given by eqs. (5.165) and (5.166), leads to

$$Pe(a_{0,0} - b_{0,0}) = d_V b_{0,0}\hat{\beta} - d_I a_{0,0}\hat{\alpha}. \tag{5.173}$$

Using eqs. (5.158) to write

$$\hat{\alpha} = \frac{\gamma^{*\prime}_{If} G^* g L}{T_m}, \tag{5.174}$$

$$\beta = \frac{\gamma^{*\prime}_{Vf} G^* g L}{T_m}, \tag{5.175}$$

eq. (5.173) can be expressed as

$$\left[\frac{V}{G}\right]_{crit} = \frac{D_I(1)C^{eq}_I(1)\gamma^{*\prime}_{If} - D_V(1)C^{eq}_V(1)\gamma^{*\prime}_{Vf}}{T_m[C^{eq}_V(1) - C^{eq}_I(1)]}. \tag{5.176}$$

where $\gamma^{*\prime}_{If}$ and $\gamma^{*\prime}_{Vf}$ are the effective dimensionless enthalpies of formation for the actual

concentration profiles, as are computed from the leading order concentration profiles developed in Section 5.6.2 for a specified thermal field set by $g(\hat{r})$. This expression is very similar to eq.(5.7), derived by Voronkov (1982) with a couple of fundamental differences. In particular, the activation energies in our expression, eq. (5.176), are those of the *actual* profiles not the equilibrium ones. Furthermore, Voronkov's analysis assumed that they were equal and simply multiplied each point defect flux by the same activation energy which, from the present analysis, is not generally true. However, Voronkov's expression does capture the essential dynamics between the interstitial and vacancy concentrations; this is a very notable result of his simplified analysis.

The point defect properties given by equations (5.89)-(5.93) give an estimate for $(V/G)_{crit}$ as

$$\frac{V}{G}_{crit} = 1.38 \times 10^{-3} \ cm^2/minK, \tag{5.177}$$

which is in excellent agreement with the experimentally determined value of $1.34 \times 10^{-3} \ cm^2/minK$. The fact that the point defect properties represented by eqs. (5.89)-(5.93) reproduce well both the OSF-ring data and the critical value for (V/G) is a good indication that the mechanistic picture of point defect dynamics as the dominant effect in OSF-ring dynamics is correct.

5.7 Summary and Discussion

The OSF-ring provides an uniquely explicit system with which a point defect dynamic model can be validated against real experimental data. The results of the simulations presented in this Chapter show that an intrinsic point defect based picture is able to model OSF-ring dynamics across a wide range of crystal growth conditions. The crucial test of the models validity lies in the ability of a *single* set of point defect parameters to reproduce OSF-ring behavior for several, widely varying configurations, implying that the model is capturing the essential physics that govern OSF-ring dynamics. It also gives the first indication that such a model is capable of quantitatively predictive calculations which is of great current technological importance.

The (V/G) parameter provides an empirical approach to estimating the ring loca-

tion. The prediction of a critical value in excellent agreement with the experimental value reinforces the validity of the fundamental assumption that the dynamics of the OSF-ring are completely determined by point defects and the structure of the thermal field near the melt/solid interface. This is an extremely important result, since the point defect properties derived through the fitting of the model to experimental data can therefore be assumed to be self-consistent. Current and future numerical simulations investigating directly the size and concentration of microdefects themselves will necessarily require such self-consistency if any quantitative progress is to be made.

R	$(\alpha + \beta)$	$(\delta + \phi)$
0.1 R	2.25	2.24
0.4 R	2.47	2.46
0.8 R	3.02	3.01

Table 5.1: Exponent Values computed from simulation by numerical fitting

R	α-simulation	α-theory	β-simulation	β-theory
0.1 R	1.11	1.10	1.13	1.15
0.4 R	1.20	1.20	1.25	1.26
0.8 R	1.40	1.44	1.60	1.58

Table 5.2: Solution exponents derived from theory compared to simulation values

Chapter 6

Continuum Scale Simulation of Point Defect Aggregation

6.1 Introduction

The model presented in Chapter 5 is based on the hypothesis that it is possible to decouple the *dynamics* of the Oxidation-Induced Stacking-Fault Ring (OSF-ring) from the details of its *composition*. It was possible to predict one aspect of microdefect behavior without explicitly computing microdefect sizes and distributions. As described in Chapter 2, todays technological challenge, however, lies not in the prediction of the location of the OSF-ring but in the minimization of the microdefect density in the region inside the OSF-ring, because all commercial CZ crystals are grown with the OSF-ring at the periphery of the crystal (Dornberger, 1997). Crystals are grown this way because even though D-defects detrimentally affect device performance by lowering Gate Oxide integrity (GOI), large dislocation loops in the interstitial-rich zone cause even more damage including complete device failure. Furthermore, in order to grow crystals exclusively in the interstitial dominated regime, crystals would have to be pulled very slowly which is economically undesirable. The goal in modeling of CZ microdefect growth is therefore currently directed at minimizing the density and size of D-defects while keeping the OSF-ring at the periphery of the crystal. The possibility of growing commercial crystals in the interstitial-dominated regime is

not entirely infeasible, however, and the investigation of interstitial aggregates is still useful especially if the size of these defects can be reduced.

A significant obstacle to modeling microdefect formation quantitatively is the fact that it is very difficult to elucidate experimentally the exact composition of most microdefect species. Only recently it was found that the D-defects that appear inside the OSF-ring were composed mainly of vacancies instead of oxygen (Itsumi *et al.*, 1995). Most earlier aggregation models attempting to estimate quantitatively the size and concentration of these microdefects were therefore invalidated (Esfandyari *et al.*, 1996). Even now, it is not clear what the role of oxygen is in the formation of D-defects in CZ silicon. Some SiO_2 is known to exist within the voids as a layer coating the surface of the defect, but whether or not oxygen influences the nucleation mechanism is unknown; a detailed discussion of the observed features of voids in CZ silicon was presented in Chapter 2. The best that any model of microdefect formation can do is to first postulate a mechanism for microdefect formation, and then test whether the predictions of the model are in accord with the experimental observations for a wide range of crystal growth operating conditions.

The concept that vacancies are the main ingredient of D-defects is very consistent with the intrinsic point defect picture described in Chapter 5, and the current Chapter will focus on the extension of the point defect model to include nucleation and aggregation mechanisms that allow for the explicit modeling of intrinsic point defect aggregates. Of particular interest is the response of the vacancy aggregate size distribution to changes in the macroscopic operating conditions. These include the pull rate and the characteristics of the thermal field within the crystal. At this level of detail, we are in a position to investigate directly the effects of the crystal growth system geometry, such as heat shield configuration, on microdefect formation.

Also of interest are the mechanisms that lead to the formation of the OSF-ring itself after wafer wet-oxidation. The results of the point defect model of Chapter 5 show that the OSF-ring likely separates regions of self-interstitial and vacancy domination but do not provide an explanation for the source of the large number of interstitial-type stacking faults in the OSF-ring region itself. Oxygen precipitation

is known to play an important part in forming these stacking faults because OSF-rings do not form in FZ crystals. However, there is no reason for oxygen to precipitate preferentially at a particular radial position in a crystal without some external driving force. A possible explanation for the origin of the OSF-ring is the distribution of point defect aggregates in this region. A large density of small vacancy clusters, for example, could provide the free volume required for such precipitation to proceed rapidly, even at high temperature; homogeneous oxygen precipitation is otherwise known to occur at relatively low temperatures as compared to interstitial and vacancy condensation. Such a picture would predict that the pattern of oxygen precipitation is actually controlled by intrinsic point defect dynamics. These issues are the focus of the model described in this Chapter.

6.2 Nucleation and Aggregation Model

Expressions for the rates of point defect aggregation constitute the core of the nucleation model. These are derived here by combining standard diffusion-limited reaction theory (Waite, 1957) as in the case of point defect recombination (see Chapter 3), with homogeneous nucleation theory (Katz and Wiedersich, 1971). Point defect clusters are assumed to be immobile and to grow and shrink by the addition or loss of single point defects, which is a good approximation if clusters are dilute and coalescence is unlikely. For point defects in silicon, which is exist in concentrations no higher than 1 part-per-million, this is certainly the case.

The aggregation process can be represented as a series of dynamical equilibria, so that

$$\cdots \overset{k_{n-1}}{\longleftrightarrow} (n-1)X + X \overset{k_n}{\longleftrightarrow} nX + X \overset{k_{n+1}}{\longleftrightarrow} (n+1)X + X \longleftrightarrow \cdots, \tag{6.1}$$

where X represents a single point defect, and nX a cluster of n point defects. Expressions for the actual aggregation steps are obtained by considering both the forward and backward aggregation rates for each cluster size and assuming that the nucle-

ation proceeds homogeneously. In eq. (6.1), the reaction represented by k_n can be decomposed into two steps; the growth of clusters of size n to size $(n+1)$ (denoted by k_{n+1}^f), and the dissolution of clusters of size $(n+1)$ to clusters of size n (denoted by k_n^b). These reactions can be expressed as bimolecular and unimolecular processes respectively;

$$k_n^f = \alpha_n C_{(n-1)X} C_X, \tag{6.2}$$

and

$$k_n^b = \beta_n C_{nX}. \tag{6.3}$$

In order to obtain expressions for α_n and β_n we need to consider the energy path that the system needs to pass through as a cluster grows. The initial configuration consists of two defects; a cluster of size n and a single point defect. As the point defects combine, they must overcome a free energy barrier which is denoted as ΔG_B. This barrier is assumed to be an adjustable parameter in our model as there is no data available on its magnitude. The final configuration of the system is a single cluster of size $n+1$. In general, the rate of a thermally activated process is given by jump rate theory as

$$K = A \exp\left(-\frac{\Delta G_{tot}}{kT}\right) \tag{6.4}$$

where ΔG_{tot} is the total free energy required for the process to occur, which can be either positive or negative, and A is the attempt frequency that characterizes the process.

6.2.1 Thermodynamics of Open and Closed Systems

Before a complete free energy description of the nucleation/aggregation process can be developed it is necessary to consider the thermodynamic environment in which aggregation proceeds. Section 3.6.6 of Chapter 3 outlined the development of a rate expression for interstitial-vacancy recombination based only on the local concentrations of point defects. The fact that the depletion of point defects results in a lowering of the *configurational* entropy of the crystal was not considered. This assumption is

now discussed in detail because it has significant implications for aggregation.

The relevant thermodynamic energy function for describing a particular system depends on which state variables are held constant, according to statistical mechanics (Allen and Tildesley, 1992; McQuarrie, 1976). The Gibbs free energy is the appropriate function for a system where the number of particles, pressure, and temperature are held constant. From a material viewpoint, this represents a closed system and, in this case, no interstitials or vacancies are allowed to enter or leave the system across its boundaries. In such a system, the equilibrium concentration of each point defect type cannot be defined as an independent quantity; it must be defined in terms of the other variables because the only mechanism for altering the number of either involves both species (recombination). In systems where aggregation also occurs, point defect equilibrium concentrations also are functions of the concentrations of all aggregate sizes. Thus, when computing the equilibrium concentration of interstitials and vacancies for a given temperature and pressure under these conditions, the total free energy is minimized under a constraint. Arising naturally from this constraint is the configurational entropy which is related to the degree of disorder in the system.

In an open system which allows for the gain or loss of point defects across its boundaries, the chemical potential rather than the number of particles is constant. Each of the point defect species can achieve equilibrium independently and the configurational entropy no longer imposes a constraint on the minimization of the total free energy. However, in this case, the Gibbs free energy is no longer the natural thermodynamic function that describes the system. Instead it is $\Omega \equiv H - \mu T$.

The macroscopic crystal is essentially an open system because its surfaces are assumed to acts as sinks and sources for point defects. However, on a microscopic scale characterized by point defect diffusion lengths during growth, the local environment can be different. Near boundaries, point defects can easily be added or removed from the system. In the bulk of the crystal, however, point defects essentially see an infinite medium which possesses the same thermodynamic characteristics of a closed system; i.e. the boundaries are too far away to make a difference during the growth process. *The analysis of Chapter 5 showed that point defect recombination proceeds*

in the region immediately next to the melt/solid interface and so recombination can be assumed to proceed in an open system.

The situation for aggregation is different as it proceeds mainly at lower temperatures away from the melt/solid interface. Thus, the consumption of point defects to form aggregates reduces the configurational entropy. The free energy of a closed system of aggregates is given by

$$G = G_0 + \sum_n z_n G^f(n) - kT \ln \Omega, \tag{6.5}$$

where G_0 is a reference free energy, $G^f(n)$ is the non-configurational free energy of formation of an aggregate, z_n is the number of aggregates of size n and Ω is the number of ways clusters can be distributed across lattice sites. This number is a function of the number of lattice sites, N, and is given by (Katz and Wiedersich, 1971)

$$\Omega \equiv \frac{(n)^{z_n}(N/n)!}{(N/n - z_n)! z_n!}. \tag{6.6}$$

The Stirling approximation gives the following expression for the chemical potential of clusters of size n:

$$\mu_n \equiv \frac{\partial F}{\partial z_n} = G^f(n) - kT \ln \left[\frac{(N - nz_n)}{z_n} \right], \tag{6.7}$$

where for dilute cluster concentrations, $nz_n << N$. The free energy change involved in the growth of a cluster from size n to size $n+1$ is then given by

$$\Delta G_{n \to n+1} = G^f(n+1) - G^f(n) - G^f(1) + kTln\frac{z_{n+1}}{z_n} - kTln\frac{z_1}{N}. \tag{6.8}$$

Equation (6.8) can be simplified by choosing the unconstrained equilibrium concentration of monomers to be the reference state of the system, and defining the supersaturation of monomers with respect to that state as $S \equiv C_1/C_1^{eq}$, where $C \equiv z_1/N$. These definitions lead to the following expression for the free energy change associated

with the incorporation of a monomer into a cluster:

$$\Delta G_{n\to n+1} \equiv G^f(n+1) - G^f(n) - kTln\frac{C_1}{C_1^{eq}}. \tag{6.9}$$

In order to derive a rate expression for the growth of a cluster we use jump rate theory (Vineyard, 1957) as in the case of point defect recombination. Cluster growth is assumed to proceed when a monomer located at the surface of an aggregate "jumps" into the aggregate. This occurs whenever the diffusing monomer possesses at least $\Delta G_{n\to n+1}$. The number of incorporation attempts is related to the size of the cluster, the mobility of the monomer (clusters are assumed to be stationary), and the distance across which the jump is to be made. The latter is assumed to be on the order of the the lattice spacing, or 0.235 nm. The rate of cluster surface reaction is then given by (Dornberger, 1997)

$$k_r(n,\mathbf{r}) = \frac{4\pi r_v^2 D_1}{\delta}\exp\left(-\frac{\Delta G_{n\to n+1}}{kT}\right). \tag{6.10}$$

The non-configurational part of the free energy of formation for a cluster is generally a function of bulk and surface energies:

$$G^f(n) = 4\pi r_c^2\sigma + \frac{4}{3}\pi r_c^3 \Delta G_b \tag{6.11}$$

where r_c is the radius of a spherical cluster, σ is the surface free energy per unit area, and ΔG_b is the bulk energy per unit volume. In the case of vacancy clusters the bulk contribution is identically zero because void space does contribute to the energy of the system.

The consumption of monomers in the region around a void can lead to a depletion layer around the void if the aggregate growth rate is fast compared to the diffusivity of the monomers; i.e. the nucleation Damköhler number is high. Thus, the actual cluster growth rate would be characterized by a surface concentration of monomers, not the bulk value. It is possible to estimate the concentration of monomers at a cluster surface under diffusion limitation by considering the steady-state spherical

diffusion problem in the vicinity of a cluster (Brabec *et al.*, 1989),

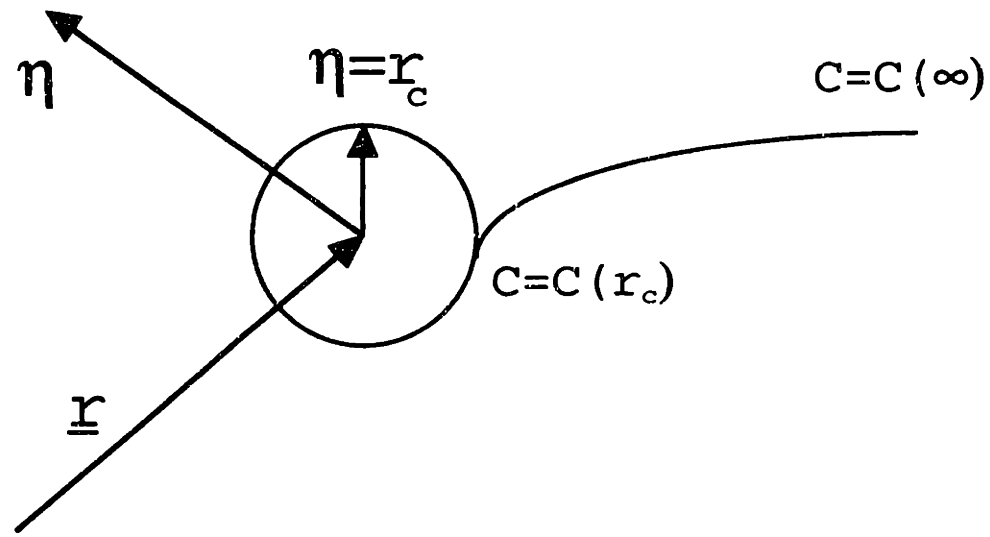

Figure 6-1: Spherical coordinate system for describing monomer depletion in the vicinity of a microdefect.

$$C_1(\infty) = C_1(bulk), \tag{6.13}$$

and at the surface of the growing cluster,

$$4\pi r_c^2 D_1 \left(\frac{\partial c}{\partial \eta}\right)_{\eta=r_c} = k_r[C_1(r_c) - C_1^{eq}(r_c)], \tag{6.14}$$

where $C_1(r_c)$ and $C_1^{eq}(r_c)$ are the actual and equilibrium concentrations of monomer at the surface of the cluster. The solution of eq. (6.12) with the above boundary conditions gives

$$C_1(r_c) = \frac{k_r C_1^{eq}(r_c) + k_d C_1(bulk)}{k_r + k_d}, \tag{6.15}$$

where k_d represents the diffusion rate of monomers to the cluster and is given by

$$k_d = 4\pi r_c D_1. \tag{6.16}$$

The equilibrium concentration of monomers at the surface of a cluster of size x is obtained by minimizing the free energy of formation if a cluster with respect to cluster size (Lifshitz and Slyozov, 1961). It is related to the bulk equilibrium concentration as

$$C_1^{eq}(r_c) = C_1^{eq}(\infty) \exp\left(\frac{\partial G^f(n)/\partial n}{kT}\right). \tag{6.17}$$

Cluster growth and dissolution rates can now be written in terms of the surface reaction rates and monomer concentrations. The growth rate is

$$g(n) = k_r(n, \mathbf{r}) C_1(r_c), \tag{6.18}$$

and the dissolution rate is

$$d(n) = k_r(n, \mathbf{r}) C_1^{eq}(r_c). \tag{6.19}$$

6.3 Point Defect and Cluster Balance Equations

The primary challenge in the direct numerical simulation of microdefect nucleation and growth lies in the large span of cluster sizes that exist within a growing crystal. In order to accommodate the very large distribution of clusters sizes present in real crystals (up to ten orders-of-magnitude), it is necessary to employ a hybrid approach, where point defects and small clusters are described by discrete rate equations (DREs) and larger clusters are described with a single Fokker-Planck equation (FPE). Such an approach has been employed by Esfandyari et al. (1996) . The transformation of a series of discrete rate equations into a single Fokker-Planck equation is common in the numerical simulation of aggregation phenomena and will be described in Section 6.3.2. The discrete and continuous representation of cluster distributions are linked together by a matching condition that equates the concentrations and fluxes at the matching

point. The FPE and DREs are also implicitly coupled to each other through the nucleation and aggregation rates.

6.3.1 Discrete Representation

Small clusters of up to about ten or twenty point defects must be treated as individual discrete clusters because the Fokker-Planck approximation is not accurate for small clusters. Esfandyari et al. (1996) have shown that ten rate equations are sufficient to ensure the validity of the Fokker-Planck approximation at the matching point. Equations (5.8) and (5.9) in Chapter 5 now contain an additional term and are written as

$$\nabla \cdot \left(-D_I(T)\nabla C_I + \frac{D_I(T)C_I Q_I^{*\prime}}{kT^2}\nabla T \right) + V\frac{\partial C_I}{\partial z} + k_{IV}(T)\left(C_I C_V - C_I^{eq}(T)C_V^{eq}(T)\right) + V\frac{\partial}{\partial z}\int_2^\infty n f_I(n,\mathbf{r})dn = 0, \qquad (6.20)$$

$$\nabla \cdot \left(-D_V(T)\nabla C_V + \frac{D_V(T)C_V Q_V^{*\prime}}{kT^2}\nabla T \right) + V\frac{\partial C_V}{\partial z} + k_{IV}(T)\left(C_I C_V - C_I^{eq}(T)C_V^{eq}(T)\right) + V\frac{\partial}{\partial z}\int_2^\infty n f_V(n,\mathbf{r})dn = 0, \qquad (6.21)$$

where the last term in each balance equation represents the incorporation of point defects into clusters of all possible sizes. Here, $f_X(n,\mathbf{r})$ is the concentration of clusters of point defect X, and size n at a given position, $\mathbf{r}$, within the crystal.

Balance equations for clusters are developed in a similar manner and are written for both interstitial and vacancy types as

$$\nabla \cdot (-D_A(T)\nabla C_{n,X} + VC_{n,X}) = J_{n,X} - J_{n+1,X}, \qquad (6.22)$$

where $J_{n,X}$ is the nucleation flux of species X, and is given by

$$J_{n,X} = g(n-1)C_{n-1,X} - d(n)C_{n,X}. \qquad (6.23)$$

The first term in eq. (6.22) represents an artificial diffusion term added to preserve the ellipticity of the equation. Even though clusters are assumed to be immobile, the inclusion of this term is necessary to stabilize the Galerkin finite element method. The magnitude of the artificial diffusion coefficient, D_A, is chosen so that the numerical Peclet number, $Pe_N \equiv hV/D_A$, is of order unity, where h is the mesh size in the direction of the convective term, and V is the crystal pull rate. The resulting solution is accurate to $O(h)$. Without artificial diffusion, the resulting profiles can exhibit very steep gradients which would require very fine computational meshes to resolve; This would be tantamount to lowering Pe_N by reducing h to extremely small values.

6.3.2 The Continuous Representation - Fokker-Planck Equation

A continuous representation for the series of discrete rate equations can be developed by assuming that the cluster size distribution can be expressed in terms of a continuous function, rather than a discrete variable. The continuous nucleation flux, I_n, is given as

$$I_n(\mathbf{r}) \equiv g(n-1)f(n-1,\mathbf{r}) - d(n)f(n,\mathbf{r}) \tag{6.24}$$

where $f(n,\mathbf{r})$ is a continuous function of size and space. The assumption that cluster size is a continuously varying function is good in the limit of large cluster size and if cluster growth and dissolution proceeds via the addition and loss of single point defects, so that the *changes* in cluster size are small compared to the cluster size.

Equation (6.22) can be rewritten in terms of transition rates, (W), and probability densities, f, to give the following expression:

$$V\frac{\partial f(n,\mathbf{r})}{\partial z} = \int W(n',n)f(n',\mathbf{r})dn' - \int W(n,n')f(n,\mathbf{r})dn'. \tag{6.25}$$

Here, $W(n',n)$ is the transition rate from state n' to state n and $f(n',\mathbf{r})$ is the probability that the system exists in state n'. The Kramers-Moyal expansion (Risken, 1984) is now performed by introducing a new variable, the jump size, $\xi \equiv n - n'$, and

a new transition rate, T, so that

$$T(n',\xi) \equiv W(n',n) \tag{6.26}$$

Rewriting eq. (6.25) in terms of T, we obtain

$$V\frac{\partial f(n,\mathbf{r})}{\partial z} = \int T(n-\xi,\xi)f(n-\xi,\mathbf{r})d\xi - f(n,\mathbf{r})\int T(n,-\xi)d\xi. \tag{6.27}$$

Expanding the first integrand in eq. (6.27) in Taylor series gives

$$T(n-\xi)f(n-\xi,\mathbf{r}) = T(n,\xi)f(n,\mathbf{r}) - \frac{(-\xi)^n}{n'!}\frac{\partial^{n'}}{\partial n^{n'}}T(n,\xi)f(n,\mathbf{r}), \tag{6.28}$$

and substituting into eq. (6.27) after truncation to second order leads to

$$\begin{aligned} V\frac{\partial f(n,\mathbf{r})}{\partial z} = &\int T(n,\xi)f(n,\mathbf{r})d\xi - \int (d\xi)\xi\frac{\partial}{\partial n}[T(n,\xi)f(n,\mathbf{r})]+ \\ &\frac{1}{2}\int (d\xi)\xi^2\frac{\partial^2}{\partial n^2}[T(n,\xi)f(n,\mathbf{r})] - f(n,\mathbf{r})\int T(n,-\xi)d\xi. \end{aligned} \tag{6.29}$$

We define

$$A'(n) \equiv \int (d\xi)\xi T(n,\xi), \qquad B'(n) \equiv \int (d\xi)\xi^2 T(n,\xi). \tag{6.30}$$

so that the resulting Fokker-Planck equation can be written as

$$V\frac{\partial f(n,\mathbf{r})}{\partial z} = -\frac{\partial}{\partial n}[A'(n)f(n,\mathbf{r})] + \frac{1}{2}\frac{\partial^2}{\partial n^2}[B'(n)f(n,\mathbf{r})]. \tag{6.31}$$

The FPE is therefore similar to a diffusion/drift equation with $A'(n)$ the drift coefficient and $B'(n)$ a diffusion coefficient. These coefficients are functions of the forward and backward nucleation rates at each cluster size. The assumption that only single point defects are mobile implies that $\xi = \pm 1$. The integrals in eqs. (6.30) are then trivial to evaluate:

$$A'(n) = g(n) - d(n), \qquad B'(n) = g(n) + d(n). \tag{6.32}$$

A slightly different (and more common) representation of the FPE will be used from here onwards,

$$V\frac{\partial f(n,\mathbf{r})}{\partial z} = -\frac{\partial}{\partial n}\left(A(n)f(n,\mathbf{r}) - B(n)\frac{\partial f(n,\mathbf{r})}{\partial n}\right), \tag{6.33}$$

where

$$A(n) = g(n) - d(n) - \frac{\partial B(n)}{\partial n}, \tag{6.34}$$

and

$$B(n) = \frac{1}{2}\left(g(n) + d(n)\right). \tag{6.35}$$

Writing the FPE in the form of eq. (6.33) makes explicit the form of the n-space flux, which is related to the steady-state nucleation rate, $I_n(\mathbf{r})$ and is defined as

$$I_n(\mathbf{r}) = A(n)f(n,\mathbf{r}) - B(n)\frac{\partial f(n,\mathbf{r})}{\partial n}. \tag{6.36}$$

Before eq. (6.33) is discretized in size and space and solved numerically, an artificial diffusion term is also added to preserve the ellipticity of the FPE in the spatial dimension. The final result is

$$V\frac{\partial f(n,\mathbf{r})}{\partial z} = D_A\frac{\partial^2 f(n,\mathbf{r})}{\partial z^2} - \frac{\partial}{\partial n}\left(A(n)f(n,\mathbf{r}) - B(n)\frac{\partial f(n,\mathbf{r})}{\partial n}\right) \tag{6.37}$$

The magnitude of the artificial diffusion coefficient is identical to the one used in the discrete rate equations described in Section 6.3.1.

6.3.3 Boundary Conditions

Boundary conditions for the point defect portion of the model are the same as those described in Chapter 5; equilibrium concentrations at the melt/solid interface and no-flux conditions at the top and cylindrical surfaces:

$$C_X(r,0) = C_X^{eq}(T_m), \tag{6.38}$$

and

$$\mathbf{n} \cdot \mathbf{J}_X|_{\partial D_{cry}} = 0, \quad X = I, V, \tag{6.39}$$

where, for negligible thermodiffusion, the flux of the point defect X is written as

$$\mathbf{J}_X \equiv -D_X(T)\nabla C_X. \tag{6.40}$$

The discrete rate equations for small clusters are subject to essential boundary conditions at the melt/solid interface and no-flux conditions at the crystal cylindrical surface, so that

$$C_{n,X}(r, 0) = 0, \tag{6.41}$$

$$\mathbf{e}_z \cdot \mathbf{J}_{n,X}|_{\partial D_{cry}} = 0, \quad X = I, V \tag{6.42}$$

where

$$\mathbf{J}_{n,X} \equiv -D_A(T)\nabla C_{n,X}. \tag{6.43}$$

The Fokker-Planck equation for large clusters is subjected to similar conditions in the spatial dimensions; essential conditions on the melt/solid interface and no-flux at the top of the crystal:

$$f(n, r, 0) = 0 \tag{6.44}$$

$$\mathbf{e}_z \cdot \mathbf{J}_{fX}|_{\partial D_{cry}} = 0, \quad X = I, V \tag{6.45}$$

In the radial dimension, transport can only occur due to the artificial diffusion term. No-flux boundary conditions are imposed at both radial boundaries of the crystal, the centerline and the cylindrical surface. As in the case of the OSF-ring model of Chapter 5, radial structure can only appear because of the variation in the thermal field. In the case of the one-dimensional simulations which will be used in the following sections, the radial component of the problem vanishes.

The size component of the FPE requires two additional boundary conditions, one each for the smallest (n_{min}) and largest (n_{max}) cluster sizes represented. At n_{max},

no-flux boundary conditions are imposed, so that

$$\left(\frac{\partial I}{\partial n}\right)_{n=n_{max}} = 0 \tag{6.46}$$

This boundary condition requires that n_{max} be chosen sufficiently large so as to minimize the effects of any conflict between the solution and the no-flux constraint. There is no *a priori* method to determine the appropriate choice for n_{max}. We pursue investigation of the solution sensitivity to this parameter in a later section.

At n_{min} both the cluster concentrations and the nucleation flux are matched to the discrete representation. The concentrations are set to be equal in all regions of the crystal and this matching condition provides an essential boundary condition for the FPE at n_{min}, which is given by

$$f_X(n_{min}, r, z) = C_{n_{min},X}(r, z), \quad X = I, V. \tag{6.47}$$

The flux equality condition gives

$$-A(n)\left[\frac{\partial f_X(n, r, z)}{\partial n}\right]_{n_{min}} + B(n) f_X(n_{min}, r, z) = J(n_{min}, r, z), \quad X = I, V, \tag{6.48}$$

where $J(n_{min}, r, z)$ is the discrete nucleation flux evaluated at the matching cluster size. Eqs. (6.47) and (6.48) replace the DRE and the FPE at n_{min}.

6.4 Numerical Method

The overall system represented by eqs. (6.20)-(6.22) and (6.37) is $(n+1)$-dimensional where n represents the number of spatial dimensions. In our particular case, the crystal geometry is represented by a two-dimensional axisymmetric cylindrical coordinate system. The spatial components of the problem are discretized using the Galerkin finite element method as discussed in Section 5.3.4. The n-space portion of the FPE, however, poses some difficulties that require an alternative approach. In particular, the positivity of the solution is not guaranteed by a standard Galerkin finite element

method. Petrov-Galerkin streamline-upwinding (SUPG) (Johnson, 1987) fails to correct this problem for nucleation rates of interest. Generally, the presence of solution oscillations in regions of steep change in diffusion-convection type problems are due to the presence of a high Peclet number, or small diffusivities. In the size dimension, this would imply that the ratio of nucleation drift to diffusion, $A(n)/B(n)$, is very large. SUPG effectively lowers this ratio (which can be regarded as a n-space Peclet number), with an additional diffusion-type term. This term is introduced by adjusting the test function in the generation of the weak form of the equation. Instead of using the linear basis function, ψ, we now use

$$\psi + \epsilon h A(n) \frac{\partial \psi}{\partial n} \tag{6.49}$$

where h is the mesh spacing. This formulation leads to an additional diffusion term that scales with the magnitude of convection and the mesh, limiting how quickly the solution can change in each element. The magnitude of the additional diffusion is controlled with the parameter ϵ. This method was found to fail when realistic nucleation rates were introduced into the problem. Strong solution oscillations led to negative values of the concentration at certain points within the domain. The region where the solution begins to oscillate could not be stabilized with increasing mesh refinement or increasing the size of the SUPG parameter.

To successfully stabilize the solution for all nucleation rates, the finite-difference method of Chang and Cooper (Chang and Cooper, 1970) is employed for discretization of the n-space problem. The resulting finite difference equations are then solved at each spatial node using bilinear finite elements as before. This discretization method is an important part of the implementation of our aggregation model and is described in detail in the following section.

6.4.1 Implementation of the Chang-Cooper Method (CC70)

The underlying goal of the CC70 (Chang and Cooper, 1970) approach is to provide a weighted difference scheme for the n-space component of eq. (6.37) which will (a)

preserve the positivity of the numerical solution for all n and (b) conserve the total number of particles in the system. Chang and Cooper (1970) showed that simple centered differencing of most FPE equations leads to an artificial loss of particles at a rate which is related to the mesh size. Thus, unless the mesh is extremely fine ($\Delta n/n_{max} < 10^{-2}$), centered differencing can lead to stability problems. Furthermore, CC70 demonstrate that the convergence of the difference scheme to the differential equation is determined by $n\Delta n$ and not Δn which leads to oscillations and negative solution spectra at large n for most reasonable meshes, and explains the difficulty encountered when using the finite element method described above.

The solution to these issues is provided by using a centered difference approximation for the total flux term in eq. (6.36) and then applying weighted differencing to the individual components of the flux. The former step leads to

$$V\frac{\partial f(n,r,z)}{\partial z} = D_A\frac{\partial^2 f(n,r,z)}{\partial z^2} - \frac{I_{m+1/2}(r,z) - I_{m-1/2}(r,z)}{\Delta x_m}, \tag{6.50}$$

for $m = 0, 1, ..., M$ where I is the nucleation flux defined in eq. (6.36). The points $\{m + 1/2\}$ are defined so that

$$x_{m+1/2} \equiv \frac{x_{m+1} + x_m}{2}, \tag{6.51}$$

and

$$\Delta x_{m+1/2} = x_{m+1} - x_m, \tag{6.52}$$

which gives

$$\Delta x_m \equiv \frac{x_{m+1} - x_{m-1}}{2}. \tag{6.53}$$

Now, the Fokker-Planck diffusion and convection coefficients can be written as

$$A_m = A(x_m), \tag{6.54}$$

and

$$A_{m+1/2} = \frac{A_m + A_{m+1}}{2}. \tag{6.55}$$

Similar expressions hold for $B(n)$. The discretized flux can now be written as

$$I_{m+1/2} = A_{m+1/2} f_{m+1/2} - B_{m+1/2} \left(\frac{\partial f}{\partial n} \right)_{m+1/2}. \tag{6.56}$$

If centered differencing were applied to the total flux in eq. (6.50) we would obtain

$$\left(\frac{\partial f}{\partial n} \right)_{m+1/2} \sim \frac{1}{\Delta n} (f_{m+1} - f_m), \tag{6.57}$$

and

$$f_{m+1/2} = \frac{1}{2} (f_{m+1} + f_m). \tag{6.58}$$

The discussion in CC70 shows that this scheme would lead to negative values of f_m as the flux term becomes small or as the size of n becomes large. To overcome this difficulty, a new description for $f_{m+1/2}$ is developed based on the requirements of particle conservation and solution positivity. This is represented by

$$f_{m+1/2} = (1 - \delta_m) f_{m+1} + \delta_m f_m, \tag{6.59}$$

where $0 \leq \delta_m \leq 1/2$. In this way, the $m+1/2$ value is located somewhere in between the m and $m+1$ values. This is tantamount to introducing a variable differencing scheme that adjusts between centered ($\delta_m = 1/2$) and forward differencing ($\delta_m = 0$). An expression for how δ_m should vary across the size domain is derived by requiring that the numerical flux approach the exact flux as the solution tends to equilibrium (defined by the zero flux condition). The result of this analysis gives the following description of the discrete flux (Chang and Cooper, 1970; Park and Petrosian, 1996) for the FPE given in eq. (6.37):

$$I_{m+1/2} = \frac{B_{m+1/2}}{\Delta x_{m+1/2}} \left[W^+_{m+1/2} f_{m+1} - W^-_{m+1/2} f_m \right], \tag{6.60}$$

where,

$$W^{\pm}_m = \frac{w_m}{2 \sinh \frac{w_m}{2}}, \tag{6.61}$$

$$w_{m+1/2} = \frac{A_{m+1/2}}{B_{m+1/2}} \Delta x_{m+1/2}. \tag{6.62}$$

If these expressions are substituted into eq. (6.50) at each of the mesh points, m, a system of linear algebraic equations results which must be solved at each point in space. These equations can be written in a tridiagonal form as (Park and Petrosian, 1996)

$$V\frac{\partial f}{\partial z} - D_A \frac{\partial^2 f}{\partial z^2} = -a_m f_{m-1} + b_m f_m - c_m f_{m+1}. \tag{6.63}$$

The coefficients on the interior mesh points, $m = 1, ..., M-1$ are given by

$$a_m = \frac{1}{\Delta x_m} \frac{B_{m-1/2}}{\Delta x_{m-1/2}} W^-_{m-1/2}, \tag{6.64}$$

$$c_m = \frac{1}{\Delta x_m} \frac{B_{m+1/2}}{\Delta x_{m+1/2}} W^+_{m+1/2}, \tag{6.65}$$

$$b_m = \frac{1}{\Delta x_m} \left[\frac{B_{m-1/2}}{\Delta x_{m-1/2}} W^+_{m-1/2} + \frac{B_{m+1/2}}{\Delta x_{m+1/2}} W^-_{m+1/2} \right]. \tag{6.66}$$

The exterior mesh points are treated separately because of the boundary conditions. The equality of concentrations and fluxes at n_{min} replace both the discrete rate and Fokker-Planck equations at that point. In the CC70 formulation, the flux is computed explicitly only at locations in between nodes. The discrete nucleation flux, however, is defined only at the boundary between the two representations which lies on a nodal position. The continuous flux is extrapolated to the matching point using a finite-difference approximation so that

$$F_0 = F_{1/2} - \frac{dF}{dn}(n_{1/2} - n_{min}) = J_{n_{min}}, \tag{6.67}$$

where n_{min} corresponds to the point $m = 0$ in the n-space discretization.

The no-flux condition at n_{max}, eq. (6.46), is written as

$$f_{m+1} = \frac{W^-_{m+1/2}}{W^+_{m+1/2}} f_m, \tag{6.68}$$

which, on substitution into the above expressions for a_m, b_m, and c_m, gives

$$c_M = 0, \tag{6.69}$$

$$b_M = \frac{1}{\Delta x_m}\left[\frac{B_{m-1/2}}{\Delta x_{m-1/2}} W^+_{m-1/2} + \frac{B_{m+1/2}}{\Delta x_{m+1/2}} W^-_{m+1/2}\right] - \frac{B_{m+1/2}}{\Delta x_{m+1/2}} \frac{W^-_{m+1/2}}{W^+_{m+1/2}}. \tag{6.70}$$

and leaves a_m unchanged.

Discretization of the n-Space Component

The discretization of the n-space mesh is performed over $M+1$ discrete points ranging from n_{min} to n_{max}. In order to accommodate the many orders-of-magnitude between these limits, the mesh is graded using a power-law expression that smoothly increases the mesh size, h, so as to minimize the required number of n-space nodes for a converged solution. The mesh scaling law is given as

$$n_{x+1} = n_x + (1+\delta)^x, \tag{6.71}$$

where δ is an adjustable parameter and controls mesh refinement across the entire domain.

6.5 Thermophysical Properties

The estimates for the point defect thermophysical properties computed in Chapter 5 are used in all simulations described here and are not subjected to any further adjustments. Thermophysical properties for aggregates are estimated from a combination of molecular dynamics simulations of large supercells as described in Chapter 4 and phenomenological models. Thermodynamic integration was used to compute temperature dependent free energies of formation for clusters up to size seven. The results were fitted to biquadratic polynomials in size and temperature. The resulting polynomials are now rescaled so that the single point defect free energies match the results of Chapter 5, namely eqs. (5.89)-(5.93). The functional form of the biquadratic

fitting polynomial is given by

$$G_X^f(n,T) = a_1^X + a_2^X T + a_3^X T^2 + a_4^X Tn + a_5^X Tn^2 + a_6^X T^2 n^2 + a_7^X T^2 n + a_8^X n + a_9^X n^2. \quad (6.72)$$

$a_1^I = 2.1000$ $a_1^V = 0.9995$	$a_2^I = -0.5082$ $a_2^V = 0.5657$	$a_3^I = 0.1000$ $a_3^V = -0.1812$
$a_4^I = 0.2733$ $a_4^V = -0.3745$	$a_5^I = -0.04451$ $a_5^V = 0.07361$	$a_6^I = 0.02679$ $a_6^V = -0.01388$
$a_7^I = -0.2888$ $a_7^V = 0.007579$	$a_8^I = 1.6440$ $a_8^V = 1.7000$	$a_9^I = 0.02060$ $a_9^V = -0.06700$

Table 6.1: Coefficients for the biquadratic polynomial fit to cluster free energies of formation. Temperature is given in degrees Kelvin.

Clusters containing more than seven point defects are assumed to be spherical in shape and their free energies of formation are described by a two-component function:

$$G_X^f = n\Delta G_B(T) + n^{2/3}\sigma(T), \quad (6.73)$$

where $\Delta G_B(T)$ is the chemical potential difference between the cluster phase and the perfect crystal and $\sigma(T)$ is the surface free energy of the cluster. The temperature dependence of both these functions is assumed to be quadratic, so that

$$\Delta G_B(T) = \gamma_1 + \gamma_2 T + \gamma_3 T^2, \quad (6.74)$$

$$\sigma(T) = \sigma_1 + \sigma_2 T + \sigma_3 T^2. \quad (6.75)$$

For self-interstitial aggregation, the six free energy coefficients are determined by matching the free energy and its derivative at three distinct temperatures for interstitial and vacancy clusters containing seven point defects. Void clusters do not possess

a volume energy and are characterized by the surface energy only. This means that, for vacancies, $\{\gamma_i\} = 0$ and only the $\{\sigma_i\}$ are available for fitting. The results of these fitting procedures are given below in Table 6.5.

$\gamma_1^I = 1.6203$ $\gamma_1^V = 0$	$\gamma_2^I = -8.266 \times 10^{-3}$ $\gamma_2^V = 0$	$\gamma_3^I = 4.3214 \times 10^{-7}$ $\gamma_3^V = 0$
$\sigma_1^I = 0.8954$ $\sigma_1^V = 3.511$	$\sigma_2^I = -1.368 \times 10^{-3}$ $\sigma_2^V = -2.494 \times 10^{-3}$	$\sigma_3^I = 9.926 \times 10^{-7}$ $\sigma_3^V = -4.094 \times 10^{-6}$

Table 6.2: Coefficients for the temperature dependence of the bulk and surface free energy contributions of spherical point defect clusters. Temperature is in Kelvin.

The results of the free energy matching are shown graphically in Figure 6-2 for clusters containing up to thirty point defects. Note that in the case of the vacancy energy curve, a jump in the gradient can be observed at the junction between the MD and phenomenological equation results because of the lack of fitting degrees of freedom. The MD results show that for very small aggregates the free energy does not scale with a spherical surface area. Equation (6.73) gives an effective surface energy of 1.25 J/m^2 for large vacancy clusters. This result, which is implicitly based on the results of the MD simulations described in Chapter 4 is in good agreement with experimental estimates for the energy of the $< 111 >$ surface of silicon which are about 1.0 J/m^2 (Nakamura *et al.*, 1996). This agreement, in conjunction with the observation that the MD results predict $< 111 >$ facetting of voids in the silicon lattice (see Chapter 4), suggests that the energy of a large void is essentially the sum of $< 111 >$ surface energies around the void. Such a model has already been employed by Dornberger (1997) in similar investigations.

The possibility that an aggregation free energy barrier exists is left as an open question because there is no data in the literature concerning this quantity. The base case simulations are performed with the assumption that the barrier height is zero, i.e. $\Delta G_B = 0$. The size of this barrier can be used as an adjustable parameter for

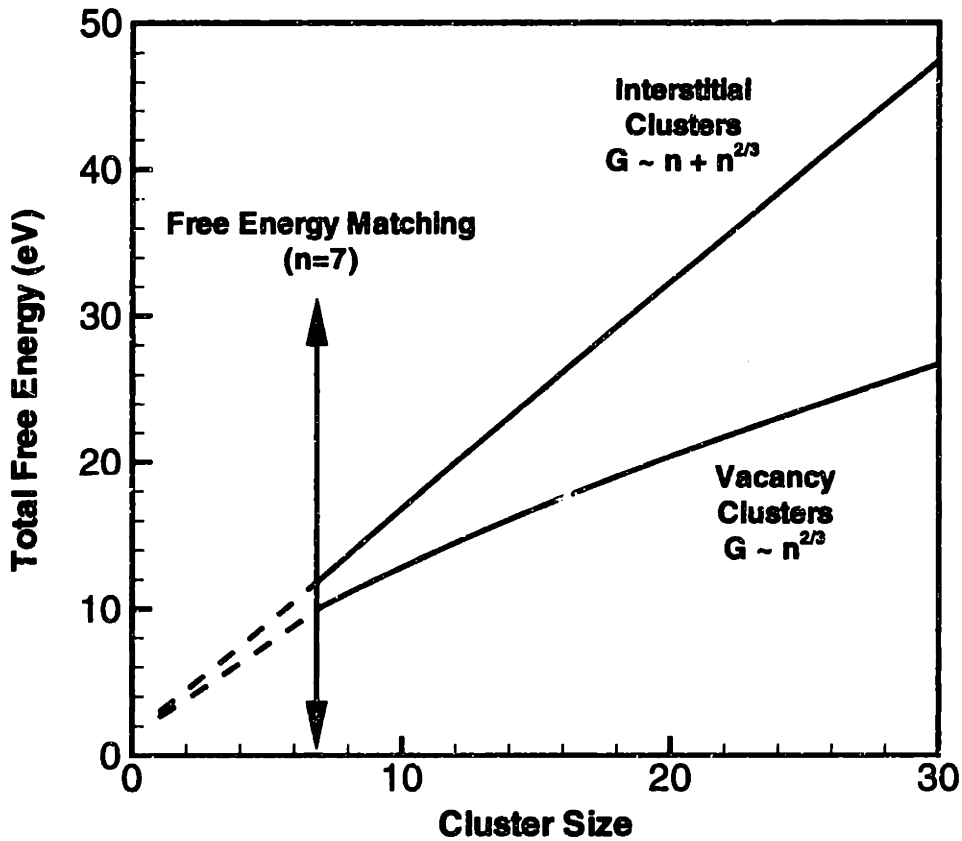

Figure 6-2: Cluster free energy of formation per point defect as a function of cluster size.

future parametric investigations.

6.6 One Dimensional Model for Vacancy Aggregation Physics

Detailed experimental data for microdefect distributions as a function of thermal and physical operating conditions, which we need to compare our predictions, are very difficult to obtain. This is in part due to the proprietary nature of such data, as well as the difficulty in isolating and identifying the constituents of a particular microdefect species. However, the recent interest in the void microdefects (D-defects) in the region inside the OSF-ring in CZ crystals has led to the publication of some experimental observations. In particular, the dependence of the size distribution of these microdefects on the crystal cooling rate in certain temperature regions is well documented, at least on a qualitative level (Dornberger and von Ammon, 1996; Takano *et al.*, 1995a). A common result of thermal history experiments is that the final D-defect density is related to the cooling profile of the crystal in a relatively narrow temperature region; the exact range of this region is still under dispute but estimates for the onset of nucleation range from about 1323 K - 1475 K, while the cessation of cluster profile evolution is thought to occur at about 1173 K - 1273 K. This temperature range will be referred to as the nucleation temperature from now on.

Characterization of the thermal field is critical for directing industrial efforts to optimize the thermal environment during crystal growth. The thermal history experiments described in Chapter 2 are aimed at determining whether the final void concentration can be described in terms of an average cooling rate across a certain temperature range. Such information can then be used to design a thermal environment that would affect the crystal cooling rate in that region alone. The cooling rate, Z, across any temperature range is defined as

$$Z \equiv V < G >, \tag{6.76}$$

where $< G >$ is an average temperature gradient in a specified interval. The results of Takano et al. (1996) and Dornberger et al. (1997) seem to indicate that the temperature interval responsible for most of the void nucleation and growth is not more than about 100-150 K. This hypothesis will be investigated in detail in Section 6.9 by computing the evolution of the cluster size distribution in terms of the temperature profile.

The detection limit for D-defects is currently about 50 *nm* (Itsumi *et al.*, 1995) and defects as large as 500 *nm* are observed under certain growth conditions. The size distribution of these defects is observed to result in a peak of cluster density at some intermediate size, which also depends on the thermal history of the crystal. Because it is not possible to observe D-defects that are smaller than 50 *nm* in diameter, only portions of the predicted size distributions can be compared. The location of a peak in cluster distribution, the total defect density, and the nucleation temperature range will be used as a basis for comparing the results of the model to experiment.

The investigation of vacancy aggregation physics in the region inside the OSF-ring is carried out with a one-dimensional model. The simulations described in Chapter 5 are first used to find a set of operating conditions that lead to a vacancy dominated crystal. These conditions are then used in simulations that include the full nucleation problem. The advantages to this approach is the relatively low computational cost of a two-dimensional calculation as compared to a three dimensional one. A fully three-dimensional system generally requires on the order of 10^3 nodes in the spatial dimensions. Reasonable n-space meshes require at least 100 nodes for each point defect species which leads to problems with up to 200,000 unknowns in the discrete problem. Problems of this size are tractable but are certainly not suited to investigative or parametric studies.

6.6.1 Thermal Fields for the One-Dimensional Model

In order to compare our results for cluster size distribution and nucleation temperature dependence directly to experimental observation, we employ one dimensional polynomial fits of the simulated thermal fields used in Chapter 5. In each case, a

tenth-order polynomial is fit to the temperature field at a particular radial position within the crystal. This polynomial order was found to be sufficiently accurate in the region near the melt/solid interface to reproduce the OSF-ring location to within one percent of the crystal radius. This method has the advantage of providing a very realistic thermal field even though the simulation domain is one-dimensional. This method also provides a direct way to account for cross-effects such as variations in the temperature gradient due to changes in the pull rate. The drawback is that detailed analysis of the effect of thermal history is difficult because it is very difficult to obtain enough simulated thermal fields to isolate various regions of the crystal cooling profile.

A possible solution to the above problem is the use of an analytic expression for the thermal field. An example is the use of a simple cooling-fin in a constant ambient which cools axially according to

$$T(z) = T_m \left(T_{amb} + \left(\frac{T_m - T_{amb} \cosh \lambda}{\sinh \lambda L} \right) \times \frac{\cosh \lambda (L - z)}{\sinh \lambda} \right). \qquad (6.77)$$

The parameter λ controls the temperature gradient by specifying how quickly the crystal cools to the ambient temperature, T_{amb}. The range $0.1 \leq \lambda \leq 0.5$ is chosen so that the thermal gradient near the melt/solid interface is physically reasonable. The thermal field given by eq. (6.77) captures the essential features of a typical thermal field by exhibiting a high temperature gradient near the melt/solid interface which levels off quickly as the crystal cools. While clearly not applicable for comparing our results to experimental data, this approach allows for more detailed study of thermal history effects, by allowing very easy modification of the cooling rate in different sections of the crystal and decoupling the effects of pull rate on microdefect formation.

The first thermal model is used for all simulations described in this Chapter. The variability in the available data is sufficient to demonstrate most of the experimentally observed D-defect responses to changes in the operating conditions. Three heat shield configurations, HS1, HS2, and HS3, which represent progressively faster cooling

environments are used. The centerline of a four inch crystal grown in the HS3 configuration at a pull rate of 1.8 mm/min is used for the base case. This configuration is used to demonstrate the effects of pull rate on D-defect distribution. Thermal history effects are investigated by comparing distributions from the centerline of two eight inch crystals grown in HS1 and HS2 environments, at a pull rate of 0.7 mm/min. Finally, the radial distribution of D-defects is probed with multiple axial sections of the base case crystal.

6.7 Mesh Selection for Computing Cluster Distributions

The convergence of the solution with respect to size and space discretization is investigated systematically. The size and space meshes are varied separately in a series of runs with progressively higher refinement. The HS3/4" crystal growth configuration at a pull rate of 1.8 mm/min is used for all mesh refinement runs. All mesh refinement studies are performed based on the characteristics of predicted void size distributions only because negligible quantities of interstitial clusters are formed in the selected portions of the crystal and this part of the solution is ignored.

The following plots show the void size distribution at the top of the one-dimensional crystal. Figure 6-3 shows the convergence of the solution with respect to the n-space mesh represented on a log-log scale showing the entire solution domain. In all cases, n_{max} is fixed at about 1×10^{10}. The transition between the DRE and FPE representations is located at $n = 10$ for all the following simulations. Sensitivity to the location of the matching region is described below. A peak in the size distribution appears at about $10^6 - 10^7$ vacancies per cluster. The concentration of clusters in the size interval $10^1 < n < 10^3$ is extremely low and is subject to significant roundoff error. The degree of roundoff error within this interval was found to increase with mesh refinement as the "true" solution was captured in more detail. Convergence of the solution in this region is not considered to be an important requirement because

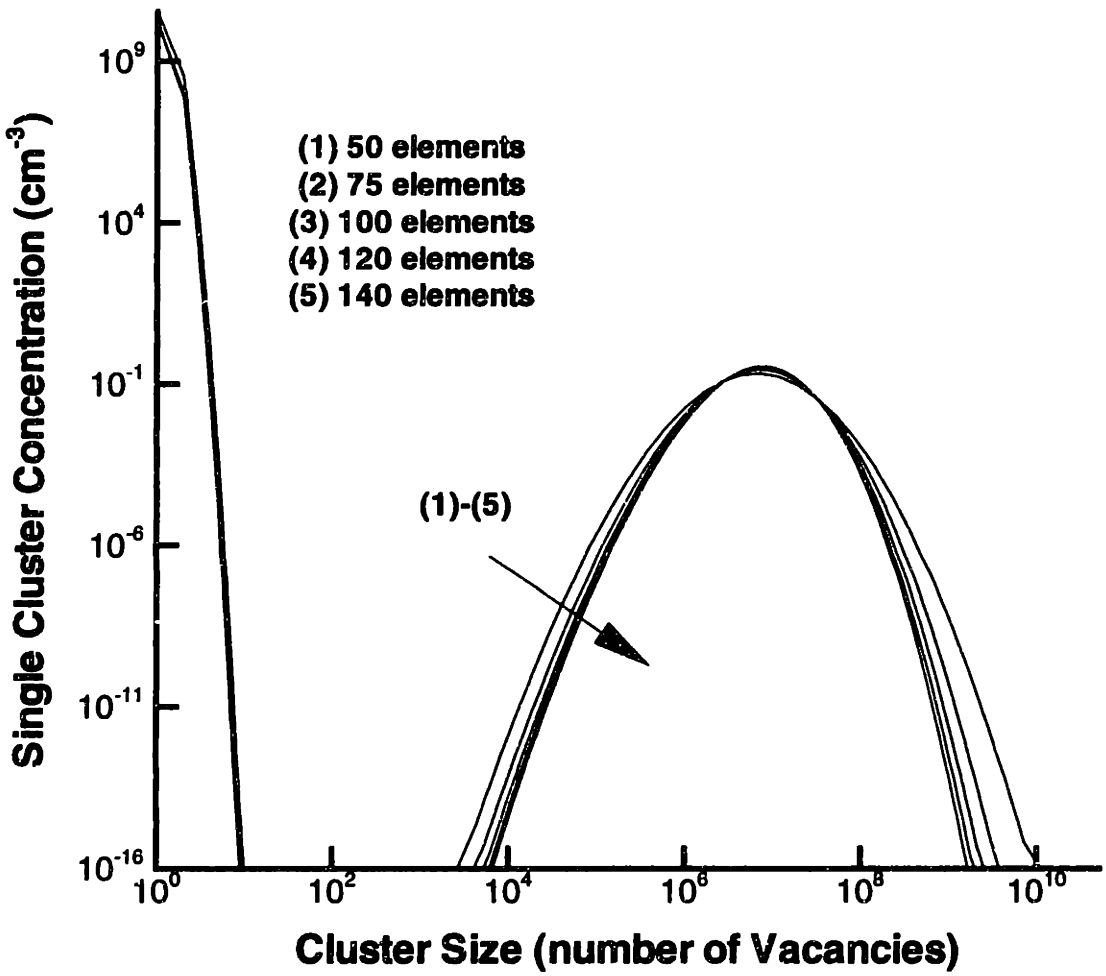

Figure 6-3: Convergence of void size distribution with respect to n-space mesh refinement.

cluster concentrations below 10^{-16} cm^{-3} are physically meaningless.

The critical test of accuracy lies in the convergence of the peak region with respect to mesh refinement. Figure 6-4 uses a log-linear scale showing only the peak in void concentration. While the peak concentration is very sensitive on the level of refinement for the given meshes, the location of the peak is essentially unaffected. The range of mesh refinement used in our investigation represents 5 n-points per

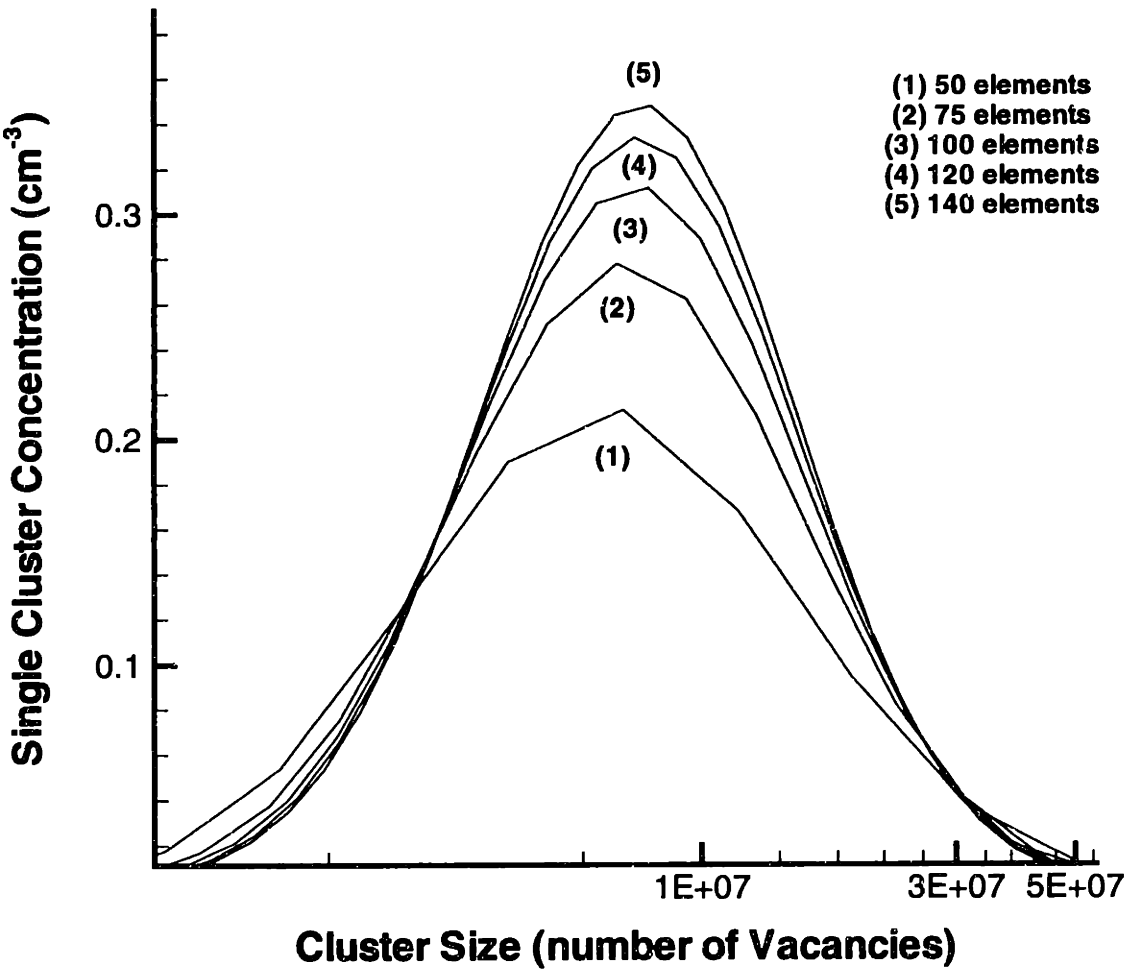

Figure 6-4: Convergence of void size distribution with respect to n-space mesh refinement.

decade of size distribution at the coarsest level upto about 16 n-points per decade for the most refined mesh. Curves (4) and (5) give peak heights that are within 5 % of each other and are considered to be sufficiently accurate to use for production runs.

The computational cost of performing an LU decomposition at each iteration in the Newton method scales as NB^2, where N is the total number of unknowns and B is the bandwidth of the banded matrix which results from the discretization of our problem. For a one-dimensional space-mesh, the bandwidth is directly proportional to the number of unknowns at each spatial node which gives the result that the cost of the computation scales cubically with the number of n-mesh points. This condition imposes a severe limit on the allowable refinement in the size dimension and is the main driving force for the selection of the mesh refinement levels shown in Figure 6-3. At the level of refinement shown in curve (5), each Newton iteration requires about 18 minutes of CPU time. This value must be considered in the context of the number of iterations required for each run. The sphere of convergence for this problem is very small and continuation from an artificial problem must be used. In particular, the nucleation and point defect recombination rates are increased from 10^{-3} of their actual values in a series of about 15-18 steps, each requiring about 4 Newton iterations. If each step takes about 18 minutes, the total time for a simulation is about 24 hours. Clearly, this represents an upper limit on the allowable length for simulations that are to be performed many times.

By contrast, spatial discretization in one dimension results in a linear scaling of the computational cost with mesh refinement. The convergence of the solution with respect to spatial mesh refinement is demonstrated in Figures 6-5 and 6-6 which represent the solution on logarithmic and linear coordinates respectively. In Figure 6-5 the solutions with the two coarsest meshes show oscillatory behavior in the region $10^1 < n < 10^3$ where the concentration of clusters is extremely small. Once again, the location of the void size distribution peak is much less sensitive to the mesh refinement than the height of the peak for the range investigated. Curves (4) and (5) differ by about 4 % for their prediction of the peak height in Figure 6-6 while curves (1) and (2) differ by about 28 % for the same difference in mesh refinement level.

In all cases, the grading of the spatial mesh is designed to capture both the rapid point defect recombination near the melt/solid interface and the nucleation further down the crystal. The mesh is graded so that the last 40 % of the crystal domain,

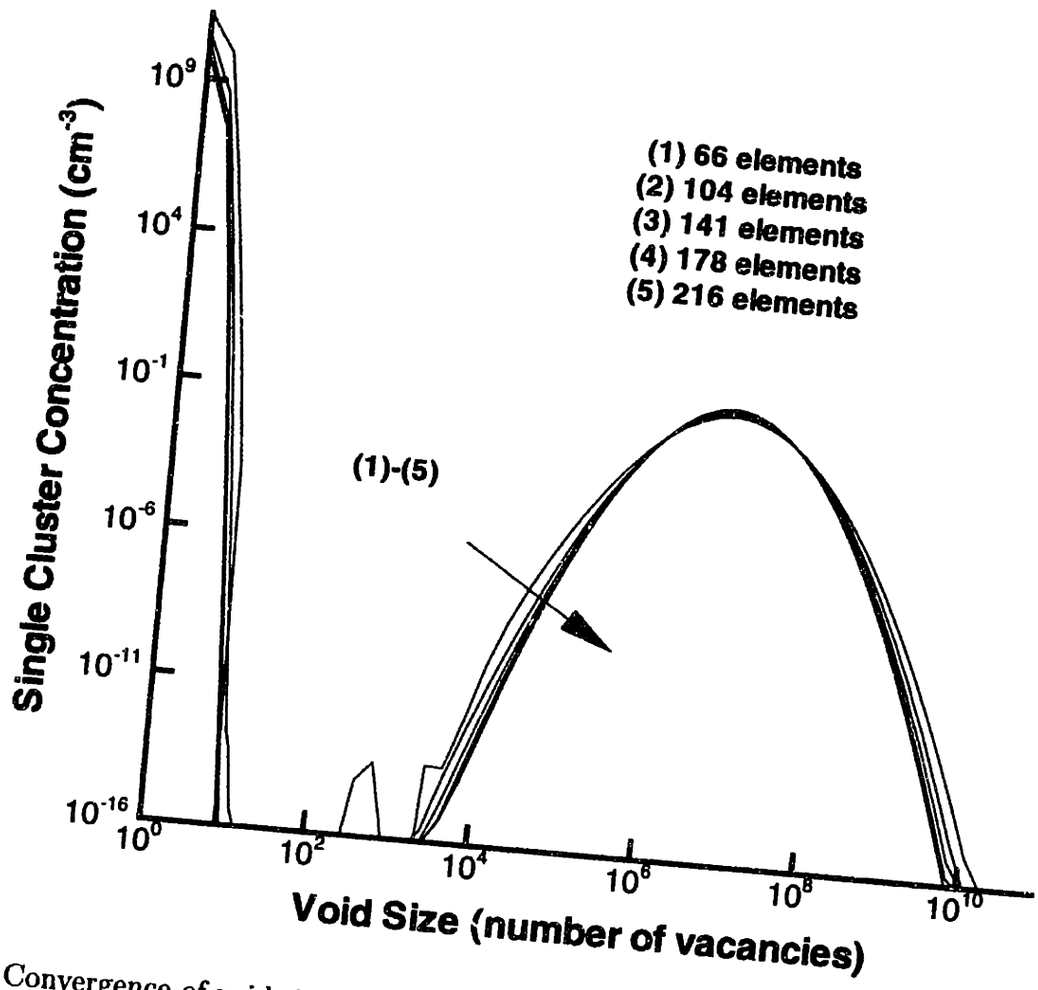

Figure 6-5: Convergence of void size distribution with respect to spatial mesh refinement. Full solution on log-log scales.

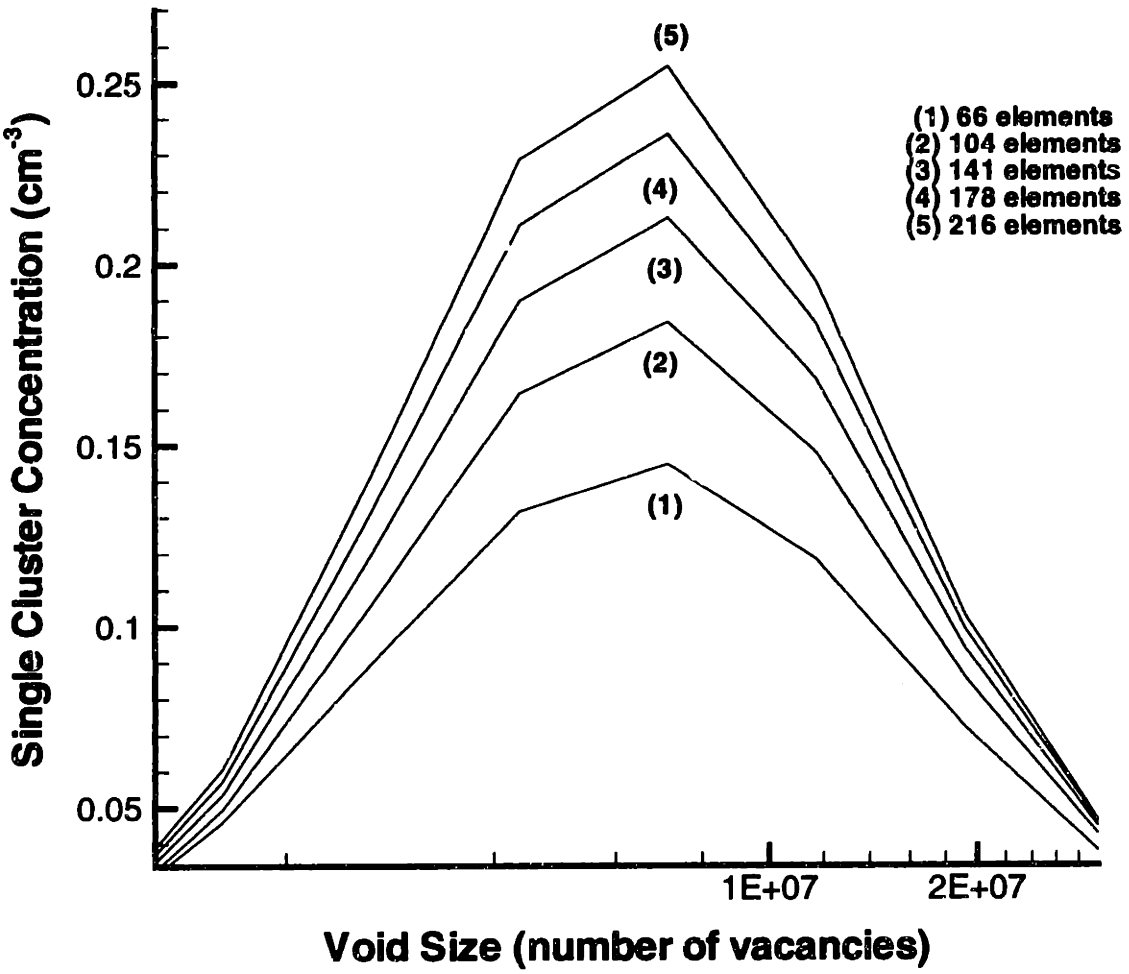

Figure 6-6: Convergence of void size distribution with respect to spatial mesh refinement. Peak section of the solution on log-linear scales.

where concentration gradients are gentler, gets progressively coarser. Convergence of the solution with respect to spatial mesh refinement is demonstrated further in Figure 6-7 which shows the axial evolution of the total cluster concentration. The total cluster concentration is simply the integral of the size distribution curve across all cluster sizes and will be considered in detail in the following section. Once again, it is clear from Figure 6-7 that curves (4) and (5) are very close to each other and support the conclusion that the refinement level represented by curve (4) in Figure 6-6 is sufficient to give a reasonably accurate solution.

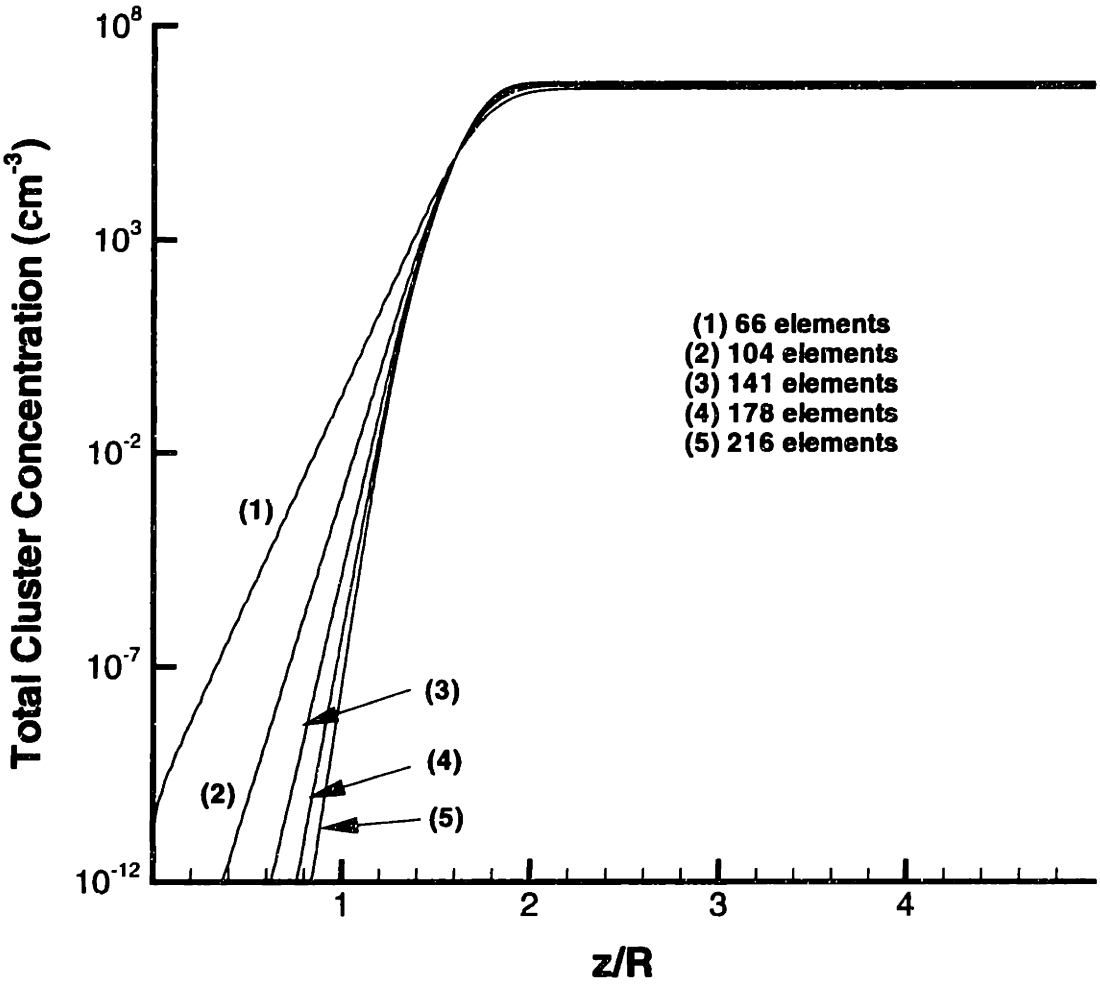

Figure 6-7: Convergence of the solution in space as a function of mesh refinement.

The results of the mesh refinement study described in this section show that

curves (4) in both the n-space and z-space tests are sufficiently accurate considering the uncertainty in the rest of the model. Of course, the most important result is that convergence with mesh *is* obtained in both dimensions and numerical uncertainty can be reduced to any level desired. All further investigations are performed with 178 z-space elements and 130 n-space elements for the vacancy portion. The interstitial n-space mesh is specified at a much coarser level in the following simulations because they consider exclusively vacancy-rich portions of the crystal, where negligible quantities of interstitials survive to form aggregrates. The matching point between the DRE and FPE representations is fixed at $n = 10$ as the results are very weakly dependent on this choice for the range $n > 10$. The insensitivity of the solution with respect to the matching point is demonstrated in Figure 6-8 which shows the peak region for matching points at $n = 10$ and $n = 20$. The resulting problem contains about 52,000 unknowns; this large number casts serious doubts on the feasibility of a three dimensional simulation with the current implementation. This point will be addressed further in a later section.

6.8 Quantitative Evaluation of Predicted Void Size Distribution for the Base Case Conditions

Characterization of the various features of our results is carried out on the base case operating conditions used in the mesh refinement study discussed in the previous section. The base case configuration corresponds to the centerline of a 4" crystal grown with the HS3 configuration at a pull rate of 1.8 mm/min.

6.8.1 Size Distribution on an Interval Basis

The raw size distributions shown in Figures 6-3 and 6-6 cannot be used to compare our predictions to experimental measurements of FPD, COP, or LSTD concentrations in real silicon crystals. Experimental measurements lead to estimates for the number of defects in a given *size interval* and not at a particular size (the raw data

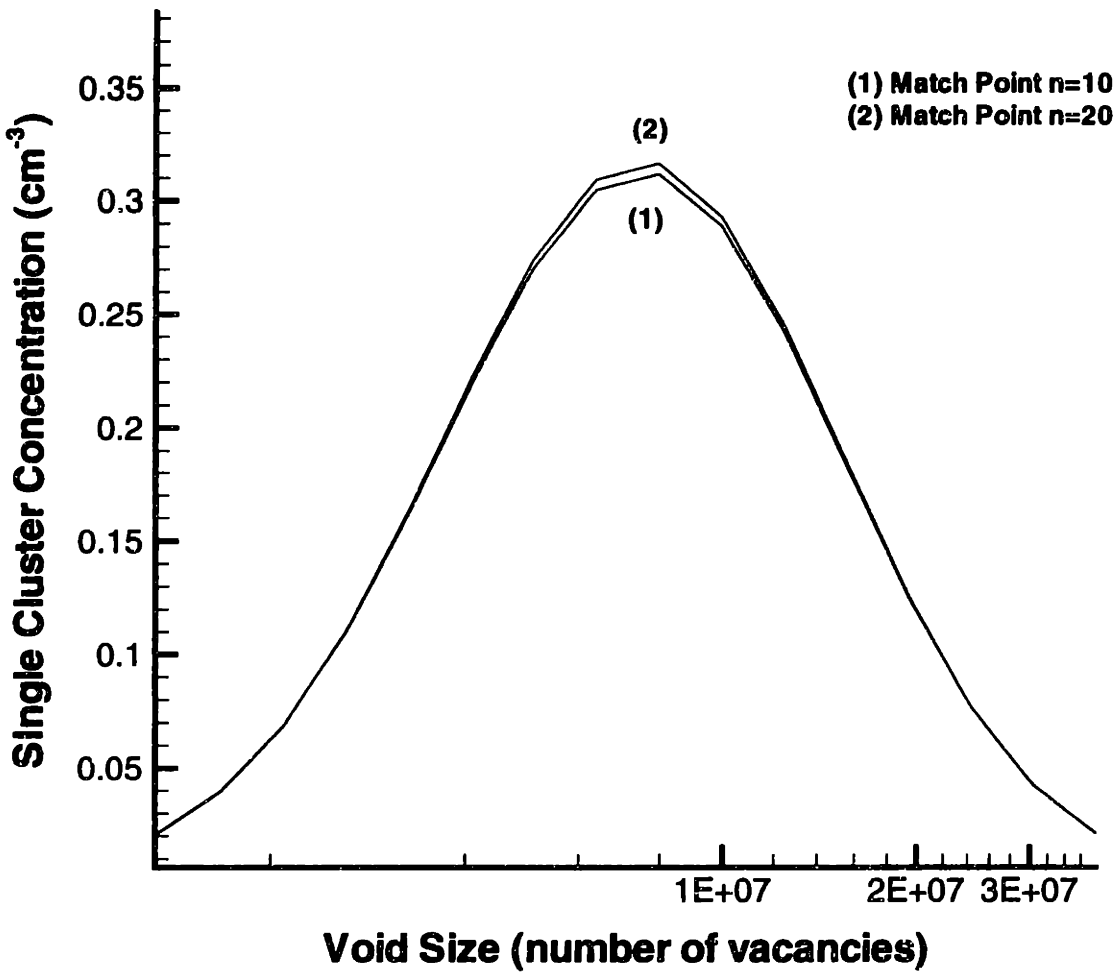

Figure 6-8: Convergence of the solution in space as a function of mesh refinement.

from the simulations). To compare the simulation predictions with experiment, the size distribution is integrated and the total number of clusters over fixed intervals is computed.

In all cases, a size interval of 1 nm diameter is chosen as the interval width and provides a consistent basis for comparing our results to each other and to experimental measurements. Figure 6-9 shows the size distribution on an interval basis for the base case operating conditions. The peak in void concentration is about 1.3×10^5 cm^{-3} and corresponds to voids of about 75 nm in diameter. Note that the absolute concentrations of these voids are not meaningful without knowledge of the interval width; larger intervals will obviously lead to higher concentrations.

6.8.2 Evolution of the Total Observable Defect Density

One of the actively researched aspects of D-defect formation in CZ-grown crystals is the dependence of defect concentration and size on the dwell time of the crystal in certain thermal regimes. Some researchers have put forward the idea that the resulting D-defect density is dependent on the time that a crystal spends in a fairly narrow temperature range somewhere between 1423 K and 1173 K. So-called dwell-time experiments have been designed to determine this critical temperature range (Dornberger *et al.*, 1996a; Takano *et al.*, 1995a; Hourai *et al.*, 1995). These experiments consist of growing several crystals under similar conditions but with different thermal environments and then comparing the total D-defect density (Dornberger *et al.*, 1996a) or by changing the crystal pull rate abruptly during growth and observing where the D-defect density changes (Takano *et al.*, 1995a). While most of these studies are in agreement that certain regions of a crystal's thermal history are more important for D-defect nucleation and growth than others, the actual temperature range is still under debate. The results of Dornberger et al. (1997) imply that nucleation is still proceeding at temperatures down to 1173 K, while most of the other investigators believe that the important temperature range is above 1273 K.

The existence of a microdefect nucleation/growth temperature region has been justified in terms of the evolution of point defect supersaturation during crystal cool-

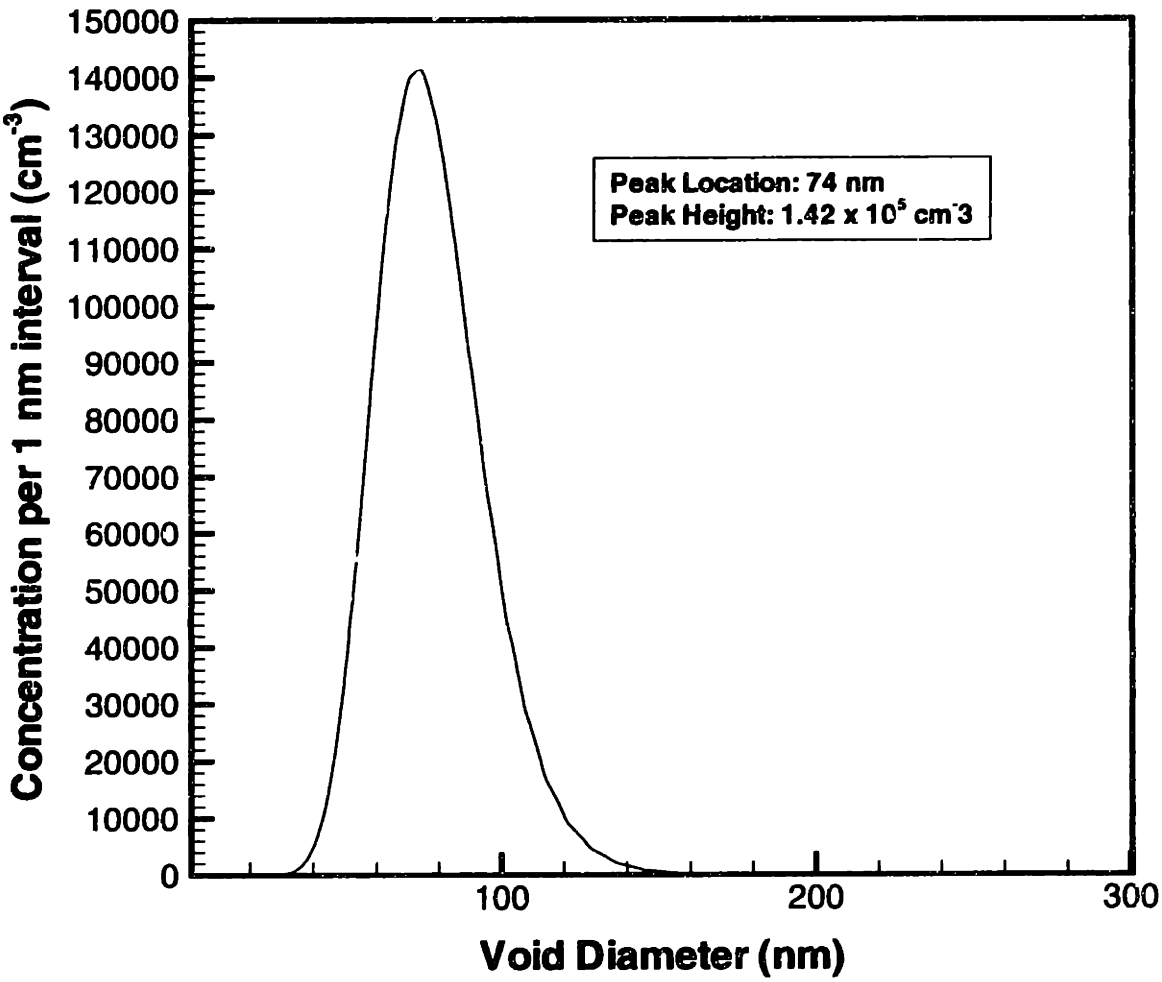

Figure 6-9: Vacancy aggregate size distribution at the crystal top as a function of cluster diameter assuming that the cluster is spherical. Interval size is 1 *nm*.

ing. Above a certain temperature point defect recombination keeps both species from achieving high enough supersaturations to initiate cluster formation. In regions away from the OSF-ring, one of the two native point defect species will rapidly be depleted and recombination ceases. The remaining species no longer has a mechanism by which to reduce its supersaturation except for aggregation into clusters which then proceeds until either the crystal becomes too cold or the point defect species is depleted.

The evolution of the cluster size distribution in space along the axis of the crystal is investigated by considering the integral of the size distribution at various axial positions within the crystal. The total cluster density is defined as

$$F(z) = \int_{n_{min}}^{\infty} f(n', z) dn' \tag{6.78}$$

The choice of n_{min} depends on the particular data that is being considered. Direct measurements of D-defect concentrations, either by light-scattering (which measures LSTDs) or laser particle counting (COPs) is limited to voids that are larger than about 50 nm in diameter. This corresponds roughly to about 3.6×10^6 vacancies per void. Measurements of flow pattern defects (FPDs) usually give concentrations that are about two orders-of-magnitude lower than those for LSTDs or COPs. The concentrations of FPDs, however, correlate very well with the other two and it is likely that all three defect types have the same origin. The apparent discrepancy between FPD and LSTD concentration can be explained by assuming that flow patterns form around a small fraction of existing voids.

Other measurements correlate the Gate-Oxide Integrity (GOI) yield to the operating conditions (Dornberger *et al.*, 1996a). In this case, n_{min} is closer to 5 nm or about 3000 vacancies per void. Figure 6-10 shows both these integral quantities evolving in terms of the crystal axial temperature profile, i.e. $F(T(r, z))$ for fixed r. Both n_{min} limits give very similar results for the base case operating conditions. The two curves are identical until the larger clusters begin to form at about 1300 K. At this point the curves diverge slightly as the clusters of size $O(10^3)$ form before clusters of size $O(10^6)$ but both integrals approach each other by about 1120 K at which point

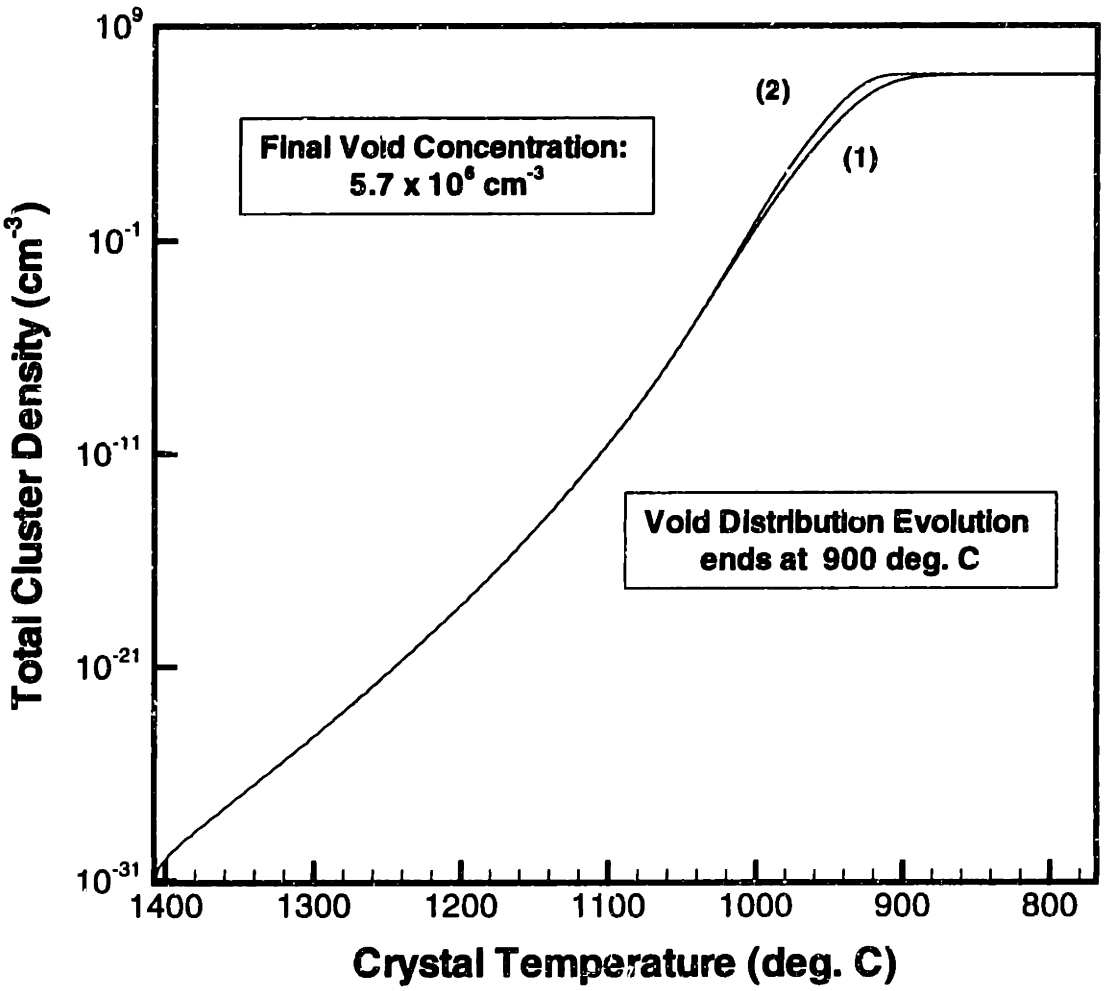

Figure 6-10: Spatial evolution of the total cluster density with n_{min} defined as (1) and (2) total *experimentally observable* clusters.

nucleation and cluster growth has ceased. The final void density for both representations is about 6×10^6 cm^{-3} which is in excellent agreement for typical LSTD or COP concentrations in fast cooled crystals (Takano *et al.*, 1995a). For crystals with high concentration of voids in the size interval, $10^3 < n < 10^6$, the two representations are expected to differ significantly.

Our base conditions correspond to a cooling rate of about 9 K/cm in the temperature interval, 1423 K - 1353 K, which has been determined to be the nucleation/growth region as defined by Takano et al. (1995). Figure 6-11 shows the results of Takano et al. (1995) which correlate the final LSTD density with the cooling rate in this temperature region. Extrapolation of the trend in this figure to the cooling rate in our base case gives an rough estimate for the LSTD density of about $8-9 \times 10^6$ cm^{-3}, which is in good *quantitative* agreement with our value. The critical point to be made here is that this agreement is achieved with *no parameter adjustment.*

Dornberger (1997) has investigated the effects of the crystal cooling rate on the residual D-defect size distribution by comparing observed GOI yields for crystals grown under different thermal conditions. The GOI yield is proportional to the number of D-defects (or voids) present in the crystal that are larger than about 5 nm as described above. It was found that the degree of correlation between the final concentration of GOI defects and the dwell time at various temperatures changed over the course of the cooling of the crystal, peaking in the region between 1173 K and 1323 K. Figure 6-12 shows the distribution of the goodness-of-fit as a function of the temperature; both GOI and FPD levels show very similar correlation patterns which is further evidence that they are related. Our prediction for the total GOI-related cluster evolution shown as curve (2) in Figure 6-10 is in complete agreement with these experimental results. Both sets of results show that there is a steep decline in cluster distribution evolution past 1173 K (900 C). This region corresponds to the point where vacancies are sufficiently depleted that thermodynamics no longer favors the rapid formation of clusters.

At temperatures higher than 1173 K, aggregates form at an increasing rate. Experimentally, this is verified by the increasing correlation between dwell time at a

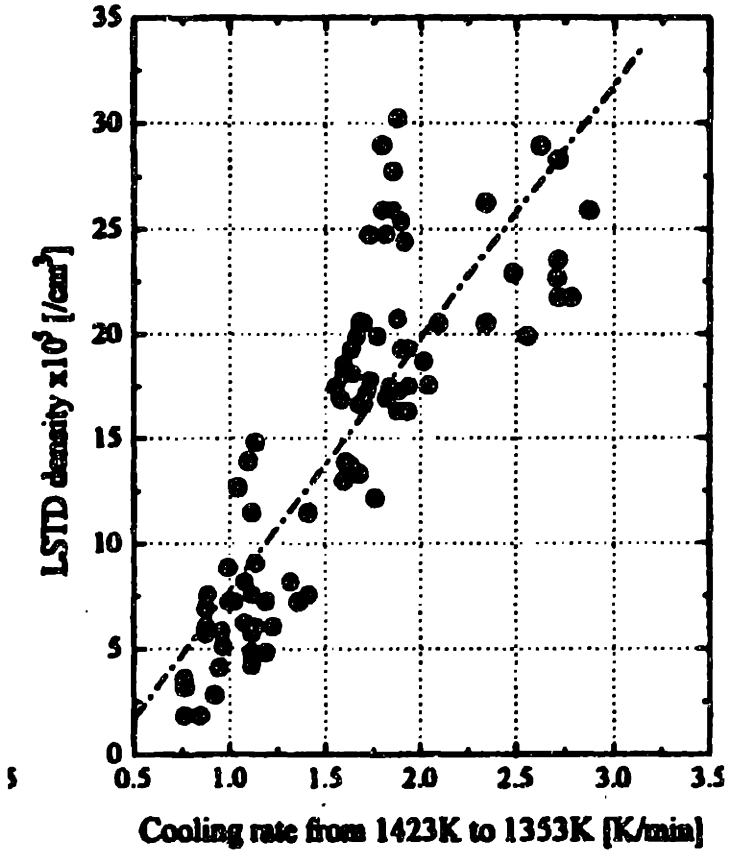

Figure 6-11: Final LSTD density as a function of crystal cooling rate in the temperature region $1353K < T1423K$. Taken from Takano et al. (1995)

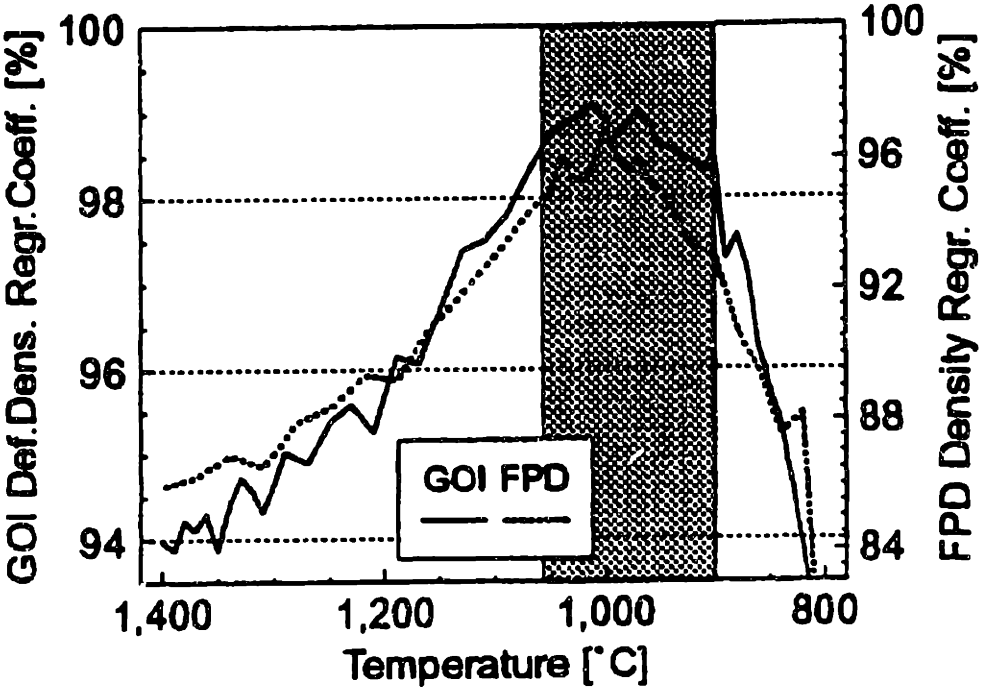

Figure 6-12: Correlation between dwell time in 10 K segments and the total density of GOI defects. Taken from Dornberger (1997).

particular temperature and the final GOI defect concentration. The simulation results show a similar trend where the logarithmic slope of the cluster concentration increases as the temperature decreases. The steepest portion of curve (2) in Figure 6-10 roughly corresponds to the peak region in Figure 6-12, in the temperature range $1323K > T > 1173K$. However, the results shown in Figure 6-10 do not indicate a temperature at which nucleation suddenly initiates. Rather, nucleation appears to take place starting near the crystal boundary, parallel to point defect recombination. Thus, the existence of a narrow temperature region for controlling D-defect formation does not seem to be justified from our analysis. The pull rate ramp experiments of Takano et al. (1995) do show that D-defect concentration vary rapidly in a narrow region but it is very possible that this phenomenon is entirely due to transient effects which cannot be captured with the present steady state analysis.

6.9 Prediction of Void Size Distributions for Different Crystal Growth Operating Conditions

The previous section has outlined several methods to present our simulation results in a way which can be used to quantitatively and directly compare our results to experimental data. These are essentially the total D-defect concentration, the binned size distribution, and the spatial evolution of cluster density as a function of crystal thermal history. The following section describes the application of the model to the investigation of the effects of crystal pull rate and thermal field on the grown-in void distribution. Also, void size distribution is compared in the base case crystal for several radial positions, ranging from the axis of the crystal to near the OSF-ring at the periphery. These comparisons have important implications in determining the radial uniformity of the defect profile in a given crystal. This issue becomes even more important in large diameter crystals. Furthermore, a radial study will be useful in attempting to explain the origin of the OSF-ring itself, which as of yet, remains only partially characterized.

6.9.1 Effect of Pull Rate Variation on Void Size Distribution

The HS3/4" crystal growth system was used to compute void distributions for pull rates of 1.8 *mm/min* (base case), 1.5 *mm/min*, and 1.2 *mm/min*. The OSF-ring model described in Chapter 5 gives OSF-ring positions for these three pull rates at about 45 %, 70 % and 90 % of the radius, respectively. The void size distribution per 1 *nm* intervals is shown part (a) of Figure 6-13. As the pull rate decreases from 1.8 to 1.2 *mm/min* the size distribution remains essentially the same, but the peak height decreases by about 50 %. The distribution is not observed to shift or spread out significantly. Part (b) of Figure 6-13 shows that the total cluster concentration is lower at all temperatures as the pull rate decreases. The reasons for this trend are evident from the evolution of the vacancy profiles in space. At lower pull rates, the crystal centerline is closer to the OSF-ring region which implies that the balance between interstitials and vacancies, while still clearly in favor of vacancies, is moving towards the equality condition, $\Delta = 0$.

Note that we use the total cluster density defined with the GOI related lower cutoff here, $n_{min} = 5$ *nm*. This quantity measures the total number of voids *that are large enough to affect device gate-oxide integrity* and not only those that can be detected with particle counters or light scattering. For most operating conditions, the large majority of voids are above 50 *nm* and the two quantities are almost identical, as shown in Figure 6-10. As a result, fewer vacancies survive after the recombination reaction is extinguished and nucleation proceeds more slowly because the supersaturation level is lower. In the case of the fastest pull rate, the cluster density stops increasing at about 1180 K (910 C). At lower pull rates, the density continues to evolve until about about 1140 K (870 C) but this is not enough to compensate for the late start and the final cluster density is lower. It is interesting to note that the higher the post-recombination vacancy level, the lower the post-nucleation level.

These results have important implications for specifying crystal growth operating conditions. High pull rates lead to greater concentrations of D-defects in the region

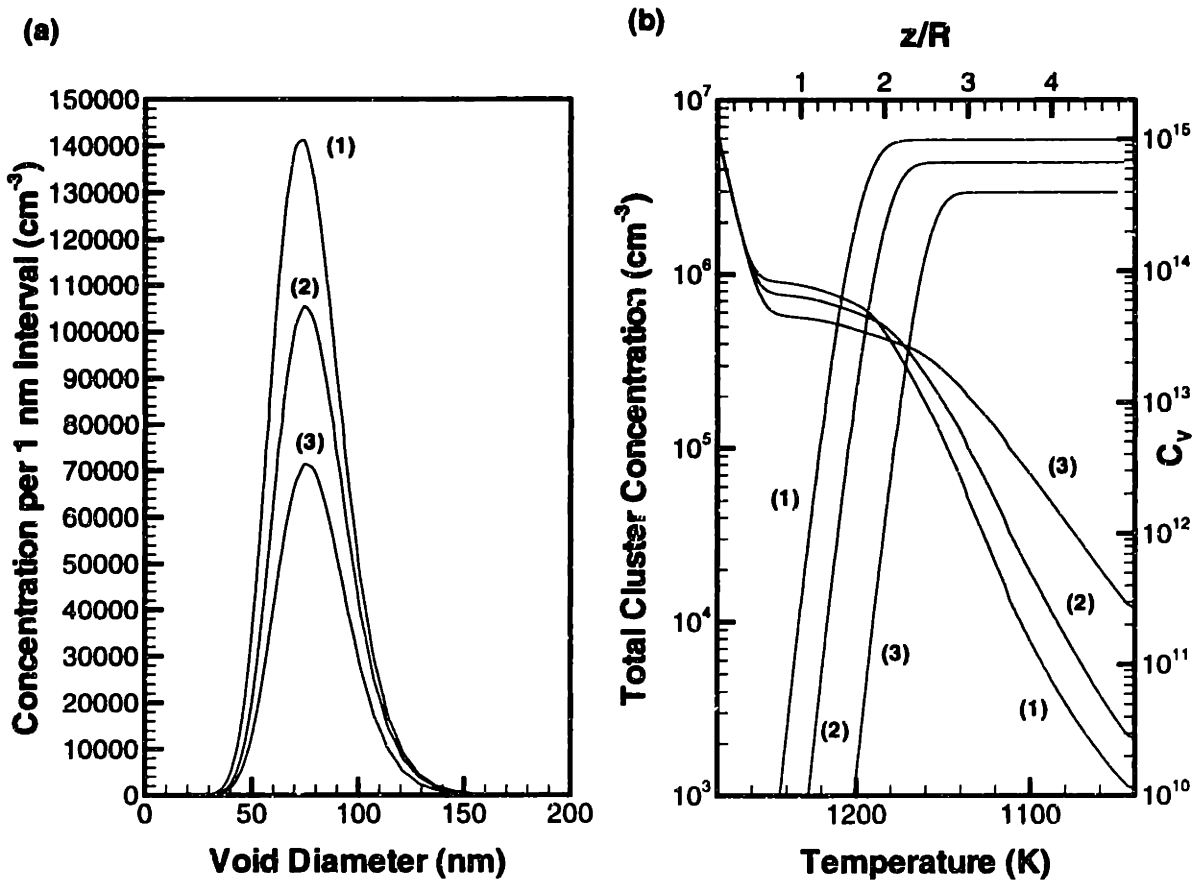

Figure 6-13: Effect of pull rate on (a) void size distributions for the HS3/4" crystal growth case. (1) $V_p = 1.8\ mm/min$, (2) $V_p = 1.5\ mm/min$, and (3) $V_p = 1.2\ mm/min$; (b) Vacancy evolution in space and Total Cluster evolution as a function of temperature.

inside the OSF-ring. Given the constraint that the OSF-ring has to be at the periphery of the crystal, the pull rate should be kept to a minimum if D-defects are to be minimized. The fact that residual monomer vacancy concentrations are higher at the lower pull rates is also interesting from the point of view of analyzing the formation of the OSF-ring. The residual monomer concentration in the OSF-ring region is very likely to be the reason for the formation of stacking faults after wet oxidation. This issue is investigated next by using the base case configuration to simulate a series of axial sections at radial positions ranging from the centerline to inside periphery of the OSF-ring.

6.9.2 Radial Variation in the Void Size Distribution

The HS3/4" crystal is used to investigate the radial variation in the void size distribution with a series of one-dimensional axial segments located at progressively larger radii. The results presented in Figure 6-14 are in the same format as those shown in Figure 6-13, and for reference, show the base case configuration as Curve (1). Part (a) of Figure 6-14 shows that as the radius of the crystal is traversed, the void size distribution changes dramatically. The predicted position of the OSF-ring is at 94 % of the crystal radius if nucleation is neglected. For radial positions up to 97 % of the crystal radius, we find that the crystal is still vacancy dominated; this implies that the predicted location of the OSF-ring is somewhat affected by the nucleation problem. This is expected to some degree as nucleation has been shown to occur all the way up to the melt/solid interface, while point defect recombination is still taking place. This effect is very small because curve (5) in Figure 6-14 is within 1 % of the interstitial/vacancy boundary, or the $\Delta = 0$ position.

At radial positions up to 90 % of the crystal radius there is relatively little change in the size distribution. The peak location moves progressively towards smaller clusters and while it initially becomes higher compared to the centerline value, it eventually starts to decrease in height. At 96 % of the crystal radius, the size distribution starts to change rapidly; the distribution becomes bimodal and the concentration of very small clusters (under 10 nm) increases dramatically. By 97 %, the second peak is

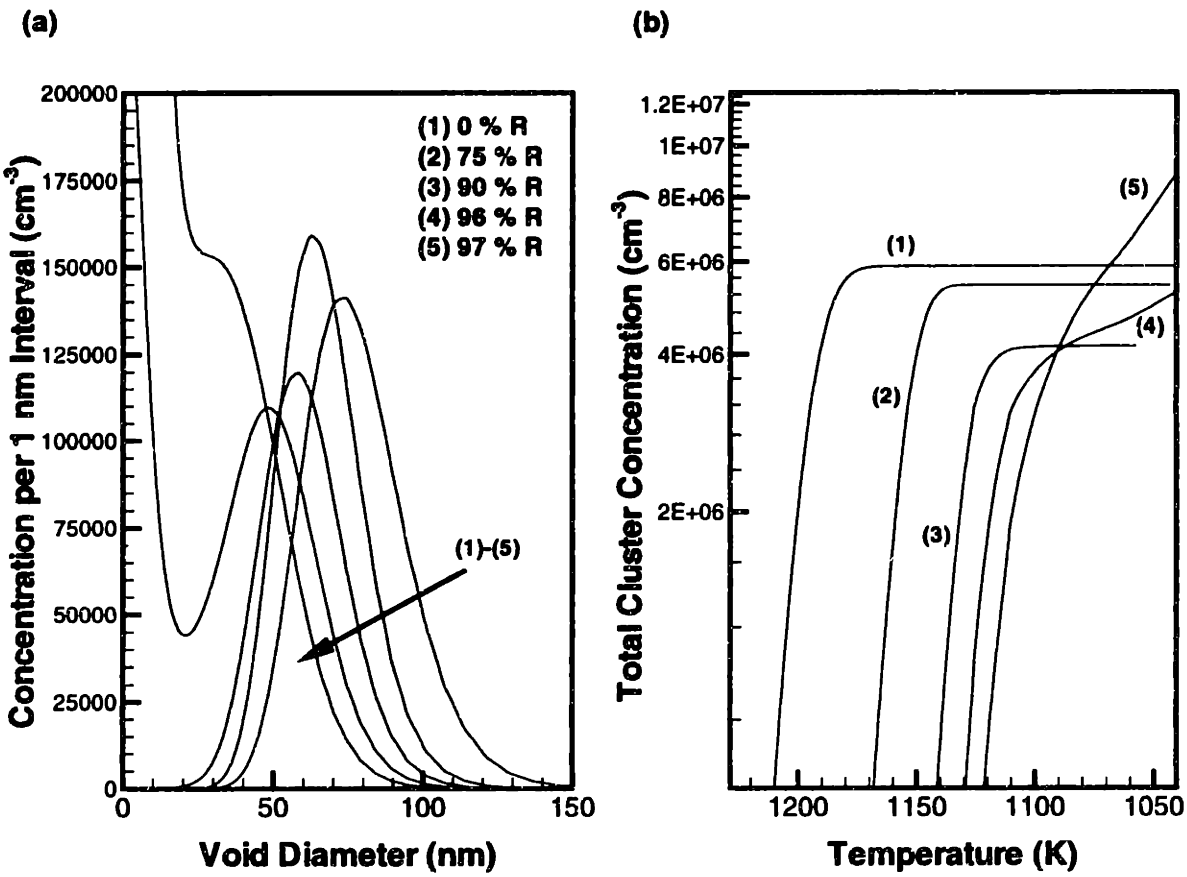

Figure 6-14: Radial variation in (a) void size distributions for the HS3/4" crystal growth case. (1) $r = 0$, (2) $r = 75\%R$, (3) $r = 90\%R$, (4) $r = 96\%R$, and (5) $r = 97\%R$; (b) Total Cluster evolution as a function of temperature.

no longer present and the dominant voids are those below 50 nm in diameter. While these would not be observed by current techniques, a large number of them would still degrade the Gate Oxide Integrity of devices fabricated in this region.

Part (b) of Figure 6-14 shows that at larger radii, the number of voids continues to evolve well past the 1173 K point. For conditions well inside the OSF-ring, 1173 K represented the temperature at which nucleation and cluster growth ceased; the results of the previous sections show that the estimate for this temperature is fairly independent of the growth conditions. However, in the vicinity of the OSF-ring, vacancy dominance is sufficiently reduced so that the onset of nucleation is severely retarded. This leads to very large supersaturations at lower temperatures which then continues to drive nucleation and growth past 1173 K. In fact, the thermal fields used in these simulations do not extend to sufficiently low temperatures because the cluster size distribution is still evolving at the top of the crystal.

The evolution of the single vacancy concentration in space is shown separately in Figure 6-15 for clarity. The residual vacancy population in the region near the OSF-ring is almost three orders-of-magnitude greater than the vacancy concentration at the center of the crystal! This huge variation is almost certainly the reason why stacking faults nucleate preferentially in the region of the OSF-ring. While the interstitial nucleation problem has not been considered in detail here, we can construct a plausible mechanism for OSF-ring formation purely on the basis of our current understanding of point defect dynamics. The OSF-ring region is unique because it is the only region of the crystal that contains high concentrations of single vacancies after the crystal has cooled. Vacancies are required to nucleate the stacking-faults because they provide the free volume needed for oxygen precipitation. It is well known that the center of the interstitial-type stacking faults that appear in the OSF-ring is composed of an oxide precipitate (Hasebe *et al.*, 1989). It is likely that the picture is slightly more complicated and that another nucleating agent is needed in addition to vacancies because of the as of yet unexplained lack of an OSF-ring in highly oxygen doped FZ silicon. The explanation of this discrepancy will surely provide the last piece of the puzzle for characterizing the OSF-ring completely.

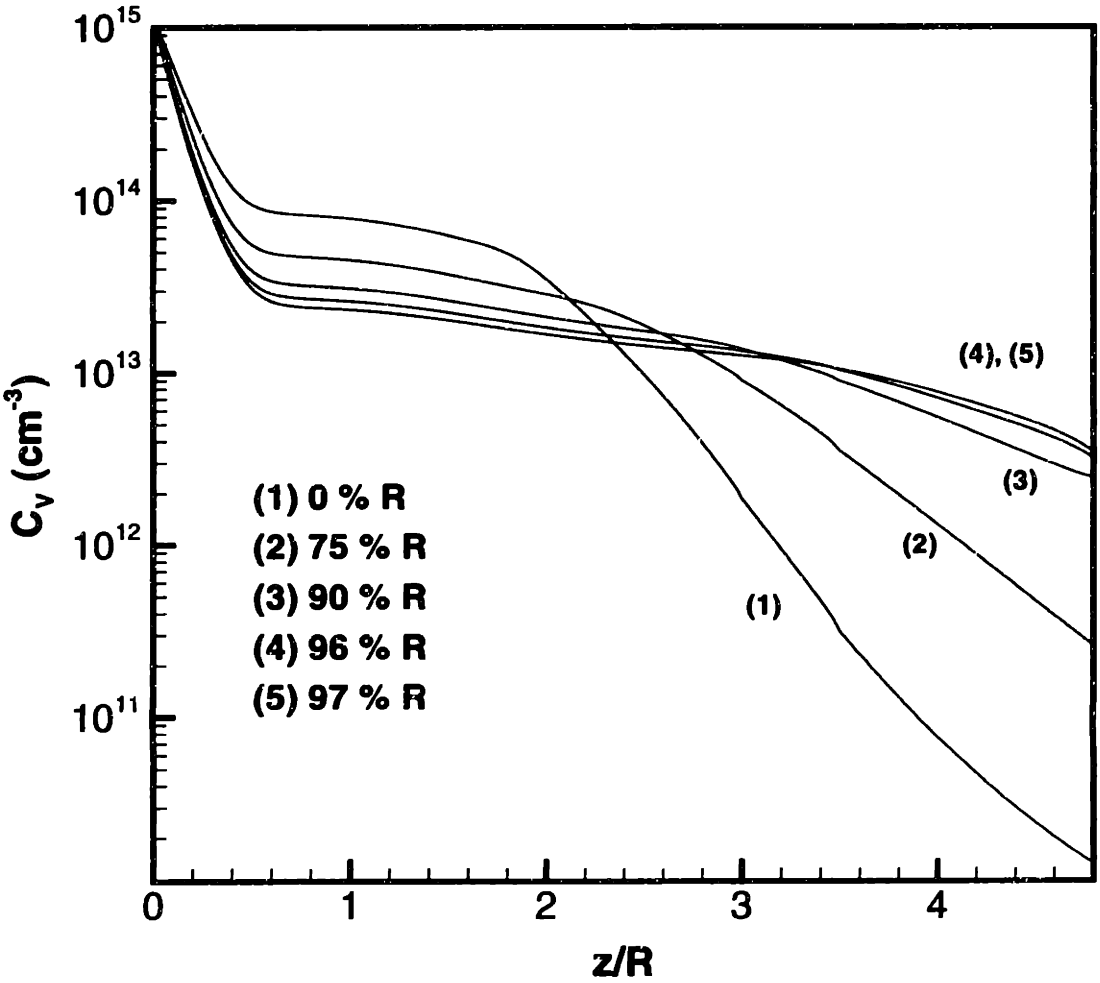

Figure 6-15: Radial variation in the single vacancy concentration for the HS3/4" crystal growth case. (1) $r = 0$, (2) $r = 75\%R$, (3) $r = 90\%R$, (4) $r = 96\%R$, and (5) $r = 97\%R$.

6.9.3 Thermal History Effect on Void Size Distribution

The effects of crystal cooling rate on the resultant void size distribution is investigated by comparing the predictions of the model at the centerline of two eight inch crystals grown at a pull rate of 0.7 mm/min with different heat shield configurations, designated as HS1 and HS2. The larger diameter of these crystals compared to the one use in previous analyses significantly lowers the cooling rate. In addition, both HS1 and HS2 are designed to lead to slower cooling as compared to the HS3 configuration, with HS1 providing the slowest cooling environment. Figure 6-16 shows the final void size distribution and the evolution of the vacancy and total cluster concentrations as a function of axial position. The size distribution in the crystal with the faster cooling rate (HS2) exhibits a higher peak at a smaller cluster size interval. Classical nucleation theory (Becker and Doring, 1935; Volmer and Weber, 1925) provides an explanation for this effect. The evolution of an aggregate size distribution starts with the formation of nuclei due to homogeneous supersaturation. These nuclei grow at different rates depending on their size. In slow cooling systems, the distribution has sufficient time to undergo the process of Ostwald ripening (Lifshitz and Slyozov, 1961). This is the process by which larger, more stable clusters grow at the expense of smaller, less stable ones. The net result of this process is a few, large clusters. The exact distribution is a function of the available supersaturation and the configurational entropy of the system. In the case of a growing crystal, the cooling rate limits the extent of Ostwald ripening and thus the size distribution does not have time to achieve equilibrium at each temperature.

In addition to the time allowed for classical nucleation to proceed, interstitial vacancy dynamics provide another mechanism to retard nucleation in the fast cooled crystal; the ratio, V/G, which was introduced in Chapter 5, is lower in the HS2 crystal because of the higher temperature gradient, G. This condition corresponds to a smaller number of vacancies surviving the recombination phase and so supersaturation builds up more slowly. This gives the crystal less time to develop large clusters (which consume smaller ones) and a distribution of many smaller clusters are frozen into the

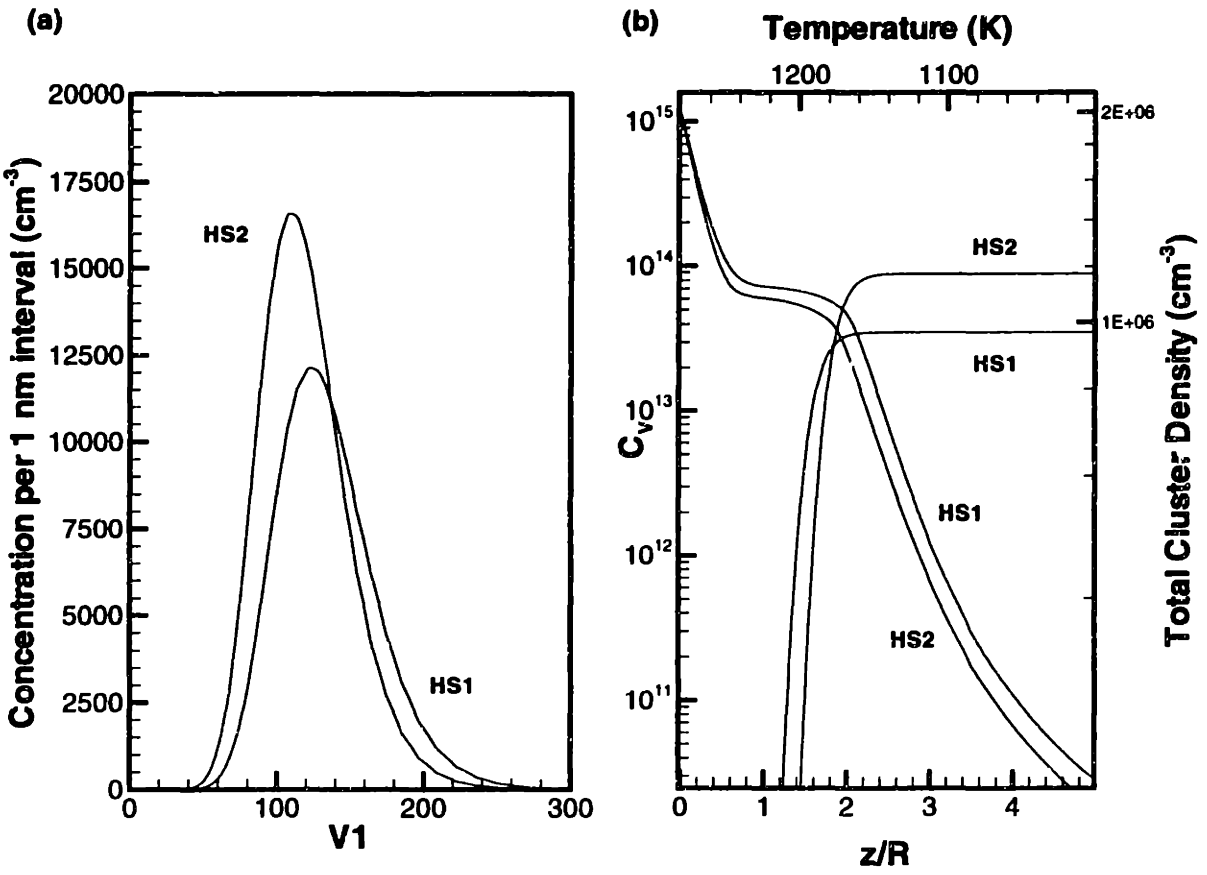

Figure 6-16: Thermal History Effect on (a) Void size Distribution; (b) Vacancy evolution in space and Total Cluster evolution as a function of temperature.

fully grown ingot.

The results for the total number of clusters larger then 5 nm are shown in part (b) of Figure 6-16. Defect concentrations of 1.2×10^6 and 9.7×10^5 cm^{-3} are present for the fast cooled and slow cooled crystals, respectively. These concentrations are compared to the experimental data of Takano et al. (1995), shown in Figure 6-11. The cooling rates in the temperature range, $1353 > T > 1423$ K, are 1.72 and 1.46 K/min for the HS2 and HS1 cases respectively. For these rates, the data of Takano gives a total LSTD concentration of about 1.6×10^6 and 1.2×10^6 cm^{-3}, respectively. Given that the experimental scatter is about 30-40 agreement between the simulation results and the data is very good. Our values are perhaps a little low and it is expected that some minor parameter adjustment would make the agreement better without sacrificing the fit to the OSF-ring data in Chapter 5.

6.9.4 Parametric Sensitivity of the Void Simulation to Thermophysical Properties

It is important to note that the point defect thermophysical properties used in the present model were fitted to very high temperature data. The validity of low temperature extrapolation of these properties is entirely dependent on the accuracy of the Stillinger-Weber interatomic potential. As previously discussed in Chapter 3, it is likely that the formation free energy prediction for the vacancy is inaccurate, and, while this did not pose a problem for the OSF-ring model, it is more important here. The processes of nucleation and aggregation continue to take place at appreciable rates at temperatures far below the melting point. For this reason, inaccuracy in the formation free energy (and the migration energy) will definitely lead to error in the predicted void size distribution. However, from the present analysis we see that if our interpretation of void formation is correct, the current point defect thermophysical properties are good enough to give quantitative agreement with experimental measurements of LSTD sizes and concentrations.

As a demonstration of model sensitivity to the point defect thermophysical prop-

erties, the vacancy diffusivity is increased by 30 % and the base case crystal used to compute void size distributions. This adjustment is chosen as an example because it does *not* affect significantly any of the OSF-ring results in Chapter 5. Its only effect is to raise the estimate for the vacancy contribution to the self-diffusion coefficient by 30 %. Figure 6-17 compares predicted void size distributions for the original and adjusted diffusivities. The perturbation in the vacancy diffusion coefficient shifts the

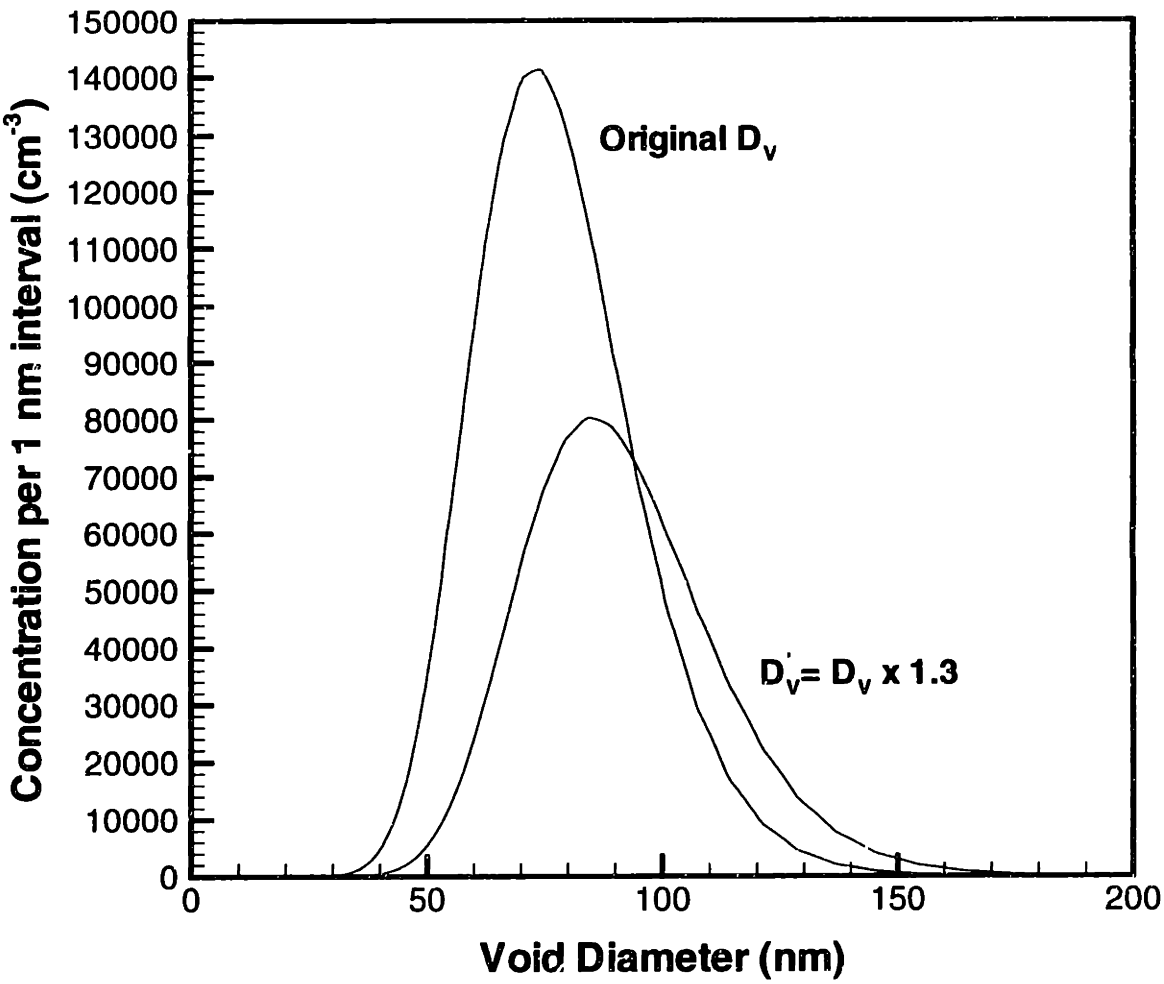

Figure 6-17: Sensitivity of the Aggregation Model to the vacancy diffusivity.

void size distribution to large sizes and lower concentrations, as expected. The peak height is observed to drop by about 50 % and is located at a void diameter of about

85 *nm* rather than 75 *nm*. This result is due to the fact that the higher mobility effectively increases the time that void can grow in any particular thermal regime. The results are clearly very sensitive on the thermophysical property estimates for point defects and the fact that good quantitative agreement is obtained without adjustment is another indicator of the overall validity of the model for the OSF-ring presented in Chapter 5.

6.10 Conclusions

This Chapter has presented a detailed model for the formation and growth of vacancy aggregates in a Czochralski crystal. Technologically important microdefects, known as D-defects, have been investigated in detail by considering crystals grown under vacancy-dominated conditions. There is significant evidence that D-defects are essentially voids with oxygen playing a secondary role in their nucleation and growth rates. Several different techniques for characterizing these defects have emerged recently, and as a result, several apparently different defect species have been assumed to exist. Recent investigations of the defect response to operating conditions during growth have shown that the various defect types observed in the crystal region inside the OSF-ring are highly correlated to each other.

Based on the excellent quantitative agreement between the simulation results and the available experimental data we conclude that D-defects are in fact very likely to be vacancy aggregates, nucleated by high levels of vacancy supersaturation. Although the presence of oxygen in the voids has been confirmed, our results indicate that the role of oxygen is secondary in the nucleation of these voids, and that the process is nearly homogeneous. The results of this Chapter are important because they provide the first indication that *technologically useful* quantitative predictions can be made using simulation techniques. The simulations described in Chapters 5 and 6 show the extent to which microdefects in Czochralski silicon are based on the *native* point defect balance; i.e. self-interstitials and vacancies. This is important because the properties of these defects are relatively well understood as compared to those of

impurities.

Void formation is shown to begin at very high temperatures, to some extent before recombination is complete. However, most of the size distribution is determined in the later, cooler section of the crystal. The final size distribution of voids is a function of two features during crystal growth. The first of these is the nature of the balance between the point defect species near the melt/solid interface. Even though vacancies are everywhere in excess in the region inside the OSF-ring, the degree to which this dominance is present is a function of the thermal field (i.e the axial temperature gradient at the melt/solid interface (G)) and the pull rate (V). The results of Chapter 5 show that the relevant quantity in this case is the ratio V/G. Once the number of surviving vacancies is fixed, the crystal cooling rate in the temperature range, 1683 K - 1173 K, determines the final void size distribution. These results imply that the void size distribution is therefore a function of both V/G and $V < G >$ in different parts of the crystal.

The origin of the OSF-ring is postulated to be a consequence of very high free vacancy concentrations in a narrow annular region within the crystal. Although no simulations have been performed in the interstitial-rich region, it is known that vacancies are strongly depleted in regions outside the OSF-ring. The high vacancy concentration in the OSF-ring region arises paradoxically from weak vacancy domination during the earlier stages of crystal growth. Vacancies are only weakly supersaturated in the region close to the OSF-ring because interstitials are present in almost sufficient strength to balance the two species. This weak supersaturation results in very slow nucleation and aggregation and by the time that the crystal has cooled to the point where thermal processes are extinguished, a relatively high concentration of vacancies still exist as monomers.

These results show the feasibility and promise of microdefect modeling in silicon. We are now rapidly approaching a point where simulation will become an essential tool for crystal growers seeking to minimize the defect levels in their product without resorting to expensive and time consuming trial-and-error experimentation.

Chapter 7

Conclusions

The focus of this thesis has been the systematic investigation of microdefect formation in single crystal silicon during growth from the melt. Microdefects are one of the leading causes of microelectronic device degradation and their control is a pressing issue in the semiconductor industry. The building blocks of all crystalline irregularities are microscopic point defects which are necessarily present in all crystals in concentrations dictated by thermodynamics. The analysis in this thesis is based on a combination of atomistic simulation for computing thermophysical property data and continuum-scale numerical schemes for computing the distribution of microdefects across an entire crystal during its growth from the melt. An important component of this framework is the incorporation of experimental data for refining the thermophysical property estimates for defect diffusion and reaction rates.

The last three decades or so have witnessed a tremendous amount of work directed at evaluating the transport and thermodynamic properties of point defects. These efforts are divided into two general classes, theoretical and experimental/modeling. Theoretical calculations are usually based on atomistic simulation using a wide variety of representations for interatomic force fields, ranging from empirically derived potentials to detailed electronic structure calculations. The term "experimental/modeling" refers to experimental measurements of indirect phenomena which are subsequently related to the presence of point defects by a continuum scale model. These approaches usually require the injection of an impurity which diffuses through the silicon lattice

with the help of intrinsic point defects. A model is postulated for the interaction mechanism and the point defect properties fitted to the diffusion profile of the impurity. The fact that these point defect thermophysical properties are still uncertain is a testament to the difficulties inherent in both approaches. Solving for the electronic structure of any material is very computationally demanding and all modern calculations must include serious assumptions to make this task feasible. Empirical potentials attempt to lump complex electronic phenomena into simple analytical functions which also leads to great uncertainty. In the case of experimental measurement, the accuracy of the fitted parameters is sensitively dependent on a correct interpretation of the underlying physics in the system, such as boundary conditions and defect recombination mechanisms, which are generally poorly understood.

The goal of this thesis is not to compute thermophysical properties but to develop a quantitative, internally consistent picture of microdefect formation. In this context, internal consistency requires that a single set of thermophysical property estimates be applicable to a wide range of crystal growth conditions. The framework in this thesis does not rely entirely on either of the two approaches described above to provide thermophysical data for defect modeling. Atomistic simulation is used to provide an estimate for the temperature dependent portion of these properties which is very difficult to obtain by other means. The absolute values are treated as adjustable parameters and determined by requiring that our point defect model for OSF-ring dynamics reproduce a single set of experimental data. The usefulness of the resulting parameters is tested throughout the rest of the thesis work. The OSF-model is applied to a wide range of operating conditions without additional parameter fitting and is observed to predict well the location of the OSF-ring for different thermal conditions and crystal radii. The OSF-ring model is analyzed in detail using asymptotic analysis which leads to a closed form expression for the location of the OSF-ring in terms of the high temperature values of the point defect concentrations and diffusivities. This analysis confirms the validity of both the physics represented by the model and the estimates for the point defect properties.

Our model for OSF-ring dynamics does not require explicit consideration of the

microdefect size distribution; it links directly the location of the OSF-ring to the point defect concentrations. This allows us to greatly reduce the number of simultaneous fitting degrees-of-freedom. The next level of detail requires that we do consider defect aggregates explicitly. The aggregation model described in Chapter 6 shows that the point defect properties used in Chapter 5 give quantitatively reasonable results for the distribution of vacancy aggregates in the region inside the OSF-ring, *without additional fitting parameters.* This is a strong indication of the internal consistency of our point defect-based picture of microdefect formation during crystal growth. This mechanistic picture is summarized in detail in the following section.

In the context of future applications, we require that these estimates be transferable to models for other processes. For example, the point defect diffusion rates and equilibrium concentrations that are computed here, if essentially correct, should be useful for modeling defect evolution during wafer annealing treatments.

7.1 Summary

The results of this thesis are best summarized by considering a sequence of events that we believe take place starting at the melt/solid interface and evolve until the crystal cools sufficiently to freeze in all thermally activated processes. The first step in the defect history of the crystal is the incorporation of self-interstitials and vacancies across the melt/solid interface. We assume that these are incorporated in equilibrium, which results in slightly higher concentrations of vacancies at the melt/solid interface. Vacancies and self-interstitials recombine rapidly in this region because their concentrations become supersaturated due to crystal cooling. Point defect recombination is so rapid that one species is usually depleted leaving the other to develop a large supersaturation.

Whether interstitials, vacancies, or both survive depends entirely on the axial thermal gradient and the pull rate in the region near the melt/solid interface. Interstitials diffuse through the lattice faster than vacancies at high temperatures which, under certain circumstances, can offset the higher equilibrium concentration of va-

cancies. The model shows that these circumstances are dependent specifically on the ratio of the pull rate to the axial temperature gradient, V/G. The axial temperature gradient changes with radial position due to crystal cooling at the surface; for typical crystal growth conditions, the axial gradient will increase as the radius is traversed from the center of the crystal to the surface, which reduces V/G accordingly. Vacancy excess is predicted for regions where $V/G > V/G_{crit} = 1.34 \times 10^{-3}\ cm^2/minK$ and self-interstitial excess for regions where $V/G < V/G_{crit}$. The radial location where $V/G = V/G_{crit}$ exhibits a sensitive balance between interstitials and vacancies and marks the location of the OSF-ring in our model.

The validity of this simple picture is confirmed with an asymptotic analysis of the point defect balance equations. Leading order approximations to these equations show that the point defect concentrations decrease exponentially in a boundary layer adjacent to the melt/solid interface while satisfying the mass action condition, $C_I C_V = C_I^{eq} C_V^{eq}$. The thickness of the boundary layer is determined by the high activation energies of the thermally activated processes in the crystal, which decay rapidly as the crystal cools. One of the most important results of the asymptotic analysis is that radial transport scales out in the boundary layer region and allows for a quasi one-dimensional treatment. This result also rules out the role of the crystal surface in shaping OSF-ring dynamics in the bulk of the crystal. By considering the axial point defect fluxes in the crystal, a closed-form expression for the value of V/G at the location of the OSF-ring is developed based on the leading-order solution for the point defect profiles and the values of the point defect properties at the melt/solid interface. The prediction for the value of V/G is seen to agree very well with the empirically determined value.

The cessation of the point defect recombination reaction due to depletion of the limiting species allows the excess species to develop high levels of supersaturation. The homogeneous nucleation model described in Chapter 6 shows that these supersaturations rapidly cause aggregation into large clusters containing up to 10^9 point defects. The size distribution of these clusters is a function of two main factors. The first is determined by the residual point defect concentration which scales according

to V/G at the melt/solid interface as discussed above. In regions where V/G is very different from the critical value, the surviving defect species is left in large excess and nucleation begins at relatively high temperatures. Conversely, near the predicted OSF-ring position, the remaining species barely survives and supersaturation due to crystal cooling builds up more slowly, delaying the onset of nucleation and cluster growth.

The second factor is related to the thermal history in the remaining portion of the crystal. A cluster size distribution will generally evolve until either kinetic or thermodynamic limits are met. For unlimited time, a certain equilibrium distribution determined by the minimum free energy will result and no further evolution will take place. This state is never achieved in the case of crystal growth because a kinetic limit is always reached first due to crystal cooling and the consequent freezing of all thermally activated processes. Thus, the crystal cooling rate, $V < G >$ determines the final cluster size distribution. In this context, $< G >$ is a measure of the characteristic gradient in the relevant temperature range. The final microdefect distribution is therefore a function of *both* (V/G) and $V < G >$ in different parts of the crystal. The former is relevant near the melt/solid interface while the latter is the appropriate quantity in the rest of the crystal. Figure 7-1 summarizes these processes in schematic form for the case of a vacancy rich crystal.

These results consolidate phenomena that were previously assumed to be disparate. It seems very plausible that native point defects do indeed control the formation of many microdefect species. The results presented in this thesis are important because they will allow for quantitative prediction of microdefect concentrations as a function of crystal growth conditions. As shown in Chapter 5, it is possible to simulate actual crystal growth systems which is of tremendous potential value for the crystal growth industry. Taking an entire CZ system off-line and growing experimental crystals which are useless commercially is extremely expensive for these companies. Simulation provides a very cheap alternative for designing optimal thermal environments and the results shown here show that this goal is certainly achievable.

It is expected that the results of the model presented in this thesis are equally

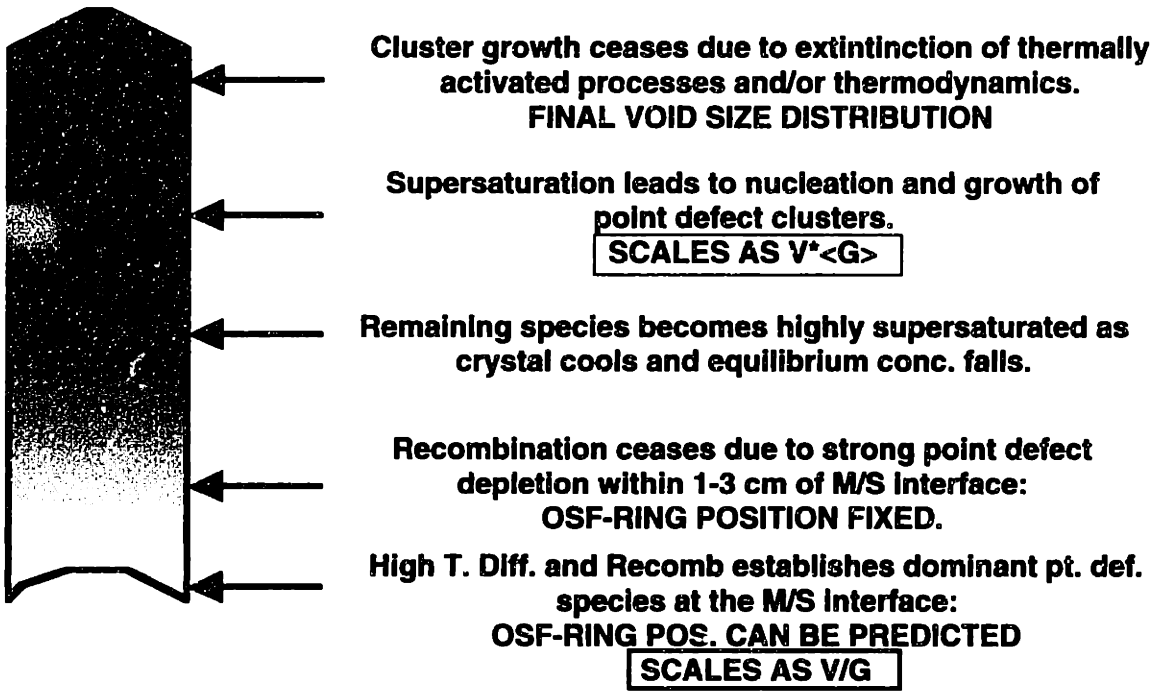

Figure 7-1: Schematic representation of the fundamental stages of microdefect size distribution evolution during crystal growth.

applicable to the case of FZ silicon which is chemically much purer than CZ silicon. There is still a need to confirm that D-defects are the same in both CZ and FZ crystals despite the different chemical environments. This link will provide strong proof that native point defects indeed are the critical species in the formation of most microdefects. Characterization of the neutral zone in FZ crystals in a similar fashion to the study described in Chapter 5 will confirm that the neutral zone and the OSF-ring represent similar regions in terms of native point defect balance. These studies will provide strong evidence that native point defects are the primary driving force for microdefect formation.

7.2 Directions for Future Research

Despite the apparent maturity of the silicon industry, we are only beginning to understand the fundamental processes that are responsible for defect formation during growth and subsequent wafer processing. While the emphasis of this thesis has been placed exclusively on the crystal growth process itself, the bulk of a wafer's thermochemical exposure occurs after growth. The following discussion of future research is therefore geared towards both aspects of wafer manufacture.

7.2.1 Two Dimensional Cluster Simulation with Interstitial-Type Microdefects and Oxygen

The fact that current crystal growers produce vacancy rich material exclusively because of the higher allowable pull rates does not mean that interstitial aggregates are not of interest. Currently, these tend to be so large that they generally lead to complete device failure. However, their concentrations are typically lower than those of D-defects and certain electronic properties are actually better in the interstitial region than they are in vacancy dominated areas (Dornberger, 1997). For these reasons, the current model should be extended to account for the formation of interstitial related microdefects.

The model presented in Chapter 6 does allow for interstitial aggregation but does not consider the fairly strong evidence that these microdefects undergo a transition from loosely bound three dimensional structures to planar geometries, such as dislocation loops (Föll and Kolbesen, 1975). Free energy descriptions for both types of structures are needed in order to obtain a complete picture for interstitial aggregation. The formation of interstitial-type perfect dislocation loops has been studied extensively in the metal irradiation literature, where radiation damage in nuclear reactor wall leads to large concentrations of these defects (Hall, 1980). Such information will be valuable in extending the present model to interstitial aggregation.

7.2.2 Time Dependent Simulations

One of the major sources of experimental data are transient experiments, which monitor the response of the microdefect distribution to rapid changes in the operating conditions. Among the most common of these experiments are pull rate ramp experiments, where the crystal growth rate is rapidly changed. If the change is fast enough, point defect diffusion becomes the limiting factor in the adjustment of the microdefect profiles and this can be used to further test the validity of the thermophysical properties used in this thesis. The effect of thermal history on microdefect formations is typically studied by detaching crystals from their melt at various points and then looking for sudden changes in the microdefect distribution. The location of these changes give much information about the temperatures at which the different defect formation processes occur (Takano *et al.*, 1995a).

The development of a time dependent simulation is also useful in the study of post-growth thermochemical anneals. This is a very active area, with many existing codes such as SUPREM (Fahey *et al.*, 1989). However, most of these codes have been developed based on the assumption that thermophysical properties would be specified by the individual user, usually by fitting the simulation results to a particular set of experiments. This has resulted in a huge number of widely varying parameter sets which only reproduce very specific data. One of the problems of current research directed at the simulation of wafer annealing is that the defect profile in the pre-

annealed wafer is not fully understood. Using an incorrect set of initial conditions will always lead to an incorrect fitting of the relevant thermophysical properties.

The application of what we have learned so far is a good start for investigating the more difficult problem of wafer annealing. Wafer annealing is usually performed at lower temperatures where electronic effects are potentially important. Furthermore, the role of impurities, especially oxygen, and electronic dopants such as boron, is more important in this case. The goal of research into this area would be to see whether the same thermophysical parameters used in the crystal growth simulations can be used in wafer processing models. If so, only impurity thermophysical properties would be fitted to the new experimental data, reducing the likelihood of large errors.

7.2.3 Atomistic Simulation of Large Defects

Atomistic simulation is clearly a viable technique for obtaining detailed structural and thermodynamic information for point defects and their aggregates. Empirical potentials such as the Stillinger-Weber potential are sufficiently accurate to give useful structural and semi-quantitative information for large defects. The ability to fully relax large aggregates will require further optimization of the approach described in Chapter 4 of this thesis. The simulation presented there was implemented on a fairly small number of processors; the use of a larger number of processors is the most obvious way to improve the performance of the code.

The transfer of information from the microscale to the continuum scale has traditionally consisted of computing transport and kinetic coefficients from atomic models and then using these coefficients in continuum models. The ultimate goal of true multiscale modeling is to be able to consider directly the mesoscale region in between these limits. The development of a hybrid molecular dynamics (or Monte Carlo) atomistic simulation in which regions of perfect crystal are treated with a mean-field type approximation will allow for more realistic concentrations of point defects and aggregates to be considered. One of the major obstacles to obtaining real kinetic data from atomistic simulations is the fact that artificially high concentrations must be considered to keep the size of the system as small as possible. In areas that are

sufficiently far away from defects there is no real need to compute forces and trajectories explicitly. This is especially true for short ranged potentials such as the Stillinger Weber potential which tend to predict short ranged strain fields in the region around defects. Such an approach will remove the diluteness constraint for these simulations and provide real time kinetic data, which has tremendous potential value for the kinds of computations described in this thesis.

Bibliography

Abe, T., *VLSI Electronics: Microstructure Science*, eds. N. G. Einspruch and H. Huff, Academic, New York, vol. 12, (1985).

Abe, T. and Harada, H., *Defects in Semiconductors II*, ed. S. Mahajan and J. W. Corbett, North-Holland, New York, (1983a).

Abe, T. and Harada, H., *Mat. Res. Soc. Symp Proc.*, **14** (1983b), 1.

Abe, T. and Takeno, H., *Mat. Res. Soc. Symp Proc.*, **262** (1992), 3.

Allen, C. E., Ditchfield, R., and Seebauer, E. G., *Phys. Rev. B*, **55** (1997), 13304.

Allen, M. P. and Tildesley, D. J., *Computer Simulations of Liquids*, Oxford Science Publications, Oxford Press, (1992).

Andersen, H. C., *J. Chem. Phys.*, **72** (1980), 2384.

Antoniadis, D. A. and Moskowitz, I., *J. Appl. Phys.*, **53** (1982), 9214.

Aris, R., *The Mathematical Theory of Diffusion and Reaction in Permeable Catalysts*, Clarendon Press, Oxford, (1975).

Ashcroft, N. W. and Mermin, D. N., *Solid State Physics*, Saunders College, Philadelphia, (1979).

Balamane, H., Halicioglu, T., and Tiller, W. A., *Phys. Rev. B*, **46** (1992), 2250.

Baraff, G. A. and Schluter, M., *Phys. Rev. B*, **30** (1984), 3460.

Baskes, M. I., Nelson, J. S., and Wright, A. F., *Phys. Rev. B*, **40** (1989), 6085.

Becker, R. and Doring, W., *Ann. Phys.*, **24** (1935), 719.

Benedek, R., Yang, L. H., Woodward, C., and Min, B. I., *Phys. Rev. B*, **45** (1992), 2607.

Billig, E., *Proc. R. Soc. A.*, **235** (1956), 37.

Biswas, R. and Hamann, D. R., *Phys. Rev. Lett.*, **55** (1985), 2001.

Blöchl, P. E., Smargiassi, E., Car, R., Laks, D. B., Andreani, W., and Pantelides, S. T., *Phys. Rev. Lett.*, **70** (1993), 2435.

Bokshtein, B. S., Bokshtein, S. Z., and Zhukhovitskii, A. A., *Thermodynamics and Kinetics of Diffusion in Solids*, Oxonian Press Pvt. Ltd., New Delhi, (1985).

Bolding, B. C. and Andersen, H. C., *Phys. Rev. B*, **41** (1990), 10568.

Bornside, D. E., Kinney, T. A., and Brown, R. A., *Int. J. Num. Meth. Eng.*, **30** (1990), 133.

Brabec, T., Schrems, M., Budil, M., Poetzl, H. W., Kuhnert, W., Pongratz, P., Stingeder, G., and Grasserbauer, M., *J. Electrochem. Soc.*, **136** (1989), 1542.

Bracht, H., Stolwijk, N. A., and Mehrer, H., *Phys. Rev. B*, **52** (1995), 16542.

Brice, J. C., *The Growth of Crystals from Liquids*, North-Holland/American Elsevier, New York, (1973).

Broughton, J. Q. and Li, X. P., *Phys. Rev. B*, **35** (1987), 9120.

Brown, R. A., *AIChE J.*, **34** (1988), 881.

Brown, R. A., Maroudas, D., and Sinno, T., *Journal of Crystal Growth*, **137** (1994), 12.

Bulatov, V. V., Yip, S., and Argon, A. S., *Phil. Mag. A.*, **72** (1995), 453.

Burton, J. J. and Jura, G., *J. Phys. Chem. Solids*, **28** (1967), 705.

Car, R. and Parinello, M., *Phys. Rev. Lett.*, **55** (1985), 2471.

Carlsson, A. E., *Solid State Physics: Advances in Research and Applications*, edited by Ehrenreich and D. Turnbull, Academic, New York, (1990).

Chaney, R. E. and Varker, C. J., *J. Cryst. Growth.*, **33** (1976), 188.

Chang, J. S. and Cooper, G., *J. Comp. Phys.*, **6** (1970), 1.

Chelikowsky, J. R. and Phillips, J. C., *Phys. Rev. B*, **41** (1990), 5735.

Chikawa, J. and Shirai, S., *J. Crystal Growth*, **39** (1977), 328.

Dannefaer, S., Mascher, P., and Kerr, D., *Phys. Rev. Lett.*, **56** (1986), 2195.

Dash, W. C., *J. Appl. Phys.*, **29** (1958), 736.

deGroot, S. R. and Mazur, P., *Non-Equilibirum Thermodynamics*, Dover Publications, Inc., New York, (1984).

deKock, A. J. R., *Appl. Phys. Lett.*, **16** (1970), 100.

deKock, A. J. R. and Wijgert, *J. Crystal Growth*, **49** (1980), 718.

Diaz de la Rubia, T., Caturla, M. J., and Tobin, M., *Microstructure of Irradiated Materials.*, Mater. Res. Soc.; Pittsburgh, PA, (1996).

Dodson, B. W., *Phys Rev. B*, **35** (1987), 2795.

Dongarra, J. J., *Computer Architecture News*, **20** (1992), 22.

Dornberger, E. (1997). *Prediction of Oxidation Induced Stacking Fault Ring and Grown-in Defects in Czochralski Silicon Crystals.* PhD thesis, U. de Louvain.

Dornberger, E., Esfandyari, J., Graf, D., Vanhellemont, J., Lambert, U., Dupret, F., and von Ammon, W., *Electrochem. Soc. Proc.*, **97** (1996a), 40.

Dornberger, E. and von Ammon, W., *J. Electrochem. Soc.*, **143** (1996), 1648.

Dornberger, E., von Ammon, W., Graf, D., Lambert, U., Miller, A., Oelkrug, H., and Ehlert, A., *Electrochem. Soc. PV*, **96** (1996b), 140.

Dupret, F., Nicodeme, P., Ryckmans, Y., Wouters, P., and Crochet, M. J., *Int. J. Heat Mass Trans.*, **33** (1990), 1849.

Esfandyari, J., Schmeiser, C., Senkader, S., Hobler, G., and Murphy, B., *J. Electrochem. Soc.*, **143** (1996), 995.

Esselink, K. and Hilbers, P. A. J., *Journal of Computational Physics*, **106** (1993), 108.

Esselink, K., Smit, B., and Hilbers, P. A. J., *Journal of Computational Physics*, **106** (1993), 101.

Ethier, S. and Lewis, L. J., *J. Mater. Res.*, **7** (1992), 2817.

Fahey, P. M., Griffin, P. B., and Plummer, J. D., *Rev. Mod. Phys.*, **61** (1989), 289.

Fincham, D., *Mol. Sim.*, **1** (1987), 1.

Fletcher, N. H., *J. Crystal Growth*, **35** (1976), 39.

Föll, H., Gösele, U., and O., K. B., *J. Crystal Growth*, **40** (1977), 90.

Föll, H., Gösele, U., and O., K. B., *J. Crystal Growth*, **52** (1981), 907.

Föll, H. and Kolbesen, B. O., *Appl. Phys.*, **8** (1975), 319.

Frank, F. C. and Turnbull, D., *Phys. Rev.*, **104** (1956), 617.

Frank, W., Gösele, U., Mehrer, H., and Seeger, A., *Diffusion in Crystalline Solids*, eds. Murch G. and Nowick, A. S., Academic, New York, (1984).

Gear, C. W., *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, (1971).

Ghandhi, S. K., *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, Wiley, New York, (1983).

Gillan, M. J., *J. Phys. Chem.*, **10** (1977), 1641.

Gilmer, G. H., Diaz de la Rubia, T., Stock, D. M., and Jaraiz, M., *Nuc. Inst. and Meth. Phys. Res. B*, **102** (1995), 247.

Girifalco, L. A., *Statistical Physics of Materials*, Wiley, New York, (1973).

Glassford, K. M., Chelikowsky, J. R., and Phillips, J. C., *Phys. Rev. B*, **43** (1991), 14557.

Gösele, U., Frank, W., and Seeger, A., *Appl. Phys.*, **23** (1980), 361.

Gösele, U. and Tan, T. Y., *Defects in Semiconductors II*, eds. Mahajan, S. and Corbett, J. W., North-Holland, New York, p. 45, (1983).

Goodwin, L., Skinner, A. J., and Pettifor, D. G., *Europhys. Lett.*, **9** (1989), 701.

Habu, R., Iwasaki, T., Harada, H., and Tomiura, A., *Jpn. J. Appl. Phys.*, **33** (1994), 1234.

Habu, R., Kojima, K. Harada, H., and Tomiura, A., *Jpn. J. Appl. Phys.*, **32** (1993a), 1747.

Habu, R. and Tomiura, A., *Jpn. J. Appl. Phys.*, **35** (1996), 1.

Habu, R., Yunoki, I., Saito, T., and Tomiura, A., *Jpn. J. Appl. Phys.*, **32** (1993b), 1740.

Haile, J. M., *Molecular Dynamics Simulation*, John Wiley & Sons, Inc., New York, (1992).

Hall, B., *J. Nuc. Materials*, **91** (1980), 63.

Hansen, J. P. and McDonald, I. R., *Theory of Simple Liquids*, Academic Press, New York, (1986).

Hasebe, M., Takeoka, Y., Shinoyama, S., and Naito, S., *Jpn. J. Appl. Phys.*, **28** (1989), L1999.

Hirsch, P. B., Howie, A., Nicholson, R. B., and Pashley, D. W., *Electron Microscopy of Thin Crystals*, Butterworths, London, (1965).

Hood, P., *Int. J. Num. Methods in Eng.*, **10** (1976), 379.

Hourai, M., Nagashima, T., Kajita, E., Miki, S., Shigematsu, T., and Okui, M., *J. Electrochem. Soc.*, **142** (1995), 3193.

Hu, S. M., *J. Appl. Phys.*, **45** (1974), 1567.

Hu, S. M., *J. Vac. Sci. Technol.*, **14** (1977), 17.

Hu, S. M., *J. Appl. Phys.*, **52** (1981), 3974.

Hu, S. M., *Mat. Sci. and Eng.*, **13** (1994), 105.

Huff, H. R. and Goodall, R. K., *Electrochem. Soc. Int.*, **5** (1996), 31.

Huff, H. R., Goodall, R. K., and Bhat, V., *Diffus. Defect Data B, Solid State Phenom.*, **47** (1995), 65.

Inoue, N., Osaka, J., and Wada, K., *J. Electrochem. Soc.*, **129** (1982), 2780.

Itsumi, M., Akiya, H., Ueki, T., Tomita, M., and Yamawaki, M., *J. Appl. Phys.*, **78** (1995), 5984.

Itsumi, M., Akiya, H., Ueki, T., Tomita, M., and Yamawaki, M., *Jpn. J. Appl. Phys.*, **35** (1996), 812.

Itsumi, M., Nakajima, O., and Shioni, N., *J. Appl. Phys.*, **72** (1992), 2185.

Izunome, K., Huang, X., Togawa, S., Terashima, K., and Kimura, S., *J. Crystal Growth*, **151** (1995), 291.

J., G. M. and Finnis, M. W., *J. Phys. Chem.*, **11** (1978), 4469.

Johnson, C., *Numerical Solutions of Partial Differential Equations by the Finite Elelement Method*, Cambridge University Press, Cambridge, (1987).

Johnson, R. A., *Phys. Rev.*, **174** (1968), 684.

Jones, C., Grout, P. J., and Lidiard, A. B., *Phil. Mag. Lett.*, **74** (1996), 217.

Kaiser, W. K. and Kech, P. H., *J. Appl. Phys.*, **28** (1957), 882.

Kakimoto, K., Yi, K., and Eguchi, M., *J. Crystal Growth*, **163** (1996), 238.

Kapila, A. K., *Asymptotic Treatment of Chemically Reacting Systems*, Pitman, London, (1983).

Katz, J. and Wiedersich, H., *J. Chem. Phys.*, **55** (1971), 1414.

Keck, P. H., van Horn, W., Soled, J., and MacDonald, A., *Rev. Sci. Instrum.*, **25** (1954), 331.

Keller, W. and Muhlbauer, A., *Floating-Zone Silicon*, Marcel Dekker, Inc., New York., (1981).

Kelly, P. J. and Car, R., *Phys. Rev. B*, **45** (1992), 6543.

Kern, W. and Puotinen, D. A., *RCA Rev.*, **31** (1970), 207.

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

Kohyama, M. and Takeda, S., *Phys. Rev. B*, **46** (1992), 12305.

Lawrence, J. E. and Huff, H. R., *In 'VLSI Electronics Microsctructure Science*, ed. N. G. Einspruch, pp. 51-102., Academic Press, New York, (1982).

Li, L. H. and Lowther, J. E., *Phys. Rev. B.*, **53** (1996), 11277.

Lifshitz, I. M. and Slyozov, V. V., *J. Chem. Phys. Solids*, **19** (1961), 35.

Lin, W. and Benson, K. E., *Ann. Rev. Mater. Sci.*, **17** (1987), 273.

Lutsko, J. F., Wolf, D., and S., Y., *J. Chem. Phys.*, **88** (1988), 6525.

Maroudas, D. (1992). *Dynamics of Point and Line Defects in Single Semiconductor Crystals Grown from the Melt.* PhD thesis, Massachusetts Institute of Technology.

Maroudas, D. and Brown, R. A., *Appl. Phys. Lett.*, **62** (1993), 172.

Matthews, J. W. and Van Vechten, J. A., *J. Crystal Growth*, **35** (1976), 343.

Matthews, M. D. and Ashby, S. J., *Phil. Mag.*, **27** (1973), 1313.

McQuarrie, D. A., *Statistical Mechanics*, Harper Collins, New York, (1976).

Mel'čuk, A. I., Giles, R. C., and Gould, H., *Computers in Physics*, **5** (1991), 311.

Mercer, J. L. and Chou, M. Y., *Phys. Rev. B*, **47** (1993), 9366.

Mizuo, S. and Higuchi, H., *J. Electrochem Soc.*, **129** (1982), 2292.

Morehead, F., *Defects in Electronic Materials*, eds. Stavola, M., Pearton, S. J., and Davies, G., MRS, Pittsburgh, (1988).

Morita, E., Ryuta, J., Tanaka, T., and Shimanuki, Y., *Mat. Res. Soc. Symp.*, **259** (1992), 182.

Moriya, K., *J. Crystal Growth*, **94** (1989), 182.

Moriya, K. and Ogawa, T., *Jap. J. Appl. Phys.*, **22** (1983), 307.

Nakajima, K., Furukawa, J., Furuya, H., and Shingyouji, T., *Semicond. Silicon 1994 Electrochem. Soc. Proc.*, **94** (1994), 168.

Nakamura, K., Saisyoji, T., Kubota, T., Iiada, T., Tomioka, J., and Nishimura, M., *Proc 2nd Int. Symp. Adv. Sci. and Tech. of Silicon Mat.*, **2** (1996), 282.

Nakano, A., Bi, L., Kalia, R. K., and Vashishta, P., *Phys. Rev. B*, **49** (1994), 9441.

Nichols, C. S., Van de Walle, C. G., and Pantelides, S. T., *Phys. Rev. B*, **40** (1989), 5484.

Pantelides, S. T., *New Dev. in Semicond. Physics. Proc. of the Third Summer School*, **1111** (1988), 143.

Park, B. D., Ahn, J., and Chang, G. S., *Sae Mulli*, **35** (1996), 742.

Park, B. T. and Petrosian, V., *Astrophys. J. Suppl. Ser.*, **103** (1996), 255.

Park, H. and Law, M. E., *J. Appl. Phys.*, **72** (1992), 3431.

Pearson, E. M., Takai, T., Halicioglu, T., and Tiller, W. A., *J. Crystal Growth*, **70** (1984), 33.

Petroff, P. M. and de Kock, A. J. R., *J. Crystal Growth*, **30** (1975), 117.

Phillpot, S. R. and Rickman, J. M., *J. Chem. Phys.*, **94** (1991), 1454.

Plaskett, T. S., *Trans. Met. Soc. AIME*, **233** (1965), 809.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., *Numerical Recipes*, Cambridge University Press, Cambridge, (1989).

Rasband, P. B., Clancy, P., and Thompson, M. O., *J. Appl. Phys.*, **79** (1996a), 8998.

Rasband, P. B., Horsfield, A. P., and Clancy, P., *Phil. Mag. B*, **73** (1996b), 71.

Ravi, K. V., *Imperfections and Impurities in Semiconductor Silicon*, John Wiley and Sons, New York, (1981).

Risken, H., *The Fokker-Planck Equation: methods of solution and applications*, Springer-Verlag, Berlin, (1984).

Roksnoer, R. J. and van den Boom, M. M. B., *J. Crystal Growth*, **53** (1981), 563.

Rozgonyi, G. A. and Pearce, C. W., *Appl. Phys. Lett.*, **32** (1978), 747.

Ryuta, J., Morita, E., Tanaka, T., and Shimanuki, Y., *Jpn. J. Appl.Phys.*, **29** (1990), 1947.

Sadamitsu, S., Umeno, S., Koike, Y., Hourai, M., Sumita, S., and Shigematsu, T., *Jpn. J. Appl. Phys*, **32** (1993), 3675.

Schaal, B. (1983). *Ein Modell des Einfrierens von Eigenfehlstellen bei der Kristallzucht.* PhD thesis, University of Stuttgart.

Seeger, A. and Chik, K. P., *Phys. Stat. Solidi*, **29** (1968), 455.

Seeger, A., Foll, H., and Frank, W., *Radiation Effects in Semiconductors 1975*, eds. N. B. Urli and J. W. Corbett, IOP, Bristol, (1977).

Shimura, F., *Semiconductor Silicon Crystal Technology*, Academic Press, Inc. San Diego, CA, (1989).

Shimura, F. and Huff, H. R., *VLSI Handbook*, ed. N. G. Einspruch, Academic Press, New York, NY, (1985).

Sinno, T., Jiang, Z. K., and Brown, R. A., *Appl. Phys. Lett.*, **68** (1995), 3028.

Song, E. G., Kim, E., Lee, Y. H., and Hwang, G. Y., *Phys. Rev. B*, **48** (1993), 1486.

Stillinger, F. H. and Weber, T. A., *Phys. Rev. B*, **31** (1985), 5262.

Stolwijk, N. A., Hölzl, J., Frank, W., Weber, E. R., and Mehrer, H., *Appl. Phys. A*, **39** (1986), 37.

Sueoka, K., Ikeda, N., and Yamamoto, T., *Appl. Phys. Lett.*, **65** (1994), 1686.

Takano, K., Kitagawa, K., Kimura, M., and Yamagishi, H., *Mat. Sci. Forum.*, **196** (1995a), 1707.

Takano, K., Kitagawa, K., Nagashima, T., Fujiwara, H., Sadamitsu, S., and Miki, S., *Defects in Semiconductors*, eds. M. Suezawa and H. Katayama, Yoshida, Japan, (1995b).

Tan, T. Y., Gardner, E. E., and Tice, W. K., *Appl. Phys. Lett.*, **30** (1977), 175.

Tan, T. Y. and Gösele, U., *Appl. Phys. A*, **37** (1985), 1.

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

Teal, G. K. and Little, J. B., *Phys. Rev.*, **78** (1950), 647.

Tersoff, J., *Phys. Rev. B*, **37** (1988a), 6991.

Tersoff, J., *Phys. Rev. B*, **38** (1988b), 9902.

Togawa, S., Shiraishi, Y., Terashima, K., and Kimura, S., *J. Electrochem. Soc.*, **142** (1995), 2844.

Ungar, P. J., Takai, T., Halicioglu, T., and Tiller, W. A., *J. Vac. Sci. Technol.*, **11** (1993), 224.

Van Vechten, J. A., *Appl. Phys. Lett.*, **26** (1975), 593.

Van Vechten, J. A., *Phys. Rev B*, **17** (1978), 3197.

Verlet, L., *Phys. Rev.*, **159** (1967), 98.

Vineyard, G. H., *Phys. Chem. Solids*, **3** (1957), 121.

Virkkunen, R., Alatalo, M., and Puska, *Comp. Mat. Sci.*, **1** (1993), 151.

Volmer, M. and Weber, A., *Z. Physik. Chem.*, **119** (1925), 277.

von Ammon, W., *Solid State Phen.*, **47** (1996), 97.

von Ammon, W., Dornberger, E., Oelkrug, H., and Weidner, H., *J. Crystal Growth*, **151** (1995), 273.

von Ammon, W., Dreier, P., Hensel, W., Lambert, U., and Köster, L., *Mat. Sci. and Eng. B*, **36** (1996), 33.

von Gottberg, F. K. (1997). *Stochastic Dynamic Simulations of Self-Assembly.* PhD thesis, Massachusetts Institute of Technology.

Voronkov, V. V., *J. Crystal Growth*, **59** (1982), 625.

Vysotskaya, V. V., Gorin, S. N., and Sidorov, Y. A., *Neorgan. Mat.*, **26** (1990), 458.

Wada, K., Nakanishi, H., Takaoka, H., and Inoue, N., *J. Crystal Growth*, **57** (1982), 535.

Waite, T. R., *Phys. Rev.*, **107** (1957), 463.

Wang, C. Z., Chan, C. T., and Ho, K. M., *Phys. Rev. Lett.*, **66** (1991), 189.

Watkins, G. D., Troxell, J. R., and Chatterjee, A. P., in *Defects and Radiation Effects in Semiconductors, Inst. of Phys. Conf. Ser.*, Ed. J. H. Albany, **46** (1979), 16.

Webb, W. W., *J. Appl. Phys.*, **33** (1962), 1961.

Wijaranakula, W., *J. Appl. Phys.*, **67** (1990), 7624.

Wijaranakula, W., *J. Electrochem. Soc.*, **140** (1993), 3306.

Winkler, R. and Sano, M., *J. Electrochem. Soc.*, **141** (1994), 1398.

Wirtz, K., *Z. Phys.*, **44** (1943), 221.

Yacout, A. and Stubbins, J. F., *J. Nuc. Mat.*, **141** (1986), 677.

Yamagishi, H., Fusegawa, I., Fujimaki, N., and Katayama, M., *Semicond. Sci. Technol.*, **7** (1992), A135.

Yu, P. Y. and Cardona, M., *Fundamentals of Semiconductors*, Springer, Berlin, Germany, (1996).

Zachariah, M. R., Carrier, M. J., and Blaisten-Barojas, E., *J. Phys. Chem.*, **100** (1996), 14856.

Zhu, J., Yang, L. H., Mailhiot, C., Diaz de la Rubia, T., and Gilmer, G. H., *Nuc. Instr. and Meth. B*, **102** (1995), 29.

Zimmermann, H. and Ryssel, H., *Appl. Phys. A*, **55** (1992), 121.

Zulehner, W., *Semiconductor Silicon*, eds. G. C. Harbeke and M. J. Schulz, Springer-Verlag, Berlin, Germany, (1989).

Zulehner, W. and Huber, D., *Crystals 8: Silicon Chemical Etching*, ed. J. Grabmeier, Springer, Berlin, Germany, (1982).